

# CGEBox – a flexible and modular toolkit for CGE modelling with a GUI

---

- Wolfgang Britz, Institute for Food Resource Economics –  
- Version September 2026 -  
(matches rev. 7840 of <https://svn1.agp.uni-bonn.de/svn/cgebox>)

## Content

|        |                                                                          |                                           |
|--------|--------------------------------------------------------------------------|-------------------------------------------|
| 1.     | Introduction .....                                                       | 9                                         |
| 1.1    | How the documentation is organized .....                                 | 9                                         |
| 1.2    | Why an implementation in GAMS? .....                                     | 9                                         |
| 1.3    | Scalability, Modularity and Extensions as basic features of CGEBOX ..... | 10                                        |
| 2.     | Basic model equations .....                                              | 17                                        |
| 2.1    | Core sets .....                                                          | 17                                        |
| 2.2    | Reminder on CES and CET equations .....                                  | 19                                        |
| 2.3    | Overview on the supply side .....                                        | 22                                        |
| 2.4    | Production block .....                                                   | 24                                        |
| 2.5    | Factor markets .....                                                     | 32                                        |
| 2.6    | Income generation and distribution - overview .....                      | 37                                        |
| 2.7    | Income generation .....                                                  | 38                                        |
| 2.8    | Income distribution .....                                                | 42                                        |
| 2.8.1  | Summary .....                                                            | <b>Fehler! Textmarke nicht definiert.</b> |
| 2.8.2  | Intialization .....                                                      | <b>Fehler! Textmarke nicht definiert.</b> |
| 2.8.3  | Dynamic updating of share parameters .....                               | <b>Fehler! Textmarke nicht definiert.</b> |
| 2.8.4  | GUI .....                                                                | <b>Fehler! Textmarke nicht definiert.</b> |
| 2.9    | Household consumption .....                                              | 46                                        |
| 2.9.1  | CDE case .....                                                           | 46                                        |
| 2.9.2  | LES or CD case .....                                                     | 49                                        |
| 2.10   | Government consumption .....                                             | 51                                        |
| 2.10.1 | Summary .....                                                            | <b>Fehler! Textmarke nicht definiert.</b> |
| 2.10.2 | Intialization .....                                                      | <b>Fehler! Textmarke nicht definiert.</b> |
| 2.10.3 | Dynamic updating of share parameters .....                               | <b>Fehler! Textmarke nicht definiert.</b> |
| 2.10.4 | GUI .....                                                                | <b>Fehler! Textmarke nicht definiert.</b> |
| 2.11   | Investments and savings .....                                            | 56                                        |
| 2.11.1 | Global bank .....                                                        | 58                                        |

|        |                                                                 |     |
|--------|-----------------------------------------------------------------|-----|
| 2.11.2 | Fixed allocation of savings .....                               | 61  |
| 2.11.3 | Capital account and balance of payments.....                    | 61  |
| 2.12   | Demand sub-nests.....                                           | 62  |
| 2.12.1 | Motivation .....                                                | 62  |
| 2.12.2 | Implementation in model.....                                    | 63  |
| 2.12.3 | Benchmarking of demand sub-nests.....                           | 66  |
| 2.12.4 | References .....                                                | 70  |
| 2.13   | International trade and domestic sales, and related prices..... | 71  |
| 2.13.1 | Overview .....                                                  | 71  |
| 2.13.2 | Individual equations .....                                      | 72  |
| 2.13.3 | Melitz model.....                                               | 81  |
| 2.14   | Price indices .....                                             | 87  |
| 2.15   | List of main prices in model.....                               | 88  |
| 2.16   | Flexible nesting in the production function.....                | 90  |
| 3.     | Modules.....                                                    | 95  |
| 3.1    | Capital vintages .....                                          | 95  |
| 3.1.1  | Summary .....                                                   | 95  |
| 3.1.2  | Intialization.....                                              | 95  |
| 3.1.3  | Dynamic updating of share parameters .....                      | 97  |
| 3.1.4  | GUI.....                                                        | 98  |
| 3.2    | myGTAP module .....                                             | 98  |
| 3.2.1  | Summary .....                                                   | 98  |
| 3.2.2  | Motivation .....                                                | 98  |
| 3.2.3  | Implementation.....                                             | 99  |
| 3.2.4  | Government debt dynamics.....                                   | 107 |
| 3.2.5  | Methodology: Household demand.....                              | 112 |
| 3.2.6  | Graphical User Interface.....                                   | 118 |
| 3.2.7  | Code implementation.....                                        | 122 |
| 3.2.8  | Integration FAO household data into CGEBOX.....                 | 125 |
| 3.2.9  | Post-model micro-simulation .....                               | 138 |
| 3.3    | A GMIG module for CGEBox .....                                  | 146 |
| 3.3.1  | Background and motivation .....                                 | 146 |
| 3.3.2  | Equations.....                                                  | 146 |
| 3.3.3  | Technical aspects.....                                          | 149 |



|        |                                                                                                           |     |
|--------|-----------------------------------------------------------------------------------------------------------|-----|
| 3.3.4  | Result exploitation.....                                                                                  | 150 |
| 3.4    | Sub-national dis-aggregation of production and factor markets in CGEBox.....                              | 150 |
| 3.4.1  | Background and motivation .....                                                                           | 150 |
| 3.4.2  | Methodology .....                                                                                         | 150 |
| 3.4.3  | Integration into the modeling framework.....                                                              | 151 |
| 3.4.4  | Implementation for European countries at NUTS2 level .....                                                | 153 |
| 3.4.5  | Result exploitation.....                                                                                  | 155 |
| 3.5    | Armington with commitment terms .....                                                                     | 157 |
| 3.5.1  | Background .....                                                                                          | 157 |
| 3.5.2  | Benchmarking and GUI setting.....                                                                         | 157 |
| 3.5.3  | Equations.....                                                                                            | 160 |
| 3.6    | Spatial Price Equilibrium instead of Armington .....                                                      | 163 |
| 3.6.1  | Background .....                                                                                          | 163 |
| 3.6.2  | Equations and technical implementation.....                                                               | 164 |
| 3.6.3  | Interface.....                                                                                            | 166 |
| 3.7    | MRIO extension .....                                                                                      | 168 |
| 3.7.1  | Background .....                                                                                          | 168 |
| 3.7.2  | Definition of consistent split factors.....                                                               | 169 |
| 3.7.3  | Model equations .....                                                                                     | 173 |
| 3.7.4  | Graphical User Interface for MRIO extension .....                                                         | 177 |
| 3.8    | Partial equilibrium closure .....                                                                         | 179 |
| 3.9    | Tariff Line Module.....                                                                                   | 180 |
| 3.9.1  | Summary .....                                                                                             | 180 |
| 3.9.2  | Background .....                                                                                          | 180 |
| 3.9.3  | Approach .....                                                                                            | 181 |
| 3.9.4  | Model equations .....                                                                                     | 181 |
| 3.9.5  | Entering the data.....                                                                                    | 184 |
| 3.9.6  | Calibration .....                                                                                         | 185 |
| 3.9.7  | Post model processing.....                                                                                | 186 |
| 3.9.8  | Integration in the GUI .....                                                                              | 187 |
| 3.10   | CO <sub>2</sub> , Non-CO <sub>2</sub> Accounting and air pollution emissions, energy use accounting ..... | 189 |
| 3.10.1 | GHG emissions in CO <sub>2</sub> equivalents.....                                                         | 189 |
| 3.10.2 | Marginal abatement curves for GHG emissions.....                                                          | 192 |
| 3.10.3 | Air pollution emissions.....                                                                              | 207 |

|         |                                                                                  |     |
|---------|----------------------------------------------------------------------------------|-----|
| 3.10.4  | Reporting of energy use in Million tons of oil equivalents (MTOE) .....          | 208 |
| 3.11    | GTAP-E.....                                                                      | 208 |
| 3.12    | GTAP-AGR.....                                                                    | 208 |
| 3.13    | Recursive-dynamics .....                                                         | 211 |
| 3.13.1  | Overview .....                                                                   | 211 |
| 3.13.2  | Dynamic updates of share parameters in factor supply .....                       | 211 |
| 3.14    | G-RDEM: Long-term baseline generation and analysis.....                          | 212 |
| 3.14.1  | Summary .....                                                                    | 213 |
| 3.14.2  | Background and introduction .....                                                | 213 |
| 3.14.3  | Overview .....                                                                   | 214 |
| 3.14.4  | An AIDADS demand system with detail for food consumption.....                    | 216 |
| 3.14.5  | Differentiated productivity growth.....                                          | 227 |
| 3.14.6  | Endogenous saving rates .....                                                    | 231 |
| 3.14.7  | Debt accumulation from foreign savings.....                                      | 235 |
| 3.14.8  | Cost-share adjustment .....                                                      | 238 |
| 3.14.9  | Endogenous expenditure shares for government and investment demand .....         | 246 |
| 3.14.10 | Recovering crop yield projections.....                                           | 248 |
| 3.14.11 | Recovering crop land expansion .....                                             | 251 |
| 3.14.12 | GDP-dependent correction of calorie content .....                                | 257 |
| 3.14.13 | Steering of G-RDEM model runs.....                                               | 258 |
| 3.14.14 | Population and GDP growth from the IIASA SSPs .....                              | 259 |
| 3.14.15 | Using the G-RDEM generated baseline for counterfactual analysis .....            | 261 |
| 3.14.16 | Specific nestings in G-RDEM .....                                                | 263 |
| 3.14.17 | Endogenous non-Hicks neutral technical progress and primary factors supply ..... | 264 |
| 3.14.18 | Analysis of G-RDEM simulation output (exploitation) .....                        | 266 |
| 3.14.19 | Computing and data base considerations for G-RDEM runs .....                     | 267 |
| 3.15    | GTAP-Melitz: Heterogenous firm module.....                                       | 269 |
| 3.15.1  | Summary .....                                                                    | 269 |
| 3.15.2  | Introduction .....                                                               | 269 |
| 3.15.3  | Brief Literature overview .....                                                  | 271 |
| 3.15.4  | Implementation of Melitz model into CGEBox .....                                 | 273 |
| 3.15.5  | Numerical stability and domestic industry demand for its own output .....        | 285 |
| 3.15.6  | Technical implementation and an example application.....                         | 286 |
| 3.15.7  | Conclusion.....                                                                  | 291 |

|        |                                                                       |     |
|--------|-----------------------------------------------------------------------|-----|
| 3.15.8 | Acknowledgments .....                                                 | 292 |
| 3.16   | GTAP-AEZ .....                                                        | 293 |
| 3.16.1 | Summary .....                                                         | 293 |
| 3.16.2 | Model equations .....                                                 | 293 |
| 3.16.3 | Parameterization.....                                                 | 302 |
| 3.16.4 | Visualization of GTAP-AEZ results.....                                | 303 |
| 3.17   | Climate Change Damages .....                                          | 305 |
| 3.17.1 | Data input .....                                                      | 305 |
| 3.17.2 | Aggregation .....                                                     | 305 |
| 3.17.3 | Graphical User Interface.....                                         | 307 |
| 3.17.4 | Shock definition .....                                                | 308 |
| 3.17.5 | References .....                                                      | 311 |
| 4.     | Data base generation .....                                            | 312 |
| 4.1    | Loading a GTAP Data Base .....                                        | 313 |
| 4.1.1  | Background .....                                                      | 313 |
| 4.2    | Importing the data into CGEBox.....                                   | 313 |
| 4.3    | GTAPAGG: The data base generation utility of CGEBox.....              | 316 |
| 4.3.1  | Background .....                                                      | 316 |
| 4.3.2  | Overview .....                                                        | 317 |
| 4.3.3  | Defining your aggregation: The aggregator component.....              | 318 |
| 4.3.4  | Aggregation in GAMS .....                                             | 322 |
| 4.3.5  | Backward compatibility.....                                           | 323 |
| 4.3.6  | Data base treatment after aggregation: Filtering or not .....         | 323 |
| 4.3.7  | Simple deletion.....                                                  | 323 |
| 4.3.8  | Filtering and Rebalancing.....                                        | 324 |
| 4.3.9  | Rebooking of agri-food intermediate demand.....                       | 338 |
| 4.3.10 | Macro updates .....                                                   | 340 |
| 4.3.11 | Generalized RAS after building the SAM.....                           | 346 |
| 4.3.12 | Gender Differentiated labor data .....                                | 348 |
| 4.3.13 | Split utility.....                                                    | 349 |
| 4.3.14 | An evolved example for a split.....                                   | 351 |
| 4.3.15 | Balancing equations and objective function in the split problem.....  | 353 |
| 4.3.16 | Proportionality assumptions for cost structures during splitting..... | 359 |
| 4.3.17 | Preparation of the split factors.....                                 | 362 |

|        |                                                                              |     |
|--------|------------------------------------------------------------------------------|-----|
| 4.3.18 | Split solution process.....                                                  | 365 |
| 4.3.19 | Processing the solution from the split problem .....                         | 367 |
| 4.3.20 | MRIO split.....                                                              | 368 |
| 4.3.21 | Storing results from GTAPAgg as a new Data base for further processing ..... | 368 |
| 4.4    | Model size and solution behavior .....                                       | 370 |
| 4.5    | Inspecting the resulting data base .....                                     | 373 |
| 4.6    | Pre- or post-model aggregation .....                                         | 374 |
| 5.     | Scenario runs .....                                                          | 376 |
| 5.1    | Defining the scenario.....                                                   | 376 |
| 5.1.1  | Implementation of shock files in GAMS.....                                   | 376 |
| 5.1.2  | Scenario editor.....                                                         | 379 |
| 5.2    | Running the scenario .....                                                   | 383 |
| 5.3    | Output options .....                                                         | 386 |
| 5.4    | Model structure, parameterization and factor markets .....                   | 388 |
| 5.4.1  | Modules and parameters.....                                                  | 388 |
| 5.4.2  | Demand system .....                                                          | 389 |
| 5.5    | Factors .....                                                                | 389 |
| 5.6    | Parameters and indices .....                                                 | 390 |
| 5.6.1  | Indices .....                                                                | 390 |
| 5.6.2  | Scenario drivers .....                                                       | 390 |
| 5.6.1  | Parameters .....                                                             | 391 |
| 5.7    | Closures .....                                                               | 392 |
| 5.8    | Imperfect competition model .....                                            | 394 |
| 5.9    | Pre-configurations .....                                                     | 394 |
| 5.9.1  | Standard GTAP model .....                                                    | 395 |
| 5.9.2  | GTAPinGAMS.....                                                              | 395 |
| 5.9.3  | GLOBE.....                                                                   | 395 |
| 5.9.4  | MIRAGE .....                                                                 | 396 |
| 5.9.5  | ENVISAGE .....                                                               | 396 |
| 5.9.6  | CGEBOX.....                                                                  | 396 |
| 5.9.7  | CGEBOX+ .....                                                                | 396 |
| 5.10   | Options for solving the model: Algorithm .....                               | 396 |
| 5.10.1 | First-order Taylor approximation to define starting points .....             | 399 |
| 5.10.2 | Variable initialization before solving .....                                 | 401 |

|        |                                                                                |     |
|--------|--------------------------------------------------------------------------------|-----|
| 5.10.3 | Pre-solves .....                                                               | 401 |
| 5.11   | GAMS .....                                                                     | 407 |
| 6.     | Post-model processing.....                                                     | 409 |
| 6.1    | Decomposing welfare changes in CGEBox .....                                    | 409 |
| 6.1.1  | Background .....                                                               | 409 |
| 6.1.2  | Decomposition with regard to overall price indices and income components ..... | 409 |
| 6.1.3  | Decomposition with a focus on prices in changes.....                           | 411 |
| 6.1.4  | Demand nests .....                                                             | 411 |
| 6.1.5  | Private household .....                                                        | 411 |
| 6.1.6  | Government and investments .....                                               | 413 |
| 6.1.7  | Utility change for the regional household and EV .....                         | 413 |
| 6.1.8  | Decomposition following Huff and Hertel, 2000 .....                            | 415 |
| 6.2    | Trade in VA indicators .....                                                   | 418 |
| 6.2.1  | Background .....                                                               | 418 |
| 6.2.2  | Post-Model processing .....                                                    | 419 |
| 6.2.3  | GUI.....                                                                       | 420 |
| 6.3    | Altax.....                                                                     | 422 |
| 6.3.1  | Background .....                                                               | 422 |
| 6.3.2  | Technical realization .....                                                    | 422 |
| 6.3.3  | Pre-configuration.....                                                         | 423 |
| 6.3.4  | Post-Model processing .....                                                    | 423 |
| 6.3.5  | Use of the updated benchmark .....                                             | 424 |
| 6.4    | Using FAOSTAT data for splits.....                                             | 425 |
| 6.4.1  | Introduction .....                                                             | 425 |
| 7.     | GUI and installation .....                                                     | 426 |
| 7.1    | Software requirements and access to the code .....                             | 426 |
| 7.2    | Some comments for runGTAP and GEMPACK users.....                               | 426 |
| 7.3    | Background .....                                                               | 427 |
| 7.4    | SVN checkout.....                                                              | 428 |
| 7.5    | Starting the GUI .....                                                         | 431 |
| 7.6    | First steps.....                                                               | 433 |
| 7.7    | Making the interface working with GAMS.....                                    | 435 |
| 7.8    | Inspecting the results .....                                                   | 439 |
| 7.8.1  | Notes for GEMPACK users or users familiar with other CGE models .....          | 441 |

|        |                                                                                    |     |
|--------|------------------------------------------------------------------------------------|-----|
| 7.8.2  | Relative difference for all variables as in runGTAP .....                          | 442 |
| 7.8.3  | Trouble shooting with the viewer.....                                              | 444 |
| 7.8.4  | Overview on existing views for exploitation.....                                   | 445 |
| 8.     | Sensitivity analysis .....                                                         | 455 |
| 8.1    | Controls for sensitivity experiments.....                                          | 457 |
| 8.2    | Parallel solves and analyzing the results .....                                    | 461 |
| 9.     | Getting an overview on the code .....                                              | 466 |
| 9.1    | HTML documentation .....                                                           | 466 |
| 9.2    | Equation and variable viewer (similar AnalyzeGE).....                              | 468 |
| 10.    | Technical documentation.....                                                       | 471 |
| 10.1   | Overview on directory structure .....                                              | 471 |
| 10.2   | GUI overview .....                                                                 | 471 |
| 10.3   | Processes .....                                                                    | 472 |
| 10.4   | Post-model processing.....                                                         | 473 |
| 10.5   | Substitution of variables based on macros .....                                    | 474 |
| 10.6   | Exploitation and flexible aggregation .....                                        | 476 |
| 10.7   | Main programs when solving a shock and integration of modules .....                | 480 |
| 10.8   | Tests and checks .....                                                             | 481 |
| 10.8.1 | Alerting users of compile and run-time errors.....                                 | 481 |
| 10.8.2 | Data base checks.....                                                              | 482 |
| 10.8.3 | Checks built into the equation system of the model and shock files for tests ..... | 482 |
| 10.8.4 | Checks for negative results and exhaustion.....                                    | 483 |
| 10.8.5 | Test suite.....                                                                    | 483 |

## 1. Introduction

### 1.1 How the documentation is organized

This document

1. describes the methodology of CGEBox simulation model, both of the implementation of the GTAP standard model and of the different optional modules, as well as of the steps to prepare the data
2. provides its technical documentation
3. details how to use the Graphical User Interface (GUI) and to introduce specific data or shocks.

The GTAP standard model and each module each have their own sections, these sections start with an introduction to the methodology, next detail the equations and the implementation in GAMS, followed (where applicable) by sub-sections on how to introduce additional data and on GUI choices and result exploitation.

Users who want to start using the basic model without having a look at details are invited to read the following two sections “Why an implementation in GAMS?” and “Scalability, Modularity and Extensions as basic features of CGEBOX” up to page 17 as an introduction, skip the following detailed sections first and continue with the section “GUI and installation”, page 426 ff.

### 1.2 Why an implementation in GAMS?

[GTAP](#) is certainly the far most widespread international data base for CGE (Computable General Equilibrium) modeling. The data base was for years only distributed with a default CGE model template in [GEMPACK](#), called the [standard GTAP model](#), as well as with many useful GEMPACK based utilities. GEMPACK as a special package for CGE modelling has certainly distinct advantages compared to less specialized Algebraic Modeling Languages<sup>1</sup>. Some users might however prefer [GAMS](#) when working with GTAP. Some features of GAMS not available with GEMPACK are used in CGEBox, such as being able to solve (several) constrained non-linear equation systems not written in their FOC in one program call, for instance to filter small entries from the global SAM and rebalancing it afterwards, see section “Filtering and Rebalancing”, or when applying the split utility, see section “Split utility”, or when using the model in recursive-dynamic simulations.

Indeed, outside the CGE community, GAMS seems much more common than GEMPACK as an Algebraic Modelling Language (AML), e.g. in Agricultural Economics (Britz and Kallrath 2012<sup>2</sup>). In 2015, the Center for Global Trade Analysis (GTAP) decided to release in parallel to the standard GTAP model in GEMPACK a version coded in GAMS, now available for GTAP Standard Version 7 (van der Mensbrugghe 2018<sup>3</sup>). Intensive testing showed that both versions produce the same results. The GAMS equations of that core model plus some extensions are discussed below in the section

---

<sup>1</sup> Horridge, M., Meeraus, A., Pearson, K., & Rutherford, T. F. (2013). “Solution software for computable general equilibrium modelling”. In: Handbook of computable general equilibrium modeling (Vol. 1, pp. 1331-1381). Elsevier

<sup>2</sup> Britz, W., Kallrath, J. (2012): Economic Simulation Models in Agricultural Economics: The Current and Possible Future Role of Algebraic Modeling Languages, in: Kallrath, J. (eds.): Algebraic Modelling Systems: Modeling and Solving Real World Optimization Problems: 199-212, Springer, Heidelberg, Germany

<sup>3</sup> Van der Mensbrugghe, D. (2018). “The Standard GTAP Model in GAMS”, Version 7. *Journal of Global Economic Analysis*, 3(1), 1-83

“Basic model equations”. These equations are mostly structurally identical to van der Mensbrugghe 2018 to which this documentation refers where applicable, while mnemonics and coding style are to a large degree harmonized as well. Further CGE models realized in GAMS are widely available. They share basic elements of the standard GTAP model while adding extensions in several directions, for instance [GTAPinGAMS](#) by Tom Rutherford<sup>4</sup>, [GLOBE](#) by Scott McDonald and Karen Thierfelder or [ENVISAGE](#) by Dominique van der Mensbrugghe. However, these models do not provide an exact replication of the standard GTAP Model.

The newly released replica of the standard GTAP model version 7 in GAMS comprised in CGEBox is not only interesting for research, but also provides together with the GTAP Data Base an excellent starting point for teaching or own projects of students, especially in the context of course programs already using GAMS. CGEBox offers additionally some extensions compared to the standard GTAP model as discussed next, inspired by features of existing CGE models, and adds modules such as *myGTAP*, *GMIG*, *GTAP-AEZ*, *GTAP-AGR*, *GTAP-E* or *GTAP-Melitz* from the GTAP family of models. These modules are detailed in dedicated sections of this documentation, they are not exact replica of related GTAP models. Furthermore, CGEBox offers a choice for certain structural elements such as with regard to the functional form in final demand or the aggregation of the Armington agents. The combination of the extensions and modules along with such flexibility allow adjusting the model’s structure to a specific application or to replicate main layout features of well-known models (see Britz and van der Mensbrugghe 2018<sup>5</sup>).

The model documentation is structured as follows. The following section “Scalability, Modularity and Extensions as basic features of CGEBOX” discusses the main features of the modeling platform including a brief overview on the different modules. The structure of the core model is discussed equation by equation in the section “Basic model equations” (page 17ff). The following section “Modules” (pages 95ff) discusses the methodological and technical details of the individual modules. It is followed by a block dealing with generating the data base of the model (“Data base generation”, pages 305ff) with issues such as filtering the global SAM or the split utility before a section on running the model (“Scenario runs”, pages 376 ff). Next, post-model processing is detailed: with the sub-section on “Decomposing welfare changes in CGEBox”, “Trade in VA indicators” or the discussion of the “Altertax” utility. Afterwards, the use of the Graphical User Interface is presented in the section “GUI overview” (pages 426ff). It discusses step-by-step how to install the model, construct a data base, define the shock, run the model and analyze results. A further section details the use of an extension on sensitivity analysis (“Sensitivity analysis”, pages 455ff), before section “Getting an overview on the code” (pages 466 ff) discusses some technical aspects. Some training videos are available at [http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox\\_e.htm#vid](http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox_e.htm#vid) – they might differ however in selected details as the Graphical User Interface was improved since their recording.

### 1.3 Scalability, Modularity and Extensions as basic features of CGEBOX

As the standard GTAP model and many of its variants, CGEBox is highly scalable in the sense that it can work with differently sized, but identically structured data bases which differ in the lists of regions, commodities, sectors and factors. Increasing the detail of the data base can lead to models of a

---

<sup>4</sup> Rutherford T. and Harbor A. (2005). GTAP6inGAMS: The Dataset and Static Model. Prepared for the Workshop: “Applied General Equilibrium Modeling for Trade Policy Analysis in Russia and the CIS”, The World Bank Resident Mission, Moscow, December 1-9, 2005, <http://www.mpsge.org/gtap6/gtap6gams.pdf>

<sup>5</sup> Britz, W., van der Mensbrugghe, D. (2018): CGEBox: A Flexible, Modular and Extendable Framework for CGE Analysis in GAMS, Journal of Global Economic Analysis 3(2): 106-176



size which is hard or even impossible to solve. The modeling platform offers several options to work with large-scale data bases. Firstly, a filter algorithm (see section “Filtering and Rebalancing”, page 324ff) can systematically remove tiny elements from the global SAM and afterwards rebalance it. Second, the GAMS code comprises a pre-solve algorithm which should speed up solution of larger models such as with more than 300.000 non-empty transactions in the global SAM. The code was intensively tested for numerical stability and acceptable solution time when applied to more disaggregated versions of the GTAP data base (Britz et al. 2015, Britz and Van der Mensbrugghe 2016<sup>6</sup>). Third, certain extensions such as the possibility to use as in GLOBE the same import and domestic shares for all Armington agents can reduce the number of variables and equations and, finally, the model codes substitute out by default many variables from the system which otherwise drives up model size.

CGEBox includes and supports extensions beyond the GTAP standard model. Using these extensions is partly driven by the underlying SAM structure, such as allowing for a multi-product make matrix to support non-diagonal relations between production activities and commodities outputted. Other extensions are switched on depending on the parameterization of the system, such as allowing for non-infinite CET between domestic sales and exports, and between export flows. Similarly, non-diagonal make matrices can be applied in conjunction with a CET nest to steer output composition of multi-output sectors and a CES nest to determine consumption shares in case several sectors produce the same commodity. Many of these extensions are inherited from ENVISAGE, however some ENVISAGE features are found in the code. The code also supports already a range of different closures, for instances, for foreign savings, private and government consumption which can be complemented by closures files provided by the user. Equally, different regional numeraires can be used.

Furthermore, the code structure has a modular design which facilitates introduction of new model features to the GAMS code<sup>7</sup>. So far, implementations of *GTAP-AGR* (see section “GTAP-AGR”, page 208ff), *GTAP-E* (see section “189”, page 189), *GTAP-AEZ* (see section “GTAP-AEZ”, page 293ff), *myGTAP* (see section “myGTAP module”, page 95ff) and CO2 emissions accounting and taxation (see section “CO2, Non-CO2 Accounting”, page 189ff, this also covers air pollution emissions and energy use accounting oil equivalents) are available as additional modules. Different impacts of climate change can be captured by the Climate Change Damages Module (see section “Climate Change Damages”, page 305ff).

The options for international trade are especially evolved with an implementation of a Melitz model (*GTAP-Melitz*, see section “GTAP-Melitz: Heterogenous firm module”, page 211ff), a *GTAP-MRIO* extension that splits up of bi-lateral import demand by total intermediate consumption and the final demand agents, (see section “MRIO extension”, page 168), the possibility to use an *Armington specification with commitment terms* at the bilateral level (see section “Armington with commitment

---

<sup>6</sup> Britz, W., van der Mensbrugghe, D. (2016): Reducing unwanted consequences of aggregation in large-scale economic models - A systematic empirical evaluation with the GTAP model, *Economic Modelling* 59: 462-473

<sup>7</sup> The code makes a differentiation between extensions which change the nesting of the production function or introduce nested CET structures for factor supply. These extensions only require set-definitions and substitution respectively transformation elasticities, but no changes in the equation structure of the model or is post-processing. In opposite to that, the implementation of what is called a module consists typically of two files: one file which declares the additional symbols (parameters, variables, equations), these includes are found in “model.gms”, and a file which calibrates the related parameters for the benchmark, these includes are found in “cal.gms”. Potentially, there are also additional statements in the post-model reporting part.

terms” page 157ff) and an implementation of a *Spatial Price Equilibrium approach* (SPE, see section “Spatial Price Equilibrium instead of Armington”, page 163ff). The user can decide for which markets the different option are activated, the MRIO option can be combined with other specification with the exemption of the SPE.

Most options for international trade can be combined with different agent aggregations in the Armington nests and with a module which dis-aggregates selected bi-lateral trade flows to the tariff line (see “Tariff Line Module”, page 180ff). The *G-RDEM* (GTAP based Recursive-Dynamics Model<sup>8</sup>) is available as a module in CGEBox to generate (long-term) baselines (see section “G-RDEM: Long-term baseline generation and analysis”, page 211 ff.), adding to the default features for recursive-dynamic analysis in CGEBox (see section “Recursive-dynamics”). Additionally, the production functions and factor markets can be broken down to *sub-national regions*, for instance, to 280 so-called administrative NUTS2 regions for the EU (see section “Sub-national dis-aggregation of production and factor markets in CGEBox”, page 146ff).

Extensions derived from GTAP-variants are mostly not set-up as a perfect replica of the GEMPACK implementations and not tested for replication of results obtained with the corresponding GEMPACK code. Examples of how these extensions and modules can be used to generate from the same code models with quite different features offer the pre-defined configurations; see section “Pre-configurations”, pages 394ff. The code supports currently GTAP9 to GTAP12 data bases with matching AEZ data. The “Split utility” can be used in conjunction with FAOSTAT data to add, for instance, agri-food detail to the GTAP SAMs. There are some specific extensions for post-model analysis: Trade in Value Added indicators (see section “Trade in VA indicators”, pages 418ff), decomposition of welfare changes (see section “Decomposing welfare changes in CGEBox”, pages 409ff), and a SDG indicator framework. The model can also, similar to the Altermat utility (see section “Altermat”, pages 422ff), derive a global SAM and other parameters from a simulated shock which can be used as a starting point for another run.

An important feature of the model’s structure are generic equations that allow the introduction of CES-nests comprised of factors, intermediates demand and sub-nests such that even complex nestings in the production function can be introduced in the model’s code via cross-set definitions without additional programming work. A similar generic implementation of nesting structures is provided for factor supply, the sourcing of a product from multiple production activities and for CES sub-nests under final demand. In all four cases, the post-model processing and reporting part reflects these nesting structures as well. These features are discussed below in more detail in the section “Flexible nesting”, pages 90ff. Pre-defined nests are available for substitution between closely related products, for instance, matching the detail of the GTAP-CE and GTAP-CM data bases, and to mimic the nesting structures of the GTAP-E and GTAP-Power models.

On demand, the production activities can be broken down in two vintages where the old vintage uses depreciated capital, only, considered sector specific, and the new vintage competes for newly formed capital. Their production functions differ in their substitution elasticities; old vintages can shed old capital based on an elasticity and the price relation between their returns to old capital and economy-wide price of new one.

---

<sup>8</sup> Roson, R., Britz, W. (2021): Simulating long run structural change with a dynamic general equilibrium model, *Int. J. Computational Economics and Econometrics* 11(4): 368-404; Britz, W., Roson, R. (2019): G-RDEM: A GTAP-Based Recursive Dynamic CGE Model for Long-Term Baseline Generation and Analysis, *Journal of Global Economic Analysis* 4(1): 50-96

The basic application mode of the model, according to the GTAP standard model, is that of a comparative-static global model of a barter economy, i.e. with fixed exchange rates. Alternatively, the model can also be run in recursive-dynamic fashion and in that case combined with the features of “G-RDEM: Long-term baseline generation and analysis”, pages 211ff.. Equally, a single regional model can be directly derived from the equation system, and the single region to run can be selected by the GUI. The code and GUI also support deriving a partial equilibrium model where prices for some commodities and related intermediate and factor demand of the activities producing these commodities are fixed along with total income and the model is solved for the remaining endogenous commodity markets.

Use of the options, extensions and modules discussed above does not require additional coding efforts; they can be directly activated from the interface or by using non-default parameter files. The set-up of the code should render it relatively straightforward to implement additional modules in a modular fashion. It should be noted here that the modules are to a large extent inter-operable, which allows setting up a model with captures simultaneously the features of GTAP-AGR, GTAP-AEZ and GTAP-E, treats some sectors a la Melitz, uses different household types etc.

Data input for the model is based on downloading zip-archives from the GTAP-center with the GTAP Data Base and additional data (land use, migration, MRIO etc.) with full detail (regions, products, sectors, factors) for a specific version and year. The content is then re-organized into a set of GDX containers into one folder (see section “Loading a GTAP Data Base”, page 313ff.), such that multiple data bases (different release version, years, standard or GTAP-Power) can be used in parallel. The data base driver of CGEBox (see section “GTAPAGG: The data base generation utility of CGEBox”, page 316ff.) allows to define the aggregation over regions, products, sectors and factor and to steer further details, such as filtering out small transaction, adding gender differentiated labor account, MRIO split factors and performing a split. The outcome is one single GDX container used by the simulation model for benchmarking and simulation which can also be exchanged between CGEBox users. Generated data bases can be stored back into a data base folder to perform data base generation in multiple steps. The modeling platform is supported by a Graphical User Interface, see section “GUI”, 426ff. which also details how to install the system.

Technically, the model is set-up as a constraint system of equations in levels. The equations and variables are to a largest extent paired in MCP style which also helps debugging the model especially during further developments. Due to the pairing, the model can be solved alternatively as a Mixed Complementary System (MCP) which allows exploiting features as such endogenous quota rents. An explorative version uses the so-called “Extended Mathematical Programming” (EMP) package of GAMS to optimize various taxes / subsidies such as to maximize a (global) welfare function, taking the equations of the CGE as constraints.

**Table 1:** Modules and extensions

| Module                                            | Status                                    | Remarks                                                                                                                                                     |
|---------------------------------------------------|-------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Data base construction (only main options)</i> |                                           |                                                                                                                                                             |
| Manual correction                                 | Tested                                    | Allows to correct data of the underlying data base, requires data filter                                                                                    |
| Data filter                                       | Tested, not used with GEMPACK based model | Optionally removes small transactions from global SAM / trade matrices while maintaining closely macro totals. Important to support model applications with |

| Module                                                         | Status                                                                | Remarks                                                                                                                                                                                                                                             |
|----------------------------------------------------------------|-----------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                |                                                                       | highly dis-aggregated data bases. Draws on code by T. Rutherford. Use generally recommended.                                                                                                                                                        |
| Gender differentiated labor accounts                           | Tested                                                                | Draws on GDLD data base by World bank for GTAP Version 10 standard. Data for missing regions / sectors in V11 and more detailed data base (GTAP-Power, GTAP-CE) derived with proportionality assumptions                                            |
| Macro updates                                                  | Tested                                                                | Allow to introduce changes to the GTAP Data base that requires a global re-balancing, such as updating the B.o.T.                                                                                                                                   |
| Split utility                                                  | Tested                                                                | Dis-aggregates a data base to further activity and or product detail based on split-factors provided by user codes. Ready to use codes to splits based on FAO data. See also Britz 2021 and Britz 2022 <sup>9</sup>                                 |
| Sub-national data (NUTS2 data for Europe)                      | Tested                                                                | Based on different Eurostat statistics. Focus on output composition and land use. Identical cost shares for sectors across sub-regions. List of sub-regions with shares on outputs and population can for other countries can be added by the user. |
| <i>Simulation runs</i>                                         |                                                                       |                                                                                                                                                                                                                                                     |
| GTAP-Standard                                                  | Tested for exact replication                                          | With extensions from ENVISAGE such as non-diagonal make matrix, CET on export side                                                                                                                                                                  |
| Completely flexible nesting structures of production functions | Tested, based on set-definitions                                      | Should allow to quickly generate variants of the standard GTAP model currently available which differ in nesting of factors / intermediate demands                                                                                                  |
| Completely flexible nesting structure of factor supply         | Tested, based on set-definitions                                      | Should allow to quickly generate variants of the standard GTAP model currently available which use nested CET structures to describe factor supply                                                                                                  |
| Completely flexible of sub-nests under final demand            | Tested, based on set-definitions, not available for the trade margins | Should allow to quickly generate variants of the standard GTAP model currently available which use CES-subnests under the top-level final demand equation                                                                                           |
| Completely flexible nesting of make structures                 | Tested, based on set-definition                                       | The composition of a product outputted by multiple activity can be depicted over multiple CES nests, for instance, to implement the CES nesting from GTAP-Power to define the electricity mix.                                                      |
| LES/CD/MAIDADS functions for final demand                      | Partly found in myGTAP                                                | Parameters derived from CDE parameterization, respectively econometrically estimated (MAIDADS)                                                                                                                                                      |
| GTAP-AEZ                                                       | Operational, not tested for exact replication                         | With land supply elasticities for natural land cover, differentiated by land cover, additive CET, optionally additional CET nests, carbon stock accounting. Differently detailed nesting structures available via                                   |

<sup>9</sup> Britz, W. (2021). Comparing penalty functions in balancing and dis-aggregating Social Accounting Matrices. *Journal of Global Economic Analysis*, 6(1), 34-81; Britz, W. (2022). Disaggregating agro-food sectors in the GTAP data base. *Journal of Global Economic Analysis*, 7(1), 44-75,

| Module                                                 | Status                                                                                               | Remarks                                                                                                                                                                                                                                                                                                                                                                                                          |
|--------------------------------------------------------|------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                        |                                                                                                      | GUI.                                                                                                                                                                                                                                                                                                                                                                                                             |
| GTAP-AGR                                               | Operational, not tested for exact replication                                                        | Applicable also to regional dis-aggregation different from original GEMPACK implementation, uses the flexible nesting approach, adjusts to sectoral detail. Implements also segmented factor markets.                                                                                                                                                                                                            |
| GTAP-E                                                 | Operational, not tested for exact replications                                                       | Based on flexible nesting approach, adjusts to sectoral detail, uses the additional nesting for power generation from GTAP-Power if the detail is available in the data base (via make CES-nests)                                                                                                                                                                                                                |
| GTAP-Melitz: Heterogenous firm module                  | Operational, similar to GTAP-HET                                                                     | Includes a fixed cost nest based on the flexible nesting approach, sector coverage can be flexibly chosen. Can also be turned in a Krugman specification                                                                                                                                                                                                                                                         |
| MRIO extension                                         | Operational                                                                                          | Differentiation of bi-lateral import demands and tariffs by total intermediate demand and each final demand agent (not a full MRIO in the model equations). Based on GTAP-MRIO data base. Can be combined with the Armington, Armington plus commitment terms and Melitz trade specification but not with SPE.                                                                                                   |
| myGTAP module                                          | Operational, not tested for exact replication                                                        | Removes the regional household, supports multiple private households which manage their own factor stocks, separate government account with debt dynamics, remittances; can be linked to household surveys for 19 countries and post-model micro-simulation, subsistence module where households at sub-regional level consume (larger part) of own production of selected activities, government debt dynamics. |
| GMIG                                                   | Operational                                                                                          | Bi-lateral migration of labor, remittances                                                                                                                                                                                                                                                                                                                                                                       |
| Aggregate Armington aggregator for intermediate demand | Operational, not part of standard GTAP model or variants thereof available from GTAP                 | Domestic and import shares for intermediate demand and related tax rates are not sector specific, removes a large share of equations                                                                                                                                                                                                                                                                             |
| Aggregate Armington aggregator for all agents          | Operational, not part of standard GTAP model or variants thereof available from GTAP, found in GLOBE | Domestic and import shares for intermediate demand and related tax rates are not agent specific, removes a large share of equations                                                                                                                                                                                                                                                                              |
| Third level nest for Armington / CET                   | Small shares can be treated as a Leontief under the second nest, found in GLOBE                      | Might avoid numerical problems with tiny shares. Found in GLOBE                                                                                                                                                                                                                                                                                                                                                  |
| Armington with commitment terms                        | Tested                                                                                               | Allows to model emerging trade flows                                                                                                                                                                                                                                                                                                                                                                             |

| Module                                             | Status                                            | Remarks                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|----------------------------------------------------|---------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Spatial Price Equilibrium instead of Armington     | Tested, but numerically not very stable           | Allows to model emerging trade flows. F.o.B. plus tariff prices of all active import flows of a model region must be identical.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Tariff Line Module                                 | Operational                                       | Allows a CET/CES dis-aggregation of selected bi-lateral trade links including an explicit TRQ mechanism; extension by T. Döbeling to optimize allocation of multiple tariff line trade flows to overarching TRQs                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Capital vintages                                   | Operational                                       | Draws on similar mechanism used in recursive-dynamic CGEs which differentiate vintage from new capital<br><br>Differentiation between depreciated capital stock (not mobile) and gross investment (mobile), capital stock can be rendered endogenous in comparative static mode                                                                                                                                                                                                                                                                                                                                                            |
| Sub-regions                                        | Tested, not available in standard GTAP model      | Breaks down production decisions and factor markets to sub-regional level, currently data available for NUTS2 administrative regions for European countries, data for other countries can be added by user                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Post-model reporting                               | Operational, EV decomposition as in GEMPACK       | Generates SAM like structure, calculates world totals, regional and sectoral totals based on additional GTAP agg file, welfare decomposition etc., feeds into GUI exploitation tools, Trade in VA indicators, welfare decomposition, nutritional indicators, CO2 and Non-CO2 emissions relevant for climate change, air pollution emissions, energy accounting in oil equivalents, micro-simulation module using FAO household surveys for current 19 countries, SDG indicator framework                                                                                                                                                   |
| Single region mode                                 | Operational, not available in standard GTAP model | Fixes import prices and let export demand react to lower Armington nest at export destinations. Alternatively, total import and exports can be driven by double-log functions.                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| G-RDEM: Long-term baseline generation and analysis | Operational, not tested for compliance with GDyn  | Econometrically estimated MAIDADS demand system, endogenous saving rates, differentiated productivity growth, debt accumulation from foreign savings, income dependent industry cost shares and expenditure shares for government and investors, possibility to fix calorie intakes to an empirical relation to per-capita demands, calorie content per dollar can be rendered depending on income per capita, crop yields and crop land expansion can be fixed to exogenous projections, calorie contents can be driven by per capita income; ready-to-use projections for GDP, demographics and educational attainment for the five SSPs |
| Partial Equilibrium closure                        | Operational                                       | Solves only one or some commodity markets and all factor markets, regional or household income exogenous                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

| Module                      | Status      | Remarks                                                                                                                                                                                                                                                                                                                                                                                               |
|-----------------------------|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CO2, Non-CO2 Accounting     | Operational | Can be combined with taxation or CO2 trading permits, considers non-CO2 emissions and carbon stock changes; difference between reported and taxed CO2 equivalent emissions possible (define sectors and products taxed, exclude Non-CO2 emissions from taxation); endogenous process emissions factor based on linear-extrapolation between points on marginal abatement cost curves from literature. |
| Air pollution emissions     | Operational |                                                                                                                                                                                                                                                                                                                                                                                                       |
| Nutrient accounting         | Operational | Estimates macro- and micro-nutrient content per USD and from there nutrient intake per capita and household                                                                                                                                                                                                                                                                                           |
| Energy use accounting       | Operational | Only post model. Energy demand can be defined in MTOE units. If active and use in combination with GTAP-E, additive versions of CES/CET functions for energy products are used                                                                                                                                                                                                                        |
| Climate Change Damages      | Operational | Impacts on labour productivity, health, tourism, energy demands, crop yields and available land from sea level, based on data from Sartori & Roson 2016                                                                                                                                                                                                                                               |
| <i>Sensitivity analysis</i> |             |                                                                                                                                                                                                                                                                                                                                                                                                       |
|                             | Tested      | Uses Latin-Hyper-Cube sampling and parallel model runs                                                                                                                                                                                                                                                                                                                                                |

## 2. Basic model equations

The basic model equations are to a large extent identical to Van der Mensbrugghe, D. (2018). “The Standard GTAP Model in GAMS”, Version 7. Journal of Global Economic Analysis, 3(1), 1-83. We refer below to the equations in that complete and excellent documentation (equation (x) in VDM 2018) and report differences where applicable. We refrain from repeating all equations in mathematical notation – the reader is invited to refer to VDM 2018. Instead, we insert screen shots of the actual code which we comment to ease understanding the technical implementation of the model.

### 2.1 Core sets

The equations of the basic model are comprised in the file “*model/model.gms*” and discussed in the following. The following general sets are used:

**Table 2: Core sets used in model equations**

| Set name  | Description                                                                       |
|-----------|-----------------------------------------------------------------------------------|
| r,rp      | Regions                                                                           |
| rnat,nat1 | Nations (to differentiate from sub-regions in the case the NUTS2 level is active) |
| disr      | Nations which are dis-aggregated to sub-regions                                   |
| subr      | Sub-regions (only populated in the case the NUTS2 level is active)                |
| aa        | Armington agents (sectors, private household, government, savings, transport)     |

| Set name | Description                                               |
|----------|-----------------------------------------------------------|
|          | modes; list differs with Armington agents are aggregated) |
| a        | production activities                                     |
| i,j,k    | products                                                  |
| t        | time                                                      |
| m        | mode of transport (by default air, water, rest)           |
| f        | factors                                                   |
| fm       | mobile factor (fully mobile or sluggish)                  |
| fnm      | immobile factor, i.e. sector specific                     |
| h        | households                                                |
| gov      | government (single item)                                  |
| Inv      | investment (single item)                                  |
| fd       | final demand groups, used in demand nests                 |
| dNest    | demand nests                                              |
| tNest    | technology nests                                          |
| fNest    | factor supply nests                                       |
| mNest    | make nests                                                |

Note that the lists of regions, activities, products and factors depend on the version of the GTAP Data Base used and the chosen aggregation respectively further dis-aggregation. The list of demand, technology, amke and factor supply nests is equally dynamic, depending on which modules are active and/or on additional nests introduced by user provided files. The myGTAP extensions might introduce several private households in the set *h* which is otherwise a singleton. In the standard layout, i.e. without using the NUTS2 extensions, all regions are defined as nations and the list of dis-aggregated regions *disr* is empty, as it the list of sub-regions *subr*.

As the model might run as a single region or as a partial equilibrium model or recursive dynamically, dynamic sets are used to indicate for which regions, product, activities and time points the equations in the current model instance should be generated:

rs        regions in current solve  
ts        time point in current solve  
aIn       activities in current model  
iIn       products in current model

Furthermore, to support sparsity, i.e. to ensure that equations and variables are only generated for non-empty items, a larger set of parameters are in use which serves as flags. The most important ones are listed here:

xtFlag(r,tNest,a)       Indicates if technology nests tNest for region r and activity a is non-empty  
xpFlag(r,a)               Indicates if activity a for region r is non-empty



|                |                                                                 |
|----------------|-----------------------------------------------------------------|
| xfFlag(r,f,a)  | Indicates if primary factor f is used by activity a in region r |
| xaFlag(r,i,aa) | Indicates if Armington agent aa demands product i in region r   |
| xwFlag(r,i,rr) | Bilateral trade flag                                            |

Given these examples, the names of the other flags should be hopefully self-explanatory

## 2.2 Reminder on CES and CET equations

Many relations in the model are based on Constant Elasticity of Substitution (CES) to depict cost minimization or utility maximization and Constant Elasticity of Transformation (CET) functions to depict revenues maximization, using constant returns to scale, i.e. the homothetic case<sup>10</sup>. That implies that average per unit costs (respectively revenues) are equal to marginal costs (respectively revenues). As a consequence, the functional form can only be used to find per unit minimal cost (respectively per unit maximal revenue), but not to define the total amount to demand or to transform. For each CES or CET nest comprising  $n$  inputs or outputs, there are  $n+1$  equations comprised in the model: one equation which defines the average price and  $n$  equations which define the  $n$  demand or output quantities under the nest. The number is increased to  $n+1$  for the additive variants where the dual price aggregator differs from the average price for the aggregate.

The  $n$  quantity equations have always the same structure, shown below for the CES case which defines the demand for  $x_i$  as follows. A given share parameter  $\alpha_i$  defined at the benchmark is multiplied with the total quantity  $y$  to distribute. This total, as explained above, must be defined by another mechanism in the model. The resulting share on the total is then updated by multiplying it with the price relation between the composite price (index)  $\tilde{p}$  and the price  $p_i$ , exponent the substitution elasticity  $\sigma$ . Note that settings the substitution elasticity to zero yields the Leontief case where the benchmark quantity shares stay constant. A shifter variable  $\lambda$  can be used to update preferences or the cost structure to reflect non-Hicks neutral technical progress. The reader should note again that the demands are homothetic: increasing the total  $y$  by 1% will increase also all individual demands  $x_i$  by 1% if prices and shifters stay constant.

$$x_i = \alpha_i y \lambda_i^{\sigma-1} \left( \frac{\tilde{p}}{p_i} \right)^\sigma \quad (1)$$

One might assume that the easiest way to define the composite price (index)  $\tilde{p}$  is to use value exhaustion, i.e. by introducing an equation  $\sum_i x_i p_i = y \tilde{p}$ . But this exhaustion equation requires the individual demands  $x_i$  being known. This is only possible if the average price  $\tilde{p}$  is known as well as it is used in the demand equation for  $x_i$ , while this average price in turn is implicitly defined by the revenue exhaustion equation. That circular relation makes it hard for the solver to find a simultaneous solution for the  $n+1$  equations.

---

<sup>10</sup> For a more elaborate discussion, see for instance, The constant Elasticity of Substitution Function, Dominique von der Mensbrugghe, [https://www.gtap.agecon.purdue.edu/models/envisage/2019ENVISAGEWrkshp\\_2\\_CES-CETFunctions.pdf](https://www.gtap.agecon.purdue.edu/models/envisage/2019ENVISAGEWrkshp_2_CES-CETFunctions.pdf)

Therefore, perhaps astonishingly, the far more non-linear dual price aggregator equation is used in most CGE models and also in CGEBox:

$$\tilde{p} = \left[ \sum_i \alpha_i \left( \frac{p_i}{\lambda_i} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (2)$$

The advantage is that the average price  $\tilde{p}$  can be defined from the individual prices, only, without requiring the total  $y$  or the individual demands  $x_i$ . That reflects the homothetic nature of the function where the total  $y$  does not affect the optimal shares. Once the average price  $\tilde{p}$  is solved for, the other  $n$  equations defining the demanded quantities  $x_i$  can be solved independently from each other. The dual price aggregator for additive case is slightly different and not equal to the average price.

The reader might note that the dual price aggregator is undefined for the CD case where the substitution elasticity is unity as the rightmost exponent will become undefined. The CD case leads to the following dual price aggregator:

$$\tilde{p} = \prod_i \left( \frac{p_i}{\alpha_i \lambda_i} \right)^{\alpha_i} \quad (3)$$

As most CES/CET nests in CGEBox allow for the CD case, both dual price aggregators are found in the model and switched on/off depending on the substitution respectively transformation elasticity. Finally, the case of infinite substitution or transformation cannot be handled by the equations above. They require instead a simply linear aggregator to add the components to the aggregate and the law of one price, i.e. the average price is equal to each component price. Many of the CES and CET functions also support both the standard case discussed above as well as the additive one. The latter is discussed in another section.

The volume preserving variant of the CES adds an additional constraint such that during the simulation the  $x_i$  continue to add-up to  $y$ . The individual demands differ from the standard formulation solely with regard to the exponent on the shifter:

$$x_i = \alpha_i y \lambda_i^{-\sigma} \left( \frac{\tilde{p}}{p_i} \right)^{\sigma} \quad (4)$$

The dual-price aggregator also differs in the exponents from the additive form:

$$\tilde{p} = \left[ \sum_i \alpha_i \lambda_i^{-\sigma} p_i^{-\sigma} \right]^{-\frac{1}{\sigma}} \quad (5)$$

The exhaustion price can be defined as seen in (6). This formulation takes the shares from (5) as the weights in the price index, i.e. dropping the multiplication with  $y$ , and multiplies these weights with the component prices  $p_i$ . The common term  $\tilde{p}^{\sigma}$  found in all share equation is moved behind the summation sign.

$$\tilde{p}_e = \left[ \sum_i \alpha_i \lambda_i^{-\sigma} p_i^{(1-\sigma)} \right] \tilde{p}^\sigma \quad (6)$$

For the transformation with a CET function, the expression for the supplied quantities  $x_i$  that maximize revenues at given transformation elasticity  $\varpi$  from stock  $y$  are quite similar to the CES case:

$$x_i = \alpha_i y \lambda_i^{-1-\varpi} \left( \frac{\bar{p}_i}{\tilde{p}} \right)^\varpi \quad (7)$$

With the dual price index defined as:

$$\tilde{p} = \left[ \sum_i \alpha_i \left( \frac{p_i}{\lambda_i} \right)^{1+\varpi} \right]^{\frac{1}{1+\varpi}} \quad (8)$$

The volume preserving formulation is discussed for the CET case. It is also called the additive variant of the CET function and maximizes the revenue  $r$  of a CET aggregator to distribute a physical stock  $v$  among competing uses  $x$ :

$$\max r = a \left[ \sum_i \alpha_i (\lambda_i p_i x_i)^v \right]^{1/v} \quad (9)$$

Where  $\alpha_i$  are share parameters,  $p_i$  are given prices,  $a$  is a scaling factor,  $x_i$  are the distributed stock quantities to be optimized and  $v$  refers to the transformation elasticity, and  $\lambda_i$  shifters. The A-CET introduces the condition that the sum of  $x$  is fixed to the given volume  $v$ , typically measured in physical units.

$$v = \sum_i x_i \quad (10)$$

The use of physical units will typically imply that the benchmark prices are no longer unity. Instead, the value of a transactions defined in USD in the SAM is divided by the related physical quantity  $x_i$  to define the prices  $p_i$ .

With a given transformation elasticity  $\varpi$ , this generates the following solutions for the competing uses  $x_i$  and the dual price aggregator  $p^c$

$$x_i = \alpha_i (a \lambda_i)^\varpi \left( \frac{p_i}{p^c} \right)^\varpi \quad (11)$$

$$p^c = a \left[ \sum_i \alpha_i (\lambda_i p_i)^\varpi \right]^{1/\varpi} \quad (12)$$

The advantage of using the modified A-CET formulation besides ensuring a linear aggregation is twofold. Firstly, the transformation parameter  $\varpi$ , for instance provided by literature based on the standard CET, can still be used for benchmarking. Secondly, structural changes to the equations model are relatively minor. Specifically, equation (6) is indeed identical to the original CET formulation with the exemption of the shifter while equation (6) only changes the exponent. However, the price index in (6) is no longer equal to the average revenue exhaustion price  $P^e$  which can be defined in a separate equation based on exhaustion:

$$P^e = \sum X_i P_i / V \quad (13)$$

## 2.3 Overview on the supply side

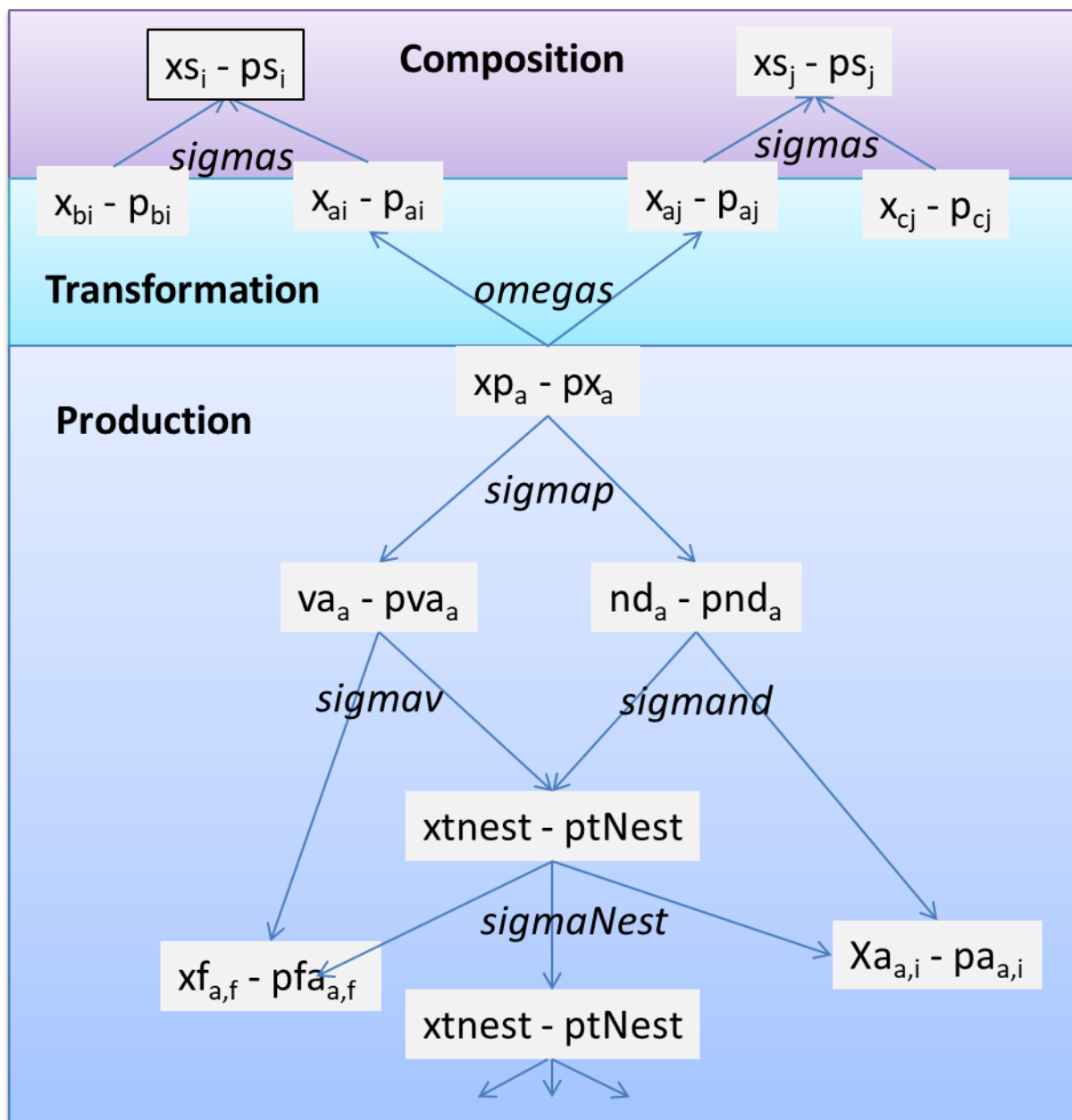
The main features of the production system are as follows. Fully flexible multi-layer CES nesting for production functions with support for volume preserving nests. Demand for endowments or intermediates can also be driven in share by multiple nests. Support for multi-output production functions. Production functions can be depicted at sub-national level. Implementation of fix-cost nests (non-constant returns to scale) in the context of the Melitz module.

The following Figure 1 depicts the quantity and price variables as well as the substitution and transformation elasticities used on the supply side of the CGE model. The bottom part is defined for the production activities  $a$  with total output denoted with  $x_p$  and the related price with  $p_x$ . It is composed of a value added composite  $va$  and an intermediate demand composite  $nd^{11}$ , that jointly define the output of an activity.

As any other technology nests  $tNest$ , the  $va$  and  $nd$  nests combine primary factor demand  $xf$ , intermediates demand defined as Armington demands  $xa$  and other technology sub-nests  $tNest$  into a sub-production function. Note here firstly that primary factors or intermediates can be present in different shares in sub-nests and, secondly, that if the Melitz / Krugmann specification is used for a sector, fix costs are present in a separate sub-nest which does not contribute to  $x_p$ .

---

<sup>11</sup> The  $va$  and the  $nd$  are the two top nests that jointly define the output of an activity. Otherwise, they follow the same logic as any other technology nest.



**Figure 1: Overview on production function nesting**

The output of the activities  $xp$  can be transformed for the non-diagonal make case to the supply of different commodities  $x$  as shown in the middle box. If several activities  $a$  produce the same commodity  $i$ , the different  $x$  can be combined into total supply  $xs$  based on a CES aggregator as shown in the top box. The CES aggregator can span over multiple make nests.

For the upper two boxes, the code supports the case of finite and infinite transformation respectively substitution where the infinite case implies a linear aggregation and the law of one price. For the production nests, only finite transformation is supported including the CD case.

Note that the  $x$  and  $xp$  are substituted from the model equations for the diagonal case.

## 2.4 Production block

The model is set up to work with non-diagonal make matrices where one activity might produce several outputs and one output might be produced by several activities. The production block therefore is defined for activities  $a$  and not for the outputs  $i$ . Furthermore, in case regions are dis-aggregated to sub-national regions, the production function is defined for these dis-aggregated regions. Accordingly, nations which are dis-aggregated to sub-regions  $disr$  are excluded from these equations. The production reflects the “Flexible nesting” approach which allows introducing CES-subnests under the value and the intermediate composite nests, or under other CES-subnests.

The nested production function for each activity  $a$  comprises a **top nest** which combines a value added  $va$  and intermediate demand  $nd$  composite with a substitution elasticity of  $p\_sigmap$ <sup>12</sup>. Note that both nests can comprise a mix of endowments and intermediates depending on the chosen nesting structures. These nests are maintained to allow a more standard implementation of the GTAP model.

The production frontier can be shifted with the variable  $axp$ . The top nest is represented by its dual price aggregator in the equation  $e\_px$ . This equation considers three cases which are shown below: (1)  $p\_sigmap$  is non-zero and different from unity with leads to the usual dual price aggregator for the CES case, (2) the CD case where  $p\_sigmap$  is unity with a different dual price aggregator and (3) the Leontief case with  $p\_sigmap$  equal to zero. The latter is introduced for technical reasons, only, as it will generate a more compact Jacobian in GAMS.

The equation is identical to VDM 2018, equation (3):

```
*
* --- Unit cost definition (also zero profit condition) equation (3) in VDM 2018
* (note: that equation is used with a non-diagonal make matrix)
*
e_px(rs,aIn(a),ts) $ (xpFlag(rs,a)
$ (not disr(rs)))..

axp(rs,a,ts)*px(rs,a,ts)*a(px,'rs,a,ts') =e=

*
* -- Non Leontief case with substitution between VA and ND
*
( ( [p_atNestT(rs,"ND",a,"top")*(ptNest(rs,"ND",a,ts)/lambdaN(rs,"ND",a,ts))**(1-p_sigmap(rs,a))] $ p_atNest(rs,"ND",a,"top")
+ [p_atNestT(rs,"VA",a,"top")*(ptNest(rs,"VA",a,ts)/lambdaN(rs,"VA",a,ts))**(1-p_sigmap(rs,a))] $ p_atNest(rs,"VA",a,"top") ) ** (1/(1-p_sigmap(rs,a))) ) s(px,'rs,a,ts')
) $ ((p_sigmap(rs,a) gt eps) $ (p_sigmap(rs,a) ne 1))

*
* --- CD case with substitution between VA and ND, p_sigmap=1
*
+ ( [p_axpCD(rs,a,ts)*s(px,'rs,a,ts')]
* (([ptNest(rs,"ND",a,ts)/p_atNestT(rs,"ND",a,"top")*lambdaN(rs,"ND",a,ts)]**p_atNest(rs,"ND",a,"top") $ p_atNest(rs,"ND",a,"top") + 1 $ (not p_atNest(rs,"ND",a,"top")))
* (([ptNest(rs,"VA",a,ts)/p_atNestT(rs,"VA",a,"top")*lambdaN(rs,"VA",a,ts)]**p_atNest(rs,"VA",a,"top") $ p_atNest(rs,"VA",a,"top") + 1 $ (not p_atNest(rs,"VA",a,"top"))))
) $ (p_sigmap(rs,a) eq 1)

*
* -- Leontief case (only introduced to ease work for solver), as in standard GTAP model
*
+ ( ( [p_atNestT(rs,"ND",a,"top")*(ptNest(rs,"ND",a,ts)/lambdaN(rs,"ND",a,ts))] $ p_atNest(rs,"ND",a,"top")
* [p_atNestT(rs,"VA",a,"top")*(ptNest(rs,"VA",a,ts)/lambdaN(rs,"VA",a,ts))] $ p_atNest(rs,"VA",a,"top") ) s(px,'rs,a,ts') ) $ (p_sigmap(rs,a) le eps)
;
```

Note that earlier version of the model substituted out the price  $p$  for the diagonal case. This was dropped with the introduction of activity and product specific production taxes and the possibility of a flexible make structure to make the code leaner.

The demand for the **value added** composite  $va$  is defined in the equation  $e\_va$ , VDM 2018, Equation (1). It comprises the same symbols as shown above in the top-level unit cost definition. Note that the equation treats the Leontief case where  $p\_sigmap$  is zero differently by removing the prices from the equation with a \$ operator which can speed up solution.

```
--- Demand for aggregate value added (equation (1) in VDM 2018)

e_va(rs,"va",aIn(a),ts) $ xtFlag(rs,"va",a) ..
xtNest(rs,"va",a,ts) s(xtNest,'rs,"va",a,ts') =e= p_atNestT(rs,"VA",a) * m_xp(rs,a,ts) s(xtNest,'rs,"va",a,ts')
* [ ((m_px(rs,a,ts)/ptNest(rs,"va",a,ts))**p_sigmap(rs,a) - 1) $ (p_sigmap(rs,a) gt eps) + 1 ]
* (axp(rs,a,ts)*lambdaN(rs,"va",a,ts))**p_sigmap(rs,a)-1 ;
```

<sup>12</sup> Generally, all GAMS parameters carry a  $p\_$  prefix to distinguish them from variables. Parameters cannot become subject to a closure swap in GAMS.

Note that the  $e\_va$  equation is scaled with the macro  $s(xtNest, 'rs, "va", a, ts')$ . Scaling factors are by default switched on and present basically in all equations relating to quantities or volumes to ease automated scaling by the solver and provide a more useful interpretation of the relative and absolute tolerances used by the solver.

The output quantity of a sector is defined in the macro  $m\_xp$ . It uses directly the commodity supply  $xs$  in case of a diagonal make matrix for that activity, i.e.  $diag(a)$  is not zero, otherwise, it introduces the activity output  $xp$  in the equation. Due to output taxation and the use of energy equivalents to define quantities, the supplied quantities of a product  $i$  can differ from the output  $xp$  even for the diagonal case. This is reflected by converting  $xs$  to  $xp$  for time point  $t$  with their benchmark relation.

```

--- some activities produce more than one output
$ifc %nDiag%>0 $macro m_xp(r,a,t) { xp(r,a,t) $ (not diag(a)) + sum(k $ xFlag(r,a,k), xs(r,k,t)*xp.l(r,a,"%t0%")/xs.l(r,k,"%t0%")) $ diag(a) }
--- all activities are producing one output, only
$ifc %nDiag%=0 $macro m_xp(r,a,t) sum(k $ xFlag(r,a,k), xs(r,k,t)*xp.l(r,a,"%t0%")/xs.l(r,k,"%t0%"))

```

The equation  $e\_nd$ , VDM 2018, Equation (2), identically structured as the  $e\_va$  equation above, drives the demand for the **intermediate demand** composite:

```

--- Demand for intermediate demand bundle (equation (2) in VDM 2018)

e_nd(rs,"nd",aIn(a),ts) $ xtFlag(rs,"ND",a) ..

    xtNest(rs,"nd",a,ts)s(xtNest,'rs,"nd",a,ts') =e= p_atNestT(rs,"ND",a) * m_xp(rs,a,ts)s(xtNest,'rs,"nd",a,ts')
    * [ {(m_px(rs,a,ts)/ptNest(rs,"nd",a,ts))*p_sigmap(rs,a) - 1} $ (p_sigmap(rs,a) gt eps) + 1 ]
    * {axp(rs,a,ts)*lambdaN(rs,"ND",a,ts))*p_sigmap(rs,a)-1} ;

```

The demand for **primary factors**  $xf$  by each activity depends on a shifter variable  $lambdaf$ , the share parameter  $p\_af$ , the total demand  $xtNest$  of the nests it belongs to and the price relation between the price of these nests  $ptNest$  and the sector specific factor price  $pfa$ , defined via the macro  $m\_pfa$ , exponent the substitution elasticity  $p\_sigmaNest$ :

```

--- Demand for factors, equation (4) in VDM 2018, with extension for commodity nests

m_xf(rs,f,aIn(a),ts) $ (xfFlag(rs,f,a) $ (not disr(rs))) ..

xf(rs,f,a,ts)s(xf,"rs,f,a,ts")
=e=

    --- demand for factor inside technology nest
    sum(tNest_f_a(ftNest,f,a) $ p_afNestT(rs,f,a,ftNest),

    --- factor cost share times demand for technology nest
    p_afNestT(rs,f,a,ftNest)*xtNest(rs,ftNest,a,ts)s(xf,'rs,f,a,ts')

    --- relative price impact (in case sigmndNest is not 0)
    * [ {([ptNest(rs,ftNest,a,ts) $ (not addCes(rs,ftNest,a)) + pdtNest(rs,ftNest,a,ts) $ addCes(rs,ftNest,a)]/%rr% m_pfa(rs,f,a,ts))*p_sigmaNest(rs,ftNest,a)
    -1 } $ (p_sigmaNest(rs,ftNest,a) gt eps) +1 ]
    * lambdaf(rs,f,a,ts)**([p_sigmaNest(rs,ftNest,a)-1] $ (not addCes(rs,ftNest,a)) - p_sigmaNest(rs,ftNest,a) $ addCes(rs,ftNest,a) ));

```

The equations support the use of the additive CES function form. A factor might also be driven in shares by different nests. For instance, part of the land could be a substitute to irrigation water and another one a substitute with yield increasing inputs. Note also the  $%rr%$  macro that will introduce the national factor tax and subsidy rates in case that the region  $rs$  is a sub-national region.

The various dollar conditions might warrant some comments. The first one is the flag  $xfFlag$  indicating that the activity  $a$  is using factor  $f$  in regions  $rs$ , while the second one excluded nation where the production is dis-aggregated to sub-nations from the equation. The dollar condition after the summation sign  $p\_aftNestT(rs,f,a,ftNest)$  ensures that only such nests are used where share parameters are present. This might catch cases where due to numerical thresholds, share parameters are set to zero despite the fact that there some tiny quantity is reported in the SAM.

The dual price aggregator equation  $e\_ptNest$  for technology nests is depicted below. It replaces also separate equations for the composite price for the value-added and intermediate nests (equations (5)

and (7) in van der Mensbrugghe 2018). The equation considers the possible components: intermediates with the related share parameter  $p_{ioNest}$ , primary factors with their share parameter  $p_{afNest}$  and finally sub-nests with share parameters  $p_{atNest}$ . It differentiates in three blocks the cases where the substitution elasticity  $p\_sigmaNest$  is (1) not unity, i.e. CES or Leontief, (2) unity, i.e. the CD case, and (3) the case where solely one input defines the

nests. It also supports the additive case.

The intermediate demand part of nests it is driven by the input coefficients  $p_{ioNest}$  and their activity specific price defined via a macro  $m_{pa}$  and product and activity specific technology shifters  $lambdaio$ , again captured by a macro. Note that the coefficients  $p_{ioNest}$  describe shares inside the intermediate nest, and not relative to total output. The price and shifters used for intermediates and primary factors are identical to those described above for the value added and intermediate composite nests.

```

ptNest(rs,tNest,aIn(a),ts) $ xtFlag(rs,tNest,a) ..

[ptNest(rs,tNest,a,ts) $ (not addCES(rs,tNest,a))
+ pdtNest(rs,tNest,a,ts) $ addCES(rs,tNest,a)]s(ptNest,'rs,tNest,a,ts')
==

%scale% %rr%

---- Standard case (not Leontief and not CD)
{
  [
    --- aggregate over intermediate inputs assigned to technology nest

    sum(tNest_i_a(tNest,i,a)
      %scale% $ p_ioNestT(rs,tNest,i,a)
      ,p_ioNestT(rs,tNest,i,a)
      * (m_paIntr(%rPrice%,i,a,ts))**(1 $ (not addCES(rs,tNest,a)) - p_sigmaNest(rs,tNest,a))
      / lambdaio(rs,i,a,ts)**[ +p_sigmaNest(rs,tNest,a) $ addCES(rs,tNest,a)
      - (p_sigmaNest(rs,tNest,a)-1) $ (not addCES(rs,tNest,a))]
      )
    %scale% $ ioTNest(tNest)

    --- aggregate over factors assigned to technology nests

    + sum(tNest_f_a(tNest,f,a)
      %scale%
      $ p_afNestT(rs,f,a,tNest)
      ,p_afNestT(rs,f,a,tNest)
      * m_pfa(rs,f,a,ts)**(1 $ (not addCES(rs,tNest,a)) - p_sigmaNest(rs,tNest,a))
      / lambdaf(rs,f,a,ts)**[ +p_sigmaNest(rs,tNest,a) $ addCES(rs,tNest,a)
      - (p_sigmaNest(rs,tNest,a)-1) $ (not addCES(rs,tNest,a))]
      )
    %scale%
    $ fTNest(tNest)

    --- aggregate over sub technology nests assigned to that technology nests

    + sum(tNest_n_a(tNest,subTNest,a)
      %scale%
      $ p_atNestT(rs,subTNest,a)
      ,p_atNestT(rs,subTNest,a)
      * ptNest(rs,subTNest,a,ts)**(1 $ (not addCES(rs,tNest,a)) - p_sigmaNest(rs,tNest,a))
      / lambdaN(rs,subTNest,a,ts)**[ +p_sigmaNest(rs,tNest,a) $ addCES(rs,tNest,a)
      - (p_sigmaNest(rs,tNest,a)-1) $ (not addCES(rs,tNest,a))]
      )
    %scale%
    $ ttNest(tNest)

  ]** (1/(1 $ (not addCES(rs,tNest,a)) - p_sigmaNest(rs,tNest,a)))
}

```



```

$$iftheni.docfile "%gams.docfile%"=="
    $ ((p_sigmaNest(rs,tNest,a) ne 1) and (p_sigmaNest(rs,tNest,a) gt eps)) or addCes(rs,tNest,a))
$$else.docfile
    $ ( (p_sigmaNest(rs,tNest,a) ne 1))
$$endif.docfile

$$iftheni.docfile "%gams.docfile%"=="

    ---- Leontief case (automatically set if the parent has one child, only)
+ [
    --- aggregate over intermediate inputs assigned to technology nest
        sum(tNest_i_a(tNest,i,a) $ p_ioNestT(rs,tNest,i,a),
            p_ioNestT(rs,tNest,i,a)*m_paIntr(%rPrice%,i,a,ts)/lambdaio(rs,i,a,ts)) $ ioTnest(tNest)
    --- aggregate over factors assigned to technology nests
        + sum(tNest_f_a(tNest,f,a) $ p_afNestT(rs,f,a,tNest),
            p_afNestT(rs,f,a,tNest)*m_pfa(rs,f,a,ts)/lambdaf(rs,f,a,ts)) $ fTnest(tNest)
    --- aggregate over sub technology nests assigned to that technology nests
        + sum(tNest_n_a(tNest,subTNest,a) $ p_atNestT(rs,subTNest,a),
            p_atNestT(rs,subTNest,a)*ptNest(rs,subTNest,a,ts)/lambdan(rs,subTNest,a,ts)) $ ttNest(tNest)
    ]
    $ (p_sigmaNest(rs,tNest,a) le eps)
$$endif.docfile
    
```

#### Notes:

- the %gams.docfile% commands are used in case that a model documentation is generated in LATEX format. For standard applications, the global %gams.docfile% is empty.
- The global %scale% serves the same purpose: it is equal to an empty string for normal model runs and equal to \$\$hidden otherwise, removing the content behind the global from compilation. It removes elements in equations that are not necessary to understand their logic.

Note that his equation is not part of the GTAP Standard as documented in VDM 2018, it reads in mathematical notation:

$$PNEST_{r,tNest,a} = \left[ \begin{aligned} & \sum_{ie \in tNest} \alpha_{r,i,a}^{tNest} \left( \frac{PA_{r,i,a}}{\lambda_{r,i,a}^{tNest}} \right)^{1-\sigma_{r,a}^{tNest}} \\ & + \sum_f \alpha_{r,f,a}^{tNest} \left( \frac{PF_{r,f,a}^a}{\lambda_{r,f,a}^{tNest}} \right)^{1-\sigma_{r,a}^{tNest}} \\ & + \sum_{tNest \in tNest} \alpha_{r,tNest,a}^{tNest} (PTNEST_{r,tNest,a})^{1-\sigma_{r,a}^{tNest}} \end{aligned} \right]^{1/(1-\sigma_{r,a}^{tNest})} \quad (13)$$

The demand for technology nests  $xtNest$  is depicted by the equation  $e_{xtNest}$  shown below. Besides the VA and ND bundles driven by total output, this demand depends on the demand of their mother nests  $ttNest$  as depicted by the cross-set  $tNest_n_a(ttNest,tNest,a)$ . The related share parameter is denoted with  $p_{atNest}$  and the related price with  $ptNest$ . Equally, in all cases, in order to reduce complexity for the solver, the price relation is taken out when the substitution elasticity is zero, i.e. the Leontief case. The equation further considers the case of the additive CES function form. A second separate block depicts the case where the nest is the sole child of its mother nest.

```

*
* --- Demand for technology nests (not part of standard GTAP model)
*
* xtNest(rs,tNest,aIn(a),ts) $ ( (not tNest_n_a("Top",tNest,a)) $ xtFlag(rs,tNest,a)) ..
*
* xtNest(rs,tNest,a,ts)s(xtNest,"rs,tNest,a,ts") =e=
*
* [
*
* --- nest is part of other nests
*
* + sum(tNest_n_a(ttNest,tNest,a),
*       p_atNestT(rs,tNest,a) * xtNest(rs,ttNest,a,ts)s(xtNest,"rs,tNest,a,ts")
*       * (((pdtNest(rs,ttNest,a,ts) $ addCES(rs,ttNest,a) + ptNest(rs,ttNest,a,ts) $ (not addCES(rs,ttNest,a)))
*         /ptNest(rs,ttNest,a,ts))**p_sigmaNest(rs,ttNest,a) - 1) $ (p_sigmaNest(rs,ttNest,a) gt eps) + 1)
*       * lambdaN(rs,tNest,a,ts)**( [p_sigmaNest(rs,ttNest,a)-1] $ (not addCES(rs,ttNest,a)) - p_sigmaNest(rs,ttNest,a) $ addCES(rs,ttNest,a)
*         ) $ subTNest(tNest)
*
* ] $ (p_atNestT(rs,tNest,a) ne 1)
*
* --- special case where upper nest only comprises the lower one
*
* + [
*       sum(tNest_n_a(tNest1,tNest,a) $ p_atNestT(rs,tNest,a),
*         xtNest(rs,tNest1,a,ts)s(xtNest,"rs,tNest,a,ts")
*         * lambdaN(rs,tNest,a,ts)**( [p_sigmaNest(rs,tNest1,a)-1] $ (not addCES(rs,tNest1,a)) - p_sigmaNest(rs,tNest1,a) $ addCES(rs,tNest1,a))
*       ] $ (p_atNestT(rs,tNest,a) eq 1);

```

Again, that equation is not part of the GTAP standard model version 7 as documented in VDM 2018. In mathematical notation it reads:

$$\begin{aligned}
 & XTNEST_{r,tNest,a} \\
 &= \sum_{tNest1} \alpha_{r,tNest,a}^{tNest1} XTNEST_{r,tNest1,a} \left( \frac{PNEST_{r,tNest1,a}}{PNEST_{r,tNest,a}} \right)^{\sigma_{r,a}^{tNest1}} \quad (13)
 \end{aligned}$$

More information on the nesting approach can be found above in the section “Flexible nesting”.

The additive functional form requires an additional equation to define the exhaustion price:

```

e_pdtNest(rs,tNest,aIn(a),ts) $ (xtFlag(rs,tNest,a) $ addCES(rs,tNest,a)) ..

ptNest(rs,tNest,a,ts)s(ptNest,'rs,tNest,a,ts') =E= %rr%

sum(tNest_i_a(tNest,i,a)
  %scale% $ p_ioNestT(rs,tNest,i,a)
  ,p_ioNestT(rs,tNest,i,a) * lambdaio(rs,i,a,ts)**(-p_sigmaNest(rs,tNest,a))
  * (m_paintr(%rPrice%,i,a,ts))**(1-p_sigmaNest(rs,tNest,a))
  ) $ ioTNest(tNest)
+ sum(tNest_f_a(tNest,f,a)
  %scale% $ p_afNest(rs,f,a,tNest,ts)
  ,p_afNestT(rs,f,a,tNest) * lambdaaf(rs,f,a,ts)**(-p_sigmaNest(rs,tNest,a))
  * m_pfa(rs,f,a,ts))**(1-p_sigmaNest(rs,tNest,a)) $ fTNest(tNest)
+ sum(tNest_n_a(tNest,subTNest,a)
  %scale% $ p_atNestT(rs,subTNest,a)
  ,p_atNestT(rs,subTNest,a) * lambdaN(rs,subTNest,a,ts)**(-p_sigmaNest(rs,tNest,a))
  * ptNest(rs,subTNest,a,ts)**(1-p_sigmaNest(rs,tNest,a))
  ) $ ttNest(tNest)
) * pdtNest(rs,tNest,a,ts)**p_sigmaNest(rs,tNest,a)s(ptNest,'rs,tNest,a,ts');

```

The **case of multiple outputs from one activity** is depicted in the equation  $e_{xq}$  (Equation (8) in VDM 2018). That case is relevant when the flag  $diag(a)$  is not unity, i.e. a not diagonal activity. In that case, the  $p\_omegas$  transformation elasticity distributes the total output  $xp$  to activity specific output  $x$  of each product  $i$  based on the share parameter  $gx$  and the activity specific prices for each product  $i$  termed  $p$  in relation to average per activity prices defined by the macro  $m\_pp$ . In case of infinite transformation, the prices have to be identical. It also considers the case of multiple vintages.

```

*
* --- Convert output into commodities (non-diagonal make matrix with CET for activities, not part of standard GTAP model)
* equation (8) in VDM 2018
*
* x(rsNat,aIn(a),i,ts) $ (xFlag(rsNat,a,i) $ (not diag(a))) ..
*
* 0 =e= (
*       x(rsNat,a,i,ts)s(xp,'rsNat,a,ts')
*       - p_gx(rsNat,a,i,ts)*xp(rsNat,a,ts)s(xp,'rsNat,a,ts') * (p(rsNat,a,i,ts)/m_pp(rsNat,a,ts))**p_omegas(rsNat,a)
*       ) $ (p_omegas(rsNat,a) ne inf)
*
* + (
*       p(rsNat,a,i,ts) $ ( (p_sigmas(rsNat,i) ne inf) or card(aVint))
*       + ps(rsNat,i,ts) $ ( (p_sigmas(rsNat,i) eq inf) and (not card(aVint)))
*       - [m_px(rsNat,a,ts) * m_prdx(rsNat,a,ts)]
*       ) $ (p_omegas(rsNat,a) eq inf)
* ;

```

The related equation  $e_{xp}$ , equation (9) in VDM 2018, considers these two cases accordingly: with infinite transformation, total output  $xp$  is equal to the sum of the commodity outputs, either  $xs$  in the diagonal case or  $x$  otherwise. With finite transformation, the producer price as defined in the macro  $m_{pp}$  is derived from the dual CET price aggregator which uses the share parameters  $gx$ , the prices  $p$  and the transformation elasticities  $p_{omegas}$ . Note that the models support the use an additive CET.

```

.....
*
*   --- Matching price definition, equation (9) in VDM 2018
*
*   e_xp(rsNat,aIn(a),ts) $ (xpFlag(rsNat,a)
*   $ $iftheni.notSubs "%subsOutVars%"=="on"
*   $ (not diag(a))
*   $ $endif.notSubs
*   ) ..
*
*   0 =e=
*
*   --- physical balance (inifinite transformation)
*
*   + [( xp(rsNat,a,ts)*px.l(rsNat,a,"%t0%")
*   - sum(xFlag(rsNat,a,i), x(rsNat,a,i,ts)*p.l(rsNat,a,i,"%t0%"))
*   )s(xp,'rsNat,a,ts')]
*   $ (p_omegas(rsNat,a) eq inf)
*
*   --- dual price aggregator, transformation elasticity is not infinite
*
*   + (px(rsNat,a,ts) - sum(xFlag(rsNat,a,i),p_gx(rsNat,a,i,ts)*p(rsNat,a,i,ts)
*   ** (1 $ (not aACET(a)) + p_omegas(rsNat,a)) ** (1/(1 $ (not aACET(a)) + p_omegas(rsNat,a))))
*   $ ((p_omegas(rsNat,a) ne inf) and (p_omegas(rsNat,a) gt eps))
*
*   + (px(rsNat,a,ts) - sum(xFlag(rsNat,a,i),p_gx(rsNat,a,i,ts)*p(rsNat,a,i,ts)))
*   $ (p_omegas(rsNat,a) le eps)
*
*
;
```

The activity specific producer price charges the production taxes  $prdtx$  on the unit costs  $px$ :

```

--- tax inclusive output price (only used in conjunction with make)

e_pp(rsNat,aIn(a),i,ts) $ (xxFlag(rsNat,a,i) $ (not diag(a))) ..

pp(rsNat,a,i,ts) =E= p(rsNat,a,i,ts)*m_prdtx(rsNat,a,i,ts);

```

The related equation  $e_p$  which aggregate the output of the same commodity from different activities is depicted next. It is only active in the non-diagonal case (not  $diag(a)$ ). It is driven by a share parameter  $p_{ax}$  on its make nest  $xMnest$ :

```

--- Aggregate commodities from activities (non-diagonal make matrix, demand side, CES, not part of standard GTAP model)
Equation (11) in VDM 2018

e_p(rsNat,aIn(a),i,ts) $ (xxFlag(rsNat,a,i) $ (not diag(a))) ..

0 =e=

--- demand for output is driven by its parent make nest

sum(mNest_a(mNest,a,i) $ (p_sigmasNest(rsNat,mNest,i) ne inf),
  x(rsNat,a,i,ts)s(x,'rsNat,a,i,ts')
  - p_ax(rsNat,a,i,ts)*xMnest(rsNat,mNest,i,ts)s(x,'rsNat,a,i,ts')
  * ((( [pmNest(rsNat,mNest,i,ts) $ (not addCES(rsNat,i,"")) + pdmNest(rsNat,mNest,i,ts) $ addCES(rsNat,i,"")]
  /pp(rsNat,a,i,ts))**p_sigmasNest(rsNat,mNest,i) $ p_sigmasNest(rsNat,mNest,i) + 1 $ (not p_sigmasNest(rsNat,mNest,i)))
)
+ sum(mNest_a(mNest,a,i) $ (p_sigmasNest(rsNat,mNest,i) eq inf),
  pp(rsNat,a,i,ts) - pmNest(rsNat,mNest,i,ts));
;
```

The output of make nests is driven by their mother nests:

```

--- ouptut of make nests
e_pMNest(rsNat,mNest,iIn(i),ts) $ xMNest.l(rsNat,mNest,i,"%t0%") ..

0 =e= {

    --- demand for make nests driven by its parent nest

    sum(mNest_n(mNest1,mNest,i) $ (p_sigmasNest(rsNat,mNest1,i) ne inf),
        xMNest(rsNat,mNest,i,ts)s(xMNest,'rsNat,mNest,i,ts')

        - p_axNest(rsNat,mNest,i,ts)*xMNest(rsNat,mNest1,i,ts)s(xMNest,'rsNat,mNest,i,ts')

        *{ ([pmNest(rsNat,mNest1,i,ts) $ (not addCES(rsNat,i,"")) + pdMNest(rsNat,mNest1,i,ts) $ addCES(rsNat,i,"")]
            /pmNest(rsNat,mNest,i,ts))*p_sigmasNest(rsNat,mNest1,i) $ p_sigmasNest(rsNat,mNest1,i)
            + 1 $ (not p_sigmasNest(rsNat,mNest1,i))}

        )

    --- infinite transformation: nest price if equal to price of its paraent

    + sum(mNest_n(mNest1,mNest,i) $ (p_sigmasNest(rsNat,mNest1,i) eq inf),
        pmNest(rsNat,mNest,i,ts) - pmNest(rsNat,mNest1,i,ts))

    } $ (not sameas(mNest,"top"))

    --- link into commodity markets: domestic supply is equal to top level make nest

    + [   xs(rsNat,i,ts)s(xs,'rsNat,i,ts')
        - xMNest(rsNat,"top",i,ts)s(xs,'rsNat,i,ts')
        ] $ sameas(mNest,"top")
;

```

The use of the additive CES case requires an additional equation  $e\_psd$  defining the exhaustion price  $ps$  That differs from the CES dual price aggregate  $psd$ .

```

*
* --- Matching price aggregator, quation (11) in VDM 2018
*
e_ps(rsNat,i,ts) $ ( (xs.range(rsNat,i,ts) ne 0) $ xsFlag(rsNat,i) $ (not diag(i)) ) ..

0 =e=
+ (xs(rsNat,i,ts)/xs.scale(rsNat,i,ts) - sum(a $ xFlag(rsNat,a,i), x(rsNat,a,i,ts))/xs.scale(rsNat,i,ts)
  ) $ (p_sigmas(rsNat,i) eq inf)

+ (ps(rsNat,i,ts)
  - sum(a $ xFlag(rsNat,a,i), p_ax(rsNat,a,i,ts)
    *p(rsNat,a,i,ts)**(1-p_sigmas(rsNat,i))**(1/(1-p_sigmas(rsNat,i)))
  ) $ (p_sigmas(rsNat,i) ne inf)
;

```

While their prices are based on (A)CES price aggregator:

```

--- price definition of make nests

e_xMNest(rsNat,mNest,iIn(i),ts) $ xMNest.l(rsNat,mNest,i,"%t0%") ..

0 =e=

    --- case of finite subst. elasticity: define dual price index

    ({ pmNest(rsNat,mNest,i,ts) $ (not addCES(rsNat,i,"")) + pdmNest(rsNat,mNest,i,ts) $ addCES(rsNat,i,"")
      - [   sum(mNest_n(mNest,mNest1,i) $ p_axNest(rsNat,mNest1,i,ts),
          p_axNest(rsNat,mNest1,i,ts)
          *pmNest(rsNat,mNest1,i,ts)**(1 $ (not addCES(rsNat,mNest,i)) - p_sigmasNest(rsNat,mNest,i)))

          + sum(mNest_a(mNest,a,i) $ p_ax(rsNat,a,i,ts),
              p_ax(rsNat,a,i,ts)*pp(rsNat,a,i,ts)**(1 $ (not addCES(rsNat,mNest,i)) - p_sigmasNest(rsNat,mNest,i)))
          ]**(1/(1 $ (not addCES(rsNat,mNest,i)) - p_sigmasNest(rsNat,mNest,i)))
        )s(pmNest,"rsNat,mNest,i,ts")
      ) $ (p_sigmasNest(rsNat,mNest,i) ne inf)

    --- infinite substitution: physical aggregation from children nests and output

    + ( xMNest(rsNat,mNest,i,ts)s(xMNest,'rsNat,mNest,i,ts')

        - sum(mNest_n(mNest,mNest1,i),
            xMNest(rsNat,mNest1,i,ts)s(xMNest,'rsNat,mNest,i,ts')

        - sum(mNest_a(mNest,a,i) $ xxFlag(rsNat,a,i),
            x(rsNat,a,i,ts)s(xMNest,'rsNat,mNest,i,ts')
        ) $ (p_sigmasNest(rsNat,mNest,i) eq inf)

    ;

```

In case of the additive CES form being used for a make nest (physical aggregation), the exhaustion price is defined as follows:

```

--- dudal price index of make nest for additive CES

e_pdmNest(rsNat,mNest,iIn(i),ts) $ (xmNest.1(rsNat,mNest,i,"%t0%") $ addCes(rsNat,i,"")) ..

pMNest(rsNat,mNest,i,ts) s (pMNest,'rsNat,mNest,i,ts')
=E= [ sum(mNest_a(mNest,a,i) $ p_ax(rsNat,a,i,ts),
      p_ax(rsNat,a,i,ts)*pp(rsNat,a,i,ts)**(1- p_sigmasNest(rsNat,mNest,i)))
    + sum(mNest_n(mNest,mNest1,i) $ p_axNest(rsNat,mNest1,i,ts),
      p_axNest(rsNat,mNest1,i,ts)*pMNest(rsNat,mNest1,i,ts)**(1- p_sigmasNest(rsNat,mNest,i)))
    ] s (pMNest,'rsNat,mNest,i,ts')*pMNest(rsNat,mNest,i,ts)**p_sigmasNest(rsNat,mNest,i);

```

The supply price  $ps$  for product  $i$  is equal to the price index of the top-level make nest:

```

--- define supply price (= price of top-level make nest)

e_ps(rsNat,iIn(i),ts) $ ( (xs.range(rsNat,i,ts) ne 0) $ xsFlag(rsNat,i) $ (not diag(i)) ) ..
      ps(rsNat,i,ts) s (ps,'rsNat,i,ts') =E= pMNest(rsNat,"top",i,ts) s (ps,'rsNat,i,ts');

```

The distribution of total output of the vintages and the related average prices is defined by the following equations based on an additive CES function. The first equation defines the share of output of the two vintages  $aVints$  (old and new) on the total output of activity  $aVint$  based on a CES function. The substitution elasticity is for the default cases chosen as 12, close to perfect substitution. The second equation defines the dual price index from the unit cost prices of the two vintages and the last one defines the exhaustion price.

```

--- share of output produced by each vintage (cost min under additive CES)

e_paVint(xxFlag(rsNat,ain(aVints),i),ts) ..

x(rsNat,aVints,i,ts) s (x,'rsNat,aVints,i,ts')
=E= sum(a_aVint(aVint,aVints),
      x(rsNat,aVint,i,ts)*p_distrVint(rsNat,aVints)
      * (pavgVintD(rsNat,aVint,i,ts)/p(rsNat,aVints,i,ts))**p_sigmaVint(rsNat,aVint)) s (x,'rsNat,aVints,i,ts');

--- dual price of vintage

e_pAvgVintD(xxFlag(rsNat,ain(aVint),i),ts) ..

pavgVintD(rsNat,aVint,i,ts)
=E= sum(a_aVint(aVint,aVints),
      p_distrVint(rsNat,aVints)*p(rsNat,aVints,i,ts)**(-p_sigmaVint(rsNat,aVint))**(-1/p_sigmaVint(rsNat,aVint)));

--- average price of vintage (exhaustion)

e_pAvgVint(xxFlag(rsNat,ain(aVint),i),ts) ..

p(rsNat,aVint,i,ts)
=E= sum(a_aVint(aVint,aVints),
      p_distrVint(rsNat,aVints)*p(rsNat,aVints,i,ts)**(1-p_sigmaVint(rsNat,aVint))
      * pavgVintD(rsNat,aVint,i,ts)*p_sigmaVint(rsNat,aVint));

```

The vintage module considers the option to sell part of the old capital  $p_{gf}$  of each activity based on a supply elasticity  $p_{etaff}$  depending on its marginal value product  $pf$  relative to the average price of new capital  $pft$ . The use of the fudging function  $ncpVusin$  ensures that not more than the depreciated capital can be used. The resulting selling scheme is defined as follows; it defines the non-sold old capital of an activity  $xf$  used in production:

```

--- define capital use to stock ratio rrat for old vintage

e_rrat(xxFlag(rs,f,a),ts(t)) $ (sameas(f,"oldCap") $ (pf.range(rs,f,a,ts) ne 0) $ (not sum(vNest_a(vNest,a),1)) $ (not disr(rs))) ..
xf(rs,"oldCap",a,t) s (xf,'rs,"oldcap",a,t')
=E= p_gf(rs,"oldCap",a,t) s (xf,'rs,"oldcap",a,t')
    * (ncpVusin(1,(pf(rs,"oldCap",a,t) / pft(rs,"newCap",t))**p_etaff(rs,"oldCap",a),1.E-5)-ncpVusin(1,1.E-5)+1);

```

It is possible to assume that several activities (such as closely related crop activities) share their old capital based on the flexible factor markets. In this case, the returns to old capital  $pf$  of each activity are assumed equal to the price of the factor market nest “oldCap”:

```

--- old capital price in case of share depreciated capital stock across activities

e_pfOldCap(xxFlag(rs,f,a),ts(t)) $ (sameas(f,"oldCap") $ (pf.range(rs,f,a,ts) ne 0) $ sum(vNest_a(vNest,a),1) $ (not disr(rs))) ..
      pf(rs,f,a,t) =E= sum(vNest_a(vNest,a),pFNest(rs,"oldCap",vNest,t));

```

The selling scheme for shared capital is defined as follows:

```
e_pfcapVnNest(rs,f,vNest,ts(t)) $ (sameas(f,"oldCap") $ (pfNest.range(rs,f,vNest,ts) ne 0) $ (not disr(rs))) ..
rratVnNest(rs,vNest,t)
=e= ncpVnSin(1,(pfNest(rs,"oldCap",vNest,t) / pft(rs,"newCap",t))**p_etafVnNest(rs,"oldCap",vNest),1.E-5)
-ncpVnSin(1,1,1.E-5)+1;

e_rratVnNest(rs,vNest,ts(t)) $ sum(vNest_a(vNest,a),xflag(rs,"oldCap",a)) ..
rratVnNest(rs,vNest,t) =E= sum(vNest_a(vNest,a) $ xflag(rs,"oldCap",a),(xf(rs,"oldCap",a,t)+1.E-8))
/ sum(vNest_a(vNest,a) $ xflag(rs,"oldCap",a),(p_gf(rs,"oldCap",a,t)+1.E-8));
```

The total amount of sold capital is defined as follows where  $p\_gf$  is depreciated capital stock of each activity and  $xf$  its actual use as defined based on the selling scheme above:

```
e_soldCap(rsNat,ts(t)) ..

soldCap(rsNat,t)s(soldCap,'rsNat,t')
=E= sum((r_r(rs,rsNat),xflag(rs,"oldCap",aOld)),p_gf(rs,"oldCap",aOld,t)-xf(rs,"oldCap",aOld,t))s(soldCap,'rsNat,t');
```

The distribution of output from nations to different sub-regions is described in the section “Integration into the modeling framework” of the chapter “Sub-national dis-aggregation of production and factor markets in CGEBox”.

## 2.5 Factor markets

The main features of the factor markets are as follows. Fully flexible nesting based on multi-layer CET function with support for factor supply elasticities. When using the myGTAP extension, households manage their own factor stocks and supply them to the multi-layer nesting structure. Land markets can be depicted at the level of Agro-Ecological Zones. Support for sub-national factor stocks.

The **supply of fully mobile or sluggish factors**  $xft$  at national level  $rsNat$  is depicted by the equation  $e\_xft$  in case where the factor supply is not fixed (.range eq 0), equation 69 in VDM 2018. It is driven by the factor price  $pft$  relative to the price index of GDP at market prices  $pgdpmp$  and the factor supply elasticity  $p\_etaf$ . If  $p\_etaf$  is zero, the price dependent part becomes a constant of unity and  $xft$  is fixed to the constant  $p\_aft$ . Note that endogenous factor supply is not part of the standard GTAP model. Demand for new capital as a new factor is part of the capital vintage module and depicted differently (not shown here).

Aggregate supply of factors (not part of standard GTAP model), equation (69) in VDM 2018  
Iso-elastic factor supply as a function of factor prices  $pft$  relative to  
global factor price index, times the factor supply elasticity  $etaf$

```
e_xft(rsNat,fm,ts) $ (xftEqFlag(rsNat,fm) $ (xft.range(rsNat,fm,ts) ne 0) $ (Not sameas(fm,"newCap"))) ..

xft(rsNat,fm,ts)s(xft,'rsNat,fm,ts')
=e= (p_aft(rsNat,fm,ts)*(pft(rsNat,fm,ts)/pgdpmp(rsNat,ts))**p_etaf(rsNat,fm))s(xft,'rsNat,fm,ts');
```

The related **economy wide average factor price**  $pft$  is defined by the equation  $e\_pft$ , equation (71) in VDM 2018, expanded by potential factor supply nests, which distinguish the Leontief case (first block), the sluggish case with a dual price aggregator (second block) and the fully mobile case where the equation ensures market clearing (third block).

```

--- Average price of mobile factors, equation (71) in VDM 2018
e_pft(rs, fm, ts) $ (xftEqFlag(rs, fm) and (pft.range(rs, fm, ts) ne 0) and (not discr(rs))) ..
0 =e=

--- sluggish factor mobility, dual price aggregator
[ pft(rs, fm, ts) s(pft, 'rs, fm, ts')

--- factors directly linked into top CET nest
- sum(a $ (xftEqFlag(rs, fm, a) $ ( (not sum(fNest_a_f(fNest, a, f), 1)) or fNest_a_f("xft", a, fm))),
      p_gf(rs, fm, a, ts) * pf(rs, fm, a, ts)) s(pft, 'rs, fm, ts')

--- subnests linked into top CET nest
- sum(fNest_n_f("xft", fNest, fm)
      %scale% $ p_gfNest(rs, fm, fNest, ts)
      , p_gfNest(rs, fm, fNest, ts) * pfNest(rs, fm, fNest, ts)) s(pft, 'rs, fm, ts')

] $ (p_omegaf(rs, fm) eq 0)

+ [ pft(rs, fm, ts) s(pft, 'rs, fm, ts')
- [

--- factors directly linked into top CET nest
sum(a $ (xftEqFlag(rs, fm, a) $ ( (not sum(fNest_a_f(fNest, a, f), 1)) or fNest_a_f("xft", a, fm))),
      p_gf(rs, fm, a, ts) * pf(rs, fm, a, ts) ** (1 + p_omegaf(rs, fm)))

--- subnests linked into top CET nest
+ sum(fNest_n_f("xft", fNest, fm)
      %scale% $ p_gfNest(rs, fm, fNest, ts)
      , p_gfNest(rs, fm, fNest, ts) * pfNest(rs, fm, fNest, ts) ** (1 + p_omegaf(rs, fm)))

] ** (1 / (1 + p_omegaf(rs, fm))) s(pft, 'rs, fm, ts')

] $ (p_omegaf(rs, fm) and (p_omegaf(rs, fm) ne inf) )

--- fully mobile factors, uniform prices => physiscal aggregation
+ [ xft(rs, fm, ts) s(xft, 'rs, fm, ts')
- sum(a $ (xftEqFlag(rs, fm, a) $ ( (not sum(fNest_a_f(fNest, a, f), 1)) or fNest_a_f("xft", a, fm))), xft(rs, fm, a, ts)) s(xft, 'rs, fm, ts')
- sum(fNest_n_f("xft", fNest, fm) $ p_gfNest(rs, fm, fNest, ts), xftNest(rs, fm, fNest, ts)) s(xft, 'rs, fm, ts')

] $ (p_omegaf(rs, fm) eq inf)
;

```

In case of sluggish factor supply ( $\$ (\text{omegaf}(r, fm) \text{ ne inf})$ ), the agent specific factor prices net of taxes  $pf$  are aggregated using the dual price aggregator in the second block, considering the case that activity specific factor demands are directly defining the economy-wide total factor stock and the case that this stock is dis-aggregated to different segment  $fNest$ . In case of fully mobile factors (the second block), the price is not directly defined in the equation, but rather indirectly via market clearing.

**Sector specific factor prices net of taxes  $pf$**  are directly or indirectly defined in the equation  $e\_pf$  shown below, equation (70) in VDM 2018. It considers five different cases. The first case considers sluggish factor supply where the sector specific factor demands  $xf$  are directly aggregated to total factor supply  $xft$  in a CET function. The usual CET distribution logic applies: the supply depends on the share parameter  $p\_gf$  and the total supply  $xft$  as well as on the relation between the price paid in the sector  $pf$  relative to the average one  $pft$ , exponent the transformation elasticity  $p\_omegaf$ . The second case where the factor is part of factor supply nest under sluggish supply is identically structured, with the difference that the average price is now nest specific, i.e.  $pfNest$  as is the transformation elasticity  $p\_omegafNest$ .

Next, we have the two cases with fully mobile supply: either in case of economy wide full mobility or in case of full mobility inside in a nest. In both cases, the sector specific price  $pf$  is equal to the average one,  $pft$  or  $pfNest$ . The last of the five cases depicts immobile factors: here, the default case is that the immobile factor supply elasticity  $p\_etaff$  is zero such that the factor demand  $xf$  must be equal to the given parameter  $p\_gf$ .



```

p pf(rs,f,aIn(A),ts) $ (xfEqFlag(rs,f,a) $ (pf.range(rs,f,a,ts) ne 0) $ (not discr(rs)) ) ..
0 =e=
*
* --- "sluggish" factor supply based on CET function, factor is part of top level nest
*
[ xf(rs,f,a,ts)/xf.scale(rs,f,a,ts)
  - p_gf(rs,f,a,ts)*xft(rs,f,ts)/xf.scale(rs,f,a,ts) * (pf(rs,f,a,ts)/pft(rs,f,ts))**p_omegaf(rs,f)
] $ (fm(f) $ (p_omegaf(rs,f) ne inf) $ ((not sum(fNest_a_f(fNest,a,f),1)) or fNest_a_f("xft",a,f)) )

*
* --- "sluggish" factor supply based on CET function, factor is part of sub-nest
* (not part of standard GTAP)
*
+ [ xf(rs,f,a,ts)/xf.scale(rs,f,a,ts)
  - sum(fNest_a_f(fNest,a,f),
    p_gf(rs,f,a,ts)*xfNest(rs,f,fNest,ts)/xf.scale(rs,f,a,ts)
    * (pf(rs,f,a,ts)/pfNest(rs,f,fNest,ts))**p_omegafNest(rs,fNest,f))
] $ (fm(f) $ sum(fNest_a_f(fNest,a,f) $ (p_omegafNest(rs,fNest,f) ne inf),1))

--- fully mobile factor: prices are equal across sectors

... factor is linked into top level nest
+ [
  pf(rs,f,a,ts) - pft(rs,f,ts)

] $ ((fm(f) $ (p_omegaf(rs,f) eq inf)) $ ((not sum(fNest_a_f(fNest,a,f),1)) or fNest_a_f("xft",a,f)))

... factor is linked into sub-nest
+ [
  pf(rs,f,a,ts) - sum(fNest_a_f(fNest,a,f), pfNest(rs,f,fNest,ts))

] $ (sum(fNest_a_f(fNest,a,f) $ (p_omegafNest(rs,fNest,f) eq inf),1))

--- immobile factors fnm, with factor supply elasticity
+ [
  xf(rs,f,a,ts)s(xf,'rs,f,a,ts') - p_gf(rs,f,a,ts)s(xf,'rs,f,a,ts')

  $$iftheni.noExpansionXF "%noExpansionXF%"=="on"

  *[( (pf(rs,f,a,ts)/pwfact(ts))**p_etaff(rs,f,a) + 1
    - Sqrt( Sqr( (pf(rs,f,a,ts)/pwfact(ts))**p_etaff(rs,f,a) - 1 + Sqr(1.E-4) ) ) )/2]
    $ p_etaff(rs,f,a) + 1 $ (not p_etaff(rs,f,a))]

  $$else.noExpansionXF

  * [

    --- these is min[ min( p**5,p**etaff), max(p**5,p**etaff) ]
    if p>1, etaft will be chosen, if p<1 the higher elasticity of 5 is active. This will let the supply of the
    natural resource drop quickly if its rent drops, stabilizing the price of the output

    ncpVusin[ ncpVUsin( (pf(rs,f,a,ts)/pwfact(ts))**5, (pf(rs,f,a,ts)/pwfact(ts))**p_etaff(rs,f,a)),
      -ncpVUsin(-(pf(rs,f,a,ts)/pwfact(ts))**5,-(pf(rs,f,a,ts)/pwfact(ts))**p_etaff(rs,f,a)),1.E-7]
      $ p_etaff(rs,f,a)
    + 1 $ (not p_etaff(rs,f,a))]
  $$endif.noExpansionXF
] $ ((fnm(f) $ (not sum(fNest_a_f(fNest,a,f) $ (p_omegafNest(rs,fNest,f) eq inf),1)))

```

The case of sector-specific factor demands, used for natural resources, there are special mechanisms for factor price depending supply. If “%noExpansionXF%” is switched on, resource use by each activity is upper bounded by the parameter  $p\_gf$ , with a fudged min operator. This is seldom used. It mimics a MCP implementation with a maximum on  $xf$ . With the price being 1 at the benchmark, the maximum use is not reached with prices  $< 1$ . Once prices exceed unity, no further quantities are supplied.

The second case is the default. It uses the given elasticity  $p\_etaff$  in case of factor price increases ( $pf/pWfact > 1$ ) and an elasticity of 5 otherwise. This will mean that the extraction of natural resources (oil fields, mines, fish stocks etc.) will drop with a high elasticity of five in case of falling demand for the factor, i.e. if rents drop, stabilizing unit costs. As long as the demand for the output of the activity using the resource is inelastic, this will increase revenues. Reducing resource abstraction of natural resources when their rents drop is hence a likely response of firms.



The **agent specific factor prices tax inclusive** are defined via the equation  $e\_pfa$  and the macro  $m\_pfa$ , equation (73) in VDM 2018:

```

--- Agent's price of factors, equation (73)

e_pfa(rs,f,aIn(a),ts) $ xfflag(rs,f,a) ..

pfa(rs,f,a,ts) =E= %rr_start% m_pfa(rs,f,a,ts) %rr_end%;

```

Note the macros `%rr_start%` and `%rr_end%` that will introduce a summation of the nations for sub-national units to introduce the factor subsidy and tax rate of the nation.

The macro  $m\_pfa$  distinguishes the case of finite transformation of factor supply and the case of full factor mobility, i.e. infinite transformation and adds the national tax rates with the macro  $m\_fctTaxes$ , i.e. subsidy rates  $fctts$ , tax rates  $fcttx$  and an economy factor tax shifter  $fcttxShift$ . The relevant price `%rPrice%` is either equal to  $rs$  when no sub-regions are present or otherwise equal to  $rsNat$ .

```

*
* --- factor price of agents, equation (73) in VDM 2018
*
*$macro m_pfa(r,f,a,t) { \
*
*   --- factors are immobile or sluggish: price before taxes differs by sector
*
*   (pf(r,f,a,t) $ (fm(f) or (p_omegaf(r,f) ne inf)) \
*
*   --- fully mobile factor: all sector face a uniform price before taxes
*
*   + pft(r,f,t) $ (fm(f) $ (p_omegaf(r,f) eq inf))) \
*
*   --- add ad-valorem factor taxes
*
*   * m_fctTaxes(%rPrice%,f,a,t) \
*
* }

```

The factor tax macro  $m\_fctTaxes$  defines these taxes ad-valorem tax and can add firstly carbon taxes related to carbon stock changes  $emispLnd$  for land use and second, carbon taxes  $emisp$  on factor use driven by emission factors per unit of factor use  $emifCo2eqTaxed$ . For details on emission accounting and taxing, see the chapter “CO2, Non-CO2 Accounting and air pollution emissions, energy use accounting”.

```

*$macro m_fctTaxes(r,f,a,t) ( 1 + fctts(r,f,a,t) + fcttx(r,f,a,t) + fcttxShift(r,t) \
*   $ fTaxShift(f,a) \
*   $ $if1 "%ModulesCO2_Emissions%"_modulesGTAP_AEZ%"="on_on" + emispLnd(r,a,t) $ lnd(f) \
*   ) \
*   $ $if1 "%ModulesCO2_Emissions%"="on" + emisp(r,t) * emifCo2eqTaxed(r,f,a,t)*(1-v_redEmiaCo2Eq(r,a,t) + p_emisShift(r,f,a,t)) \
*   )

```

The “Flexible nesting” approach allows introducing factor supply nests which can also be linked into other factor supply nests. The equation  $e\_pfNest$  defines the average factor price of such a nest. It distinguishes three cases: Leontief, finite factor transformation and infinite one. In the first block, the average price  $pfNest$  is defined via a dual price aggregator, taking the share parameters ( $p\_gf$  for factors and  $p\_gfNest$  for sub-nest), the prices ( $pf$  for factors and  $pfNest$  for sub-nests) and the transformation elasticity  $p\_omegaFNest$  into account. The case of infinite transformation is handled by the last block. Here, the price is indirectly defined from the adding up-condition of the factor quantities. This equation is not part of the GTAP standard model.

```

*
* ---- CET supply sub-nests (not part of standard GTAP model)
*
* e_pfNest(rs,f,fNest,ts) $ p_gfNest(rs,f,fNest,ts) ..
*
* 0 =e=
*
* --- sluggish factor mobility, dual price aggregator for sub-nests
*
* [ pfNest(rs,f,fNest,ts)s(pfNest,'rs,f,fNest,ts')
*
* --- factors linked into sub nest
*
* - [ sum(fNest_a_f(fNest,a,f)
*       %scale% $ p_gf(rs,f,a,ts)
*       , p_gf(rs,f,a,ts)*pf(rs,f,a,ts)**(1+p_omegafNest(rs,fNest,f)))
*
* --- subnests linked into sub nest
*
* + sum(fNest_n_f(fNest,fNest1,f)
*       %scale% $ p_gfNest(rs,f,fNest1,ts)
*       , p_gfNest(rs,f,fNest1,ts)*pfNest(rs,f,fNest1,ts)**(1+p_omegafNest(rs,fNest,f)))
*
* ]**(1/(1+p_omegafNest(rs,fNest,f)))s(pfNest,'rs,f,fNest,ts')
*
* ] $ ( p_omegafNest(rs,fNest,f) and (p_omegafNest(rs,fNest,f) ne inf))
*
* + [ pfNest(rs,f,fNest,ts)s(pfNest,'rs,f,fNest,ts')
*
* --- factors linked into sub nest
*
* - sum(fNest_a_f(fNest,a,f)
*       %scale% $ p_gf(rs,f,a,ts)
*       , p_gf(rs,f,a,ts)*pf(rs,f,a,ts))s(pfNest,'rs,f,fNest,ts')
*
* --- subnests linked into sub nest
*
* - sum(fNest_n_f(fNest,fNest1,f)
*       %scale% $ p_gfNest(rs,f,fNest1,ts)
*       , p_gfNest(rs,f,fNest1,ts)*pfNest(rs,f,fNest1,ts))s(pfNest,'rs,f,fNest,ts')
*
* ] $ (p_omegafNest(rs,fNest,f) eq 0)
*
* --- fully mobile factors, uniform prices => physical aggregation
*
* + (xfNest(rs,f,fNest,ts)s(xfNest,'rs,f,fNest,ts')
*   - sum(fNest_a_f(fNest,a,f) $ p_gf(rs,f,a,ts), xf(rs,f,a,ts))s(xfNest,'rs,f,fNest,ts')
*   - sum(fNest_n_f(fNest,fNest1,f) $ p_gfNest(rs,f,fNest1,ts), xfNest(rs,f,fNest1,ts))s(xfNest,'rs,f,fNest,ts')
*   ) $ (p_omegafNest(rs,fNest,f) eq inf)
*
* ;

```

The total supply of a factor to such a nest  $xfNest$  as defined in the equation  $e_{xfNest}$  considers the two cases of finite and infinite transformation elasticity. In case of finite transformation in the first block, total supply either depends on the sector wide supply of that factor of  $xft$  and the price relations or on the amount supplied to the nest  $fNest1$  to which the sub-nest belongs. In case of infinite transformation, the sub-nest price is either equal to the sector-wide factor price  $pft$  or to the price of the upper nest factor nest price  $pfNest$  indexed with  $fNest1$ .

```

e_xfNest(rs,f,fNest,ts) $ (p_gfNest(rs,f,fNest,ts) $ (sum(fNest_n_f(fNest1,fNest,f),1) or sum(fNest_n_f("xft",fNest,f),1))
  $ (not isControlledByMyGtap(rs,fNest,f)) ) ..
*
* 0 =e=
*
* --- "sluggish" factor supply from top level nest (xft,pft) to current nest
*
* ( [xfNest(rs,f,fNest,ts)s(xfNest,'rs,f,fNest,ts')
*   - p_gfNest(rs,f,fNest,ts) * xft(rs,f,ts)s(xfNest,'rs,f,fNest,ts')
*   * (pfNest(rs,f,fNest,ts)/pft(rs,f,ts))**p_omegaf(rs,f)
*   ] $ p_omegaf(rs,f)
*
* + [xfNest(rs,f,fNest,ts)s(xfNest,'rs,f,fNest,ts')
*   - p_gfNest(rs,f,fNest,ts) * xft(rs,f,ts)s(xfNest,'rs,f,fNest,ts')
*   ] $ (p_omegaf(rs,f) eq 0)
*
* ) $ ( (p_omegaf(rs,f) ne inf) $ fNest n f("xft",fNest,f))

```

```

*
*
*      --- "sluggish" factor supply from another nest fNest1 to the current nest
*
+ [
  { sum(fNest_n_f(fNest1,fNest,f), xfNest(rs,f,fNest,ts)s(xfNest,'rs,f,fNest,ts')
    - p_gfNest(rs,f,fNest,ts) * xfNest(rs,f,fNest1,ts)s(xfNest,'rs,f,fNest,ts')
    * (pfNest(rs,f,fNest,ts)/pfNest(rs,f,fNest1,ts)) ** p_omegafNest(rs,fNest1,f))
  } $ p_omegaf(rs,f)
+
  { sum(fNest_n_f(fNest1,fNest,f), xfNest(rs,f,fNest,ts)s(xfNest,'rs,f,fNest,ts')
    - p_gfNest(rs,f,fNest,ts) * xfNest(rs,f,fNest1,ts)s(xfNest,'rs,f,fNest,ts'))
  } $ (p_omegaf(rs,f) eq 0)

] $ ( sum(fNest_n_f(fNest1,fNest,f) $ (p_omegafNest(rs,fNest1,f) ne inf),1))

*
*      --- non "sluggish" factor supply from top level nest (xft,pft) to current nest
*
+ ( pfNest(rs,f,fNest,ts)s(pfNest,'rs,f,fNest,ts') - pft(rs,f,ts)s(pfNest,'rs,f,fNest,ts'))
  $ ( (p_omegaf(rs,f) eq inf) $ fNest_n_f("xft",fNest,f))

*
*      --- non "sluggish" factor supply from another nest to current nest
*
+ sum(fNest_n_f(fNest1,fNest,f) $ (p_omegafNest(rs,fNest1,f) eq inf),
  pfNest(rs,f,fNest,ts)s(pfNest,'rs,f,fNest,ts') - pfNest(rs,f,fNest1,ts)s(pfNest,'rs,f,fNest,ts') )
;

```

The factor supply from nation to sub-regions is described in the section “Integration into the modeling framework”, pages 151ff of the chapter “Sub-national dis-aggregation of production and factor markets in CGEBox”. Note also that the “GTAP-AEZ” module (pages 293ff) will introduce land transformation at the level of Agro-Ecological Zones and replace some of the equations detailed above.

## 2.6 Income generation and distribution - overview

The default approach to income generation and distribution is the regional household approach found in the GTAP standard model. It is switched off if the *myGTAP* module is added to the model. An overview on the regional household approach is depicted in Figure 2 below. Regional income is sourced (1) by factor income *facty* (factor remuneration including direct taxes, but net of taxes paid by firms on factor use) minus depreciation (*valDep*) and (2) by indirect taxes *yTaxInd*, i.e. all tax flows *yTaxTot* minus direct taxes *yTax<sub>dt</sub>* and transfers *yTax<sub>tg</sub>* which are already comprised in the factor income. Regional income plus depreciation is equal to GDP at market prices.

Regional household income *regy* is distributed to final demand expenditures of an aggregate private household *yc*, government *yg* and regional net savings *rsav*. Adding the values of depreciation *valDep* and of foreign savings *valSavf* to regional net savings *rsav* yields investment demand expenditures *yi*. The distribution of the final demand expenditures to the Amington demands for each product  $x_{ai}$  is based on CES demand systems for investments and the government. Depending on the chosen substitution elasticity, they can hence also encompass the CD or Leontief case. For private households, a CDE, CD, LES or (M)AIDADS demand system can be used to distribute private household expenditure *yc*.

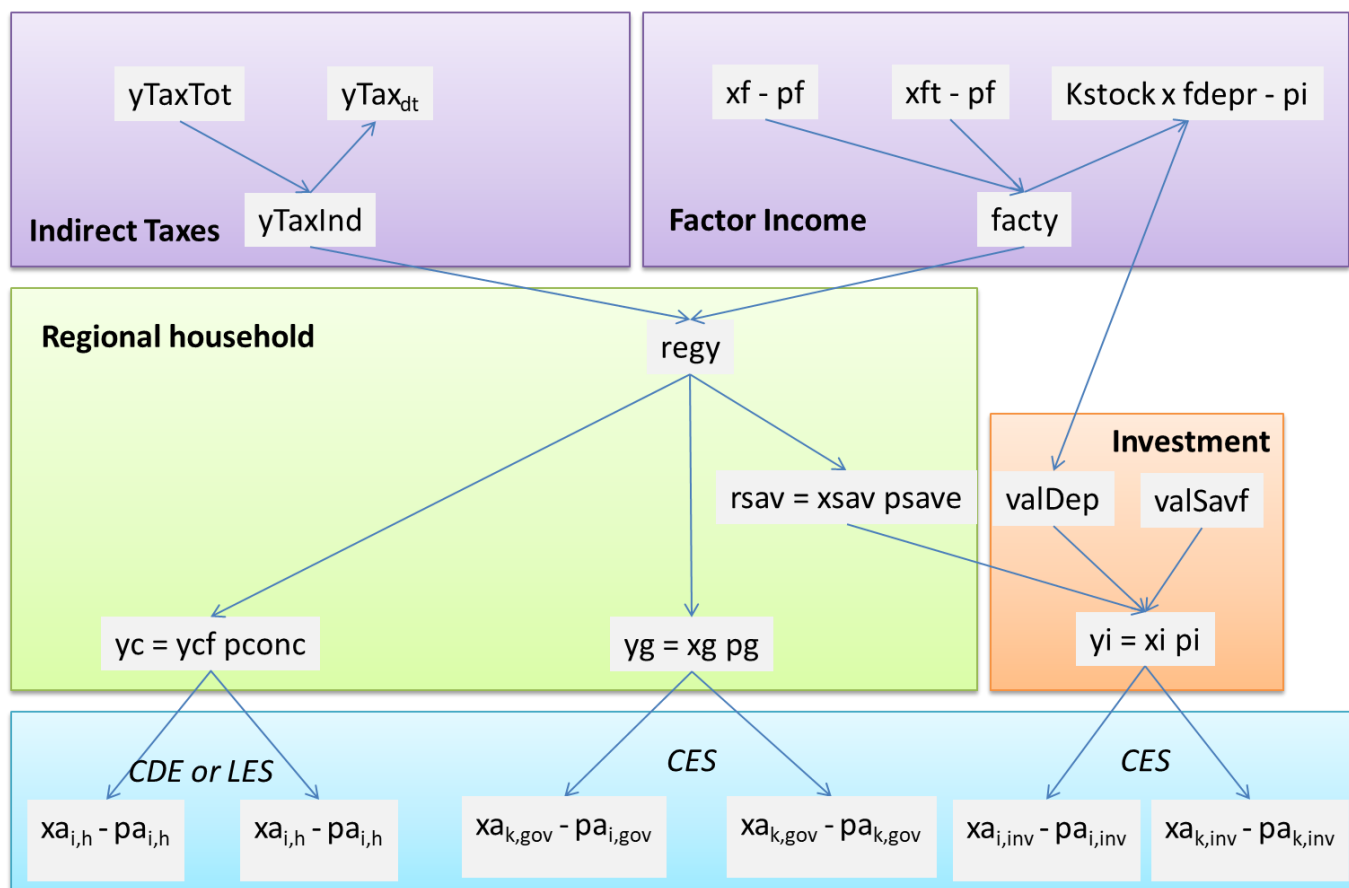


Figure 2: Overview in income generation and distribution

## 2.7 Income generation

**Regional income**  $regy$ , i.e. economy wide income which can be spent on regional savings net of depreciation and final consumption by government and private households, is generated from factor income including direct taxes  $facty$  and indirect taxes  $yTaxInd$  as defined in the  $e\_regy$ , equation (26) in VDM 2018. The CGEBox implementation considers three further potential incomes flows discussed next. Note that regional income plus depreciation is equal to GDP at market prices.

```

--- regional income, equation (26) in VDM 2018
e_regy(rsNat,ts) $ (regy.range(rsNat,ts) ne 0) ..

regy(rsNat,ts)s(regy,'rsNat,ts') =e= (factY(rsNat,ts)
+ yTaxInd(rsNat,ts)
- invFromCapInc(rsNat,ts)
+ finc(rsnat,ts)
- p_capTrans(rsNat,ts)*pnun(ts))s(regy,'rsNat,ts');
```

The model allows setting a parameter  $p\_invFromCapIncShr$  which directly convert part of returns to capital into savings, for instance, to cover depreciation):

```

e_invFromCapInc(rsNat,ts) $ (invFromCapInc.range(rsNat,ts) ne 0) ..

invFromCapInc(rsNat,ts)s(invFromCapInc,'rsNat,ts')
=e= sum(cap $ sum(a,xftFlag(rsNat,ts,a)),vft(rsNat,ts,ts))
* p_invFromCapIncShr(rsNat,ts)s(invFromCapInc,'rsNat,ts');
```

The foreign income exchange variable *finc* is fixed to zero in the standard model, but can be activated with the *myGTAP* model. The parameter *p\_capTrans* captures exogenous income exchanges with the rest-of-the-world as part of the debt accumulation mechanism of G-RDEM. *Finc* and *p\_captrans* are further elements in the Balance of Payments not found in the standard GTAP model.

**Factor income** including direct taxes *factY* is defined by the equation *e\_factY*, equivalent to equation (25) in VDM 2018. It considers returns to primary factors net of factor taxes defined by the variable *vft*. Note that factor income comprises direct taxes and government net transfers to households. As returns to capital also cover depreciation, the value of depreciation is deducted:

```
--- Factor income, net of depreciation, equation (25) in VDM 2018

e_factY(rsNat,ts) $ (factY.range(rsNat,ts) ne 0) ..

    factY(rsNat,ts) s(factY,'rsNat,ts')
    =e= [ sum(f $ sum(a,xfFlag(rsNat,f,a)),vft(rsNat,f,ts))
          - valDep(rsNat,ts) ] s(factY,'rsNat,ts') ;
```

**Indirect tax income** *yTaxInd* is calculated by the equation *e\_ytaxInd* from total tax revenues *ytaxTot*, corrected for direct taxes (index “dt”) and transfers (index “tg”), two tax streams that are comprised in factor income *factY* as defined above, see equation (24) in VDM 2018:

```
e_ytaxInd(rsNat,ts) ..

yTaxInd(rsNat,ts) s(yTaxInd,'rsNat,ts') =e= (ytaxTot(rsNat,ts) - ytax(rsNat,"dt",ts) - ytax(rsNat,"tg",ts)) s(yTaxInd,'rsNat,ts');
```

**Total tax income** *yTaxTot* considers all tax flow *gy* depicted in the model (see next equation) and is defined by the equation *e\_ytaxTot*, equivalent to equation (23) in VDM 2018:

```
--- Total tax revenue, equation (23) in VDM 2018

e_ytaxTot(rsNat,ts) ..

ytaxTot(rsNat,ts) s(yTaxTot,'rsNat,ts') =e= sum(gy, ytax(rsNat,gy,ts)) s(yTaxTot,'rsNat,ts');
```

**Different tax flows** *yTax* for the different types of tax flow *gy* are defined by the equation *e\_ytax*, defined in equation (13)-(22) in VDM 2018. The tax flows fall in the following blocks:

1. **Production taxes** “pt”, levied with the relative tax rate *prdtX* on output value, i.e. output *m\_xp* times the related producer price *px* in case of the case of a diagonal production activity and output and product specific prices *p* and quantities *x* in case of a non-diagonal one. Note that fix cost might be present in the model if the Melitz/Krugmann extension is used, depicted by the technology nest “top” on which also production taxes are charged. Rates of production taxes are currently not differentiated across sub-national units.

```
--- Tax revenues from production tax, equation (13) in VDM 2018

+ sum((a,i)
    $scale% $ (prdtX.1(rsNat,a,i,ts) $ xxFlag(rsNat,a,i))
    ,prdtX(rsNat,a,i,ts) * (
        (p(rsNat,a,i,ts) + x(rsNat,a,i,ts)) $ (not diag(a))
        + (px(rsNat,a,ts) + m_xp(rsNat,a,ts)) $ diag(a)
    )
    $ $if "%ModulesGTAP_Melitz%"=="ON" + sum(tNest $ (tNest_n_a("Top",tNest,a) $ (not (sameas(tNest,"VA") or sameas(tNest,"ND")))),
    $ $if "%ModulesGTAP_Melitz%"=="ON" xtNest(rsNat,tNest,a,ts) * ptNest(rsNat,tNest,a,ts)
    )
    ) $ sameas(gy,"pt")
```

2. **Indirect taxes on private consumption “*pc*”**, levied with rate *dintx* on domestic consumption of private household *xd* and *xn* times the related domestic price *m\_domPrice*, and with rate *mintx* on imports by private households *m\_xm* times the average price of imports *m\_pmt*:

```

$$iftheni.hsubr %hSubr%==on
+ [sum((r_x(rs,rsNat),i,h),
+ [ (dintx(rs,i,h,ts) + itxshft(rs,i,ts)) * m_domPrice(rs,i,h,ts)*xd(rs,i,h,ts) ] $ (m_alphad(rs,i,h,ts) gt eps)
+ [ (dintx(rsNat,i,h,ts) + itxshft(rsNat,i,ts)) * m_domPrice(rsNat,i,h,ts)*xn(rs,i,h,ts) ] $ (m_alphan(rs,i,h,ts) gt eps)
+ [ (mintx(rs,i,h,ts) + itxshft(rs,i,ts)) * m_pmt(rsNat,i,h,ts)*xm(rs,i,h,ts) ] $ (m_alpham(rs,i,h,ts) gt eps)
] $ sameas(gy,"pc")
$$else.hsubr
+ [sum((i,h),
+ [ (dintx(rsNat,i,h,ts) + itxshft(rsNat,i,ts)) * m_domPrice(rsNat,i,h,ts)*xd(rsNat,i,h,ts) ] $ (m_alphad(rsNat,i,h,ts) gt eps)
+ [ (mintx(rsNat,i,h,ts) + itxshft(rsNat,i,ts)) * m_pmt(rsNat,i,h,ts)*xm(rsNat,i,h,ts) ] $ (m_alpham(rsNat,i,h,ts) gt eps)
] $ sameas(gy,"pc")
$$endif.hsubr

```

Note the product specific general tax shifter *itxShft* which can be applied to multiple products which is normally not active and allows for specific closure swaps. The macro *m\_domprice* shown below in the equation above differentiates between products subject and not subject to the Melitz module. Non-Melitz products (second line in if-clause case and macro definition in else clause) use the supply price *ps* in case of infinite transformation between domestic sales and exports (*p\_omegax* is infinite). For finite transformation (*p\_omegax* not infinite), the price of domestic origin *pd* is applied. For Melitz products, two cases are differentiated depending on the indicator parameter *p\_melDomUse*, see Melitz module for detail.

```

$$iftheni.mel "%melitz%"=="on"
$$macro m_domPrice(r,i,aa,t) [ (pd(r,i,t) $ (p_omegax(r,i) ne inf) + ps(r,i,t) $ (p_omegax(r,i) eq inf)) $ (not iMel(i)) \
+ (pd(r,i,t) $ p_melDomUse(r,aa,i) + m_pfirm(r,i,"domes",t) $ (not p_melDomUse(r,aa,i))) $ iMel(i) ]
$$else.mel
$$macro m_domPrice(r,i,aa,t) (pd(r,i,t) $ (p_omegax(r,i) ne inf) + ps(r,i,t) $ (p_omegax(r,i) eq inf))
$$endif.mel

```

Two other blocks apply the very same logic to **government and investment consumption**, resulting in tax flows “*gc*” and “*ic*”.

```

* --- Indirect tax revenues from government consumption, equation (16) in VDM 2018
+ [sum((i,gov), [ (dintx(rsNat,i,gov,ts) + itxshft(rsNat,i,ts)) * m_domPrice(rsNat,i,gov,ts)*xd(rsNat,i,gov,ts) ] $ m_alphad(rsNat,i,gov,ts)
+ [ (mintx(rsNat,i,gov,ts) + itxshft(rsNat,i,ts)) * m_pmt(rsNat,i,gov,ts)*xm(rsNat,i,gov,ts) ] $ m_alpham(rsNat,i,gov,ts)
] $ sameas(gy,"gc")
* --- Indirect tax revenues from investment consumption, equation (17) in VDM 2018
+ [sum((i,inv), [ (dintx(rsNat,i,inv,ts) + itxshft(rsNat,i,ts)) * m_domPrice(rsNat,i,inv,ts)*xd(rsNat,i,inv,ts) ] $ m_alphad(rsNat,i,inv,ts)
+ [ (mintx(rsNat,i,inv,ts) + itxshft(rsNat,i,ts)) * m_pmt(rsNat,i,inv,ts)*xm(rsNat,i,inv,ts) ] $ m_alpham(rsNat,i,inv,ts)
] $ sameas(gy,"ic")

```

3. **Direct taxes “*dt*”** are levied with factor specific rates *kappaF* on factor income, i.e. factor prices *pf* times factor use *xf* respectively. An endogenous or exogenous direct tax shifter *kappaf* can be added:

```

--- Direct tax revenues, equation (22) in VDM 2018
+ sum((f,a)
+ [scale% $ xfFlag(rsNat,f,a)
+ [kappaf(rsNat,f,ts)+kappaShft(rsNat,ts)] * pf(rsNat,f,a,ts)*xf(rsNat,f,a,ts)
] $ sameas(gy,"dt")

```

4. **Transfers from the government account to households “*tg*”** are levied with ad-valorem rate *kappaFG* on total regional income. The tax rate *kappafG* is endogenous in the myGTAP module where a dedicated equation determines the level of transfers from the government account to different households:

```

--- Gov transfers (determined in myGTAP module if active)
+ kappafG(rsNat,ts)*regy(rsNat,ts)
$ sameas(gy,"tg")

```

5. **Export tax revenues** “*et*” are based on bi-lateral export tax rates *exptx* and potentially a commodity specific export tax shifter *etax*. Note that depending on how exports are depicted (infinite transformation or not), different prices are used: (a) the bi-lateral export price *pe* if there is finite transformation across model regions, (b) the average export price *pet* in case of infinite transformation across different model regions, but not between exports and domestic sales and (c) the supply price *ps* if all transformations are infinite. Finally, if the Melitz module is active, the firm price defined in the macro *m\_pFirm* is used. The related bi-lateral export quantity index is defined in the macro *m\_xws*.

```
*      --- Export tax revenues, equation (21) in VDM 2018

+      [sum((i,rNat1) $ ( (exptx.1(rsNat,i,rNat1,ts) + etax.1(rsNat,i,ts)) $ m_xwFlag(rsNat,i,rNat1) ),
        (exptx(rsNat,i,rNat1,ts) + etax(rsNat,i,ts))
        * [ (
            pe(rsNat,i,rNat1,ts) $ psFlag(rsNat,i,"pe")
          + pet(rsNat,i,ts)      $ psFlag(rsNat,i,"pet")
          + ps(rsNat,i,ts)       $ psFlag(rsNat,i,"ps")
        )
        $$ifthen.iMel "%melitz%"=="on"
        $ (not iMel(i))
        + (m_pFirm(rsNat,i,rNat1,ts) $ iMel(i))
        $$endif.iMel
        ]
        *m_xws(rsNat,i,rNat1,ts))] $sameas(gy,"et")
```

6. **Import taxes** “*mt*” are defined from bi-lateral import taxes *imptx*, a commodity specific import tax shifter *mtax*, the bi-lateral c.i.f. prices defined by the global %pmcif% and the bi-lateral flows *xw*.

```
*      --- Import tax revenues, equation (20) in VDM 2018

+      [sum((i,rNat1) $ m_xwFlag(rNat1,i,rsNat), (imptx(rNat1,i,rsNat,ts) + mtax(rsNat,i,ts))
        *%pmcif%(rNat1,i,rsNat,ts)*xw(rNat1,i,rsNat,ts))] $sameas(gy,"mt")
```

7. **Factor taxes** “*ft*” paid by each activity *a* are levied on the activity specific factor price *pf* and use *xf* with the rates *fcttx* and a model region specific factor tax shifter *fcttxShift*:

```
*      ---- Tax revenues from factor taxes, equation (18) in VDM 2018

+      [sum((f,a) $ xfFlag(rsNat,f,a), (fcttx(rsNat,f,a,ts) + fcttxShift(rsNat,ts) $ fTaxShift(f,a))
        * pf(rsNat,f,a,ts) * xf(rsNat,f,a,ts))] $sameas(gy,"ft")
```

8. The same logic (less the shifter) applies for factor subsidies “*fs*”:

```
*      ---- Revenue costs from factor subsidies, equation (19) in VDM 2018

+      [sum((f,a) $ xfFlag(rsNat,f,a), fctts(rsNat,f,a,ts)
        * pf(rsNat,f,a,ts) * xf(rsNat,f,a,ts))] $sameas(gy,"fs")
```

9. Finally, **emission taxes** “*emis*” can be introduced, levied on CO2 and non-Co2 emissions in CO2 equivalents *emisTaxed* with the potentially endogenous price *emisP*. In case of the AEZ module being active, an extension under development allows taxing carbon stocks, converted with factor 3.67 to CO2.

```
$$iftheni.co2 "%ModulesCO2_Emissions%"=="on"
+      [(emisTaxed(rsNat,ts)

      --- if the carbon stock grows, there are no subsidies paid to un-managed land: we have to subtract
      the change in carbon, converted to CO2 equivalents, from the tax base

      $$if i "%modulesGTAP_AEZ%"=="on" + [dCarbon(rsNat,ts)-dCarbonBuiltup(rsNat,ts)+dCarbonTaxedN(rsNat,ts)]*0.001*3.67
      $$if i "%modulesGTAP_AEZ%"=="on" $ sum(a $ ((emisPLnd.range(rsNat,a,ts) ne 0) $ xfFlag(rsNat,lnd,a)),1)

      ) * emisP(rsNat,ts) * emisPShr(rsNat,ts) ] $ sameas(gy,"emis")

$$endif.co2
```



## 2.8 Income distribution

The main elements with regard to final demand are as follows. Fully flexible multi-layer CES nestings under the final agent's top-level demand system with support of the volume-reserving CES variant. Multiple households can be depicted with the myGTAP module with support of sub-national households and demand for output from the same sub-national region (subsistence production). Choice between functional form for households (CDE, MAIDADS, LES, CD). CD or CES functions for savings and government demand.

If the regional household approach is used, savings, government and household demand are distributed based on a modified CD utility function where the private household demand share is driven by the elasticity of total private expenditure with regard to utility  $\phi_iP$  and the original private demand share  $\beta_iP$ , part of equation (28) of VDM 2018:

```
*
*   --- intermediate terms to help solver, part of equation (28) in VDM 2018
*
e_betaPhi(rsNat,ts) ..
    betaPhi(rsNat,ts) =E= betaP(rsNat,ts)/phiP(rsNat,ts);
```

The updated share parameter  $\beta_iP$  related to private consumption is termed  $\beta_i\phi_i$ . Updates imply that the shares as defined in the benchmark do not add to unity any longer. Therefore, an intermediate variable  $\phi_i\text{RegY}$  is defined which scales regional income to reflect the updated sum of the shares. Assume that  $\beta_i\phi_i$  is increased compared to the benchmark. That implies that the second term on the LHS exceed unity. The correction factor  $\phi_i\text{RegY}$  is accordingly proportionally decreased such that the three expenditure shares still sum up to unity and maintain the exhaustion condition, part of equation (28) in VDM 2018:

```
e_phiRegY(rsNat,ts) ..
    phiRegY(rsNat,ts)*(betaPhi(rsNat,ts) + betaG(rsNat,ts) + betaS(rsNat,ts))
    =E= 1;
```

Beside the CD case depicted in the GTAP Standard model, CGEBox can consider the general CES case where a price index  $pu$  for top-level utility is defined as follows:

```
*
*   --- price index related to CES utility function
*
e_pu(rsNat,ts) $ (p_sigmaU(rsNat) ne 1) ..
*
    pu(rsNat,ts) =E= ( betaPhi(rsNat,ts)*phiRegY(rsNat,ts) * pCons(rsNat,"hhsld",ts)**(1-p_sigmaU(rsNat))
    + betas(rsNat,ts) * phiRegY(rsNat,ts) * pSave(rsNat,ts) ** (1-p_sigmaU(rsNat))
    + betag(rsNat,ts) * phiRegY(rsNat,ts) * pg(rsNat,ts) ** (1-p_sigmaU(rsNat))) ;
```

The amount spent for private consumption  $yc$  is defined in the equation  $e_{yc}$ , (equation (29) in VDM 2018 for the first CD related part). Note that in the case of the CES function, the exponent of  $pCons$  adds unity to multiply from the defined quantity the spendings.

```
--- Total private consumption expenditure, equation (29) in VDM 2018
e_yc(rsNat,h,ts) ..
    yc(rsNat,h,ts)s(yc,'rsNat,h,ts')
    =E= [ betaPhi(rsNat,ts)*phiRegY(rsNat,ts) * regy(rsNat,ts)s(yc,'rsNat,h,ts') ] $ (p_sigmaU(rsNat) eq 1)
    + [ betaPhi(rsNat,ts)*phiRegY(rsNat,ts) * (regy(rsNat,ts)/pu(rsNat,ts))s(yc,'rsNat,h,ts')
    * pCons(rsNat,"hhsld",ts)**(1-p_sigmaU(rsNat))] $ (p_sigmaU(rsNat) ne 1) ;
```

The amount spent for government consumption  $yg$  is defined in the equation  $e_{yg}$  (equation (30) in VDM 2018 for the first CD related part):



```

--- Total government consumption expenditure, equation (30) in VDM 2018

e_yg(rsNat,ts) ..
    [ xg(rsNat,ts) $ (p_sigmaU(rsNat) ne 1) + yg(rsNat,ts) $ (p_sigmaU(rsNat) eq 1)]s(xg,'rsNat,ts')
    =e= [betaG(rsNat,ts)*phiRegy(rsNat,ts) * regy(rsNat,ts)s(xg,'rsNat,ts')] $ (p_sigmaU(rsNat) eq 1)
        + [betaG(rsNat,ts)*phiRegy(rsNat,ts) * (regy(rsNat,ts)/pu(rsNat,ts))s(xg,'rsNat,ts')
            * pg(rsNat,ts)**(-p_sigmaU(rsNat)))] $ (p_sigmaU(rsNat) ne 1) ;

```

And finally, the quantity index  $xsav$  respectively spending  $rsav$  of regional savings is depicted in the equation  $e\_rsav$  (equation (31) in VDM 2018 for the first CD related part):

```

--- Total nominal saving, equation (31) in VDM 2018

e_rsav(rsNat,ts) ..
    [ xsav(rsNat,ts) $ (p_sigmaU(rsNat) ne 1) + rsav(rsNat,ts) $ (p_sigmaU(rsNat) eq 1)]s(xsav,'rsNat,ts')
    =e= [ betaS(rsNat,ts) * phiRegy(rsNat,ts) * regy(rsNat,ts)s(xsav,'rsNat,ts')] $ (p_sigmaU(rsNat) eq 1)
        + [ betaS(rsNat,ts) * phiRegy(rsNat,ts) * (regy(rsNat,ts)/pu(rsNat,ts))s(xsav,'rsNat,ts')
            * pSave(rsNat,ts)**(-p_sigmaU(rsNat)))] $ (p_sigmaU(rsNat) ne 1) ;

```

Note that the regional household approach can be replaced by separate accounts for households and the government; see the sub-section “Model equations” in the section “Capital vintages

## 2.8.1 Summary

Capital vintages dis-aggregate a production activity into two variants: one working with new capital generated in the same period from investments and an old one working with capital generated in earlier period. The equations in the model relating to the capital vintages are described in the section “Production block”.

## 2.8.2 Initialization

The initialization of the capital vintages is performed in ‘*cal/cal\_vint.gms*’. Based on the user selection on the GUI (see section below), the set  $aVint(a)$  comprises the data base activities dis-aggregated into an old and new vintage, showns on the sets  $aOld$  and  $aNew$  and their union set  $aVints$ . Activities outputting products in the SPE and Melitz model extensions are removed from the set  $aVint$  as defined selection by the user, warnings are added to the listing as shown below for the case of the Melitz extension:

```

$ifthen.mel defined iMel
set remVintA(a);
remVintA(aVint) $ sum((rnat,aVint(aVint,aVints),iMel),sam0(rnat,aVint,iMel)) = YES;

if ( card(remVintA),
    $gbatinclude 'util/final_log' "Warning in %system.fn%, line: %system.incline%: using vintage activities for Melitz products not possible. These activities are set back to standard "
    display "Warning in %system.fn%, line: %system.incline%: sing vintage activities for Melitz products not possible. . Such activities are removed ",remVintA;
    vintageA_sel(remVintA) = no;
);
aVint(a) $ sum((rnat,iMel),sam0(rnat,a,iMel)) = no;

```

The remaining data base activities  $aVint$  are dropped from the SAM and from other data referring to production activities based on the set  $aDrop$ :

```

*
*   --- list of activities to drop from SAM and other data parameters
*
aDrop(aVint) = yes;
*

```

While all other activities (not a new or old capital vintages or an aggregate thereof) are kept according to the set  $aKeep$ :

```

*
*   --- list of activities to keep in SAM and other data parameters
*
  aKeep(a)      = yes;
  aKeep(aDrop)  = no;
  aKeep(aNew)   = no;
  aKeep(aOld)   = no;
*

```

The distribution of the data to the old and new variants at the benchmark is based on the default depreciation rate of 4% per annum times the number of depreciation years. This number is equal to unity for recursive-dynamic analysis and chosen by the user for comparative-static one. The distribution shares are set accordingly:

```

--- distribution shares for old and new vintage

p_distrVint(rs,aNew)      = p_depr(rs,"%t0%")+0.04*(p_nDeprYears("%t0%")-1);

p_distrVint(rs,aOld)      = 1-sum(a_avint(a,aNew) $ a_aVint(a,aOld),p_distrVint(rs,aNew));

```

These shares distribute the given SAM (and other data) to the old and new vintages while the original activities are removed, as shown for the case of the SAM:

```

--- distributes outputs to old and new vintages with distribution keys

SAM0(rs,aVints,is) = sum(a_aVint(a,aVints),SAM0(rs,a,is)*p_distrVint(rs,aVints));
SAM0(rs,aDrop,is)  = 0;

```

Similar statements distribute the data referring to the domestic and imported demand, for factor taxes and subsidies and for production taxes. New and old capital are treated differently: old vintages receive all capital minus the new capital allocated to the new vintage. Activities not split up into the old and new vintage receive both old and new capital and a CES nests that allows for substitution between the two.

The substitution elasticities for the new vintage are set 2.5 higher higher compared to the parameters in a model without the vintage extension for technology nests comprising old or new capital as a production factor. The substitution elasticities for the old vintages are downscaled accordingly to keep the average reaction of the model to changes in price relations:

:

```

--- downscale subst. elas of old vintage such that in average over the two vintages, the given subst.
    elasticity is maintained

p_sigmaNest(rs,tNest,aOld) $ sum(tNest_f_a(tNest, cap, aOld),1)
  = p_sigmaNest(rs,tNest,aOld) * (1 - (1-p_distrVint(rs,aOld))*p_sigmaMult(rs,aOld));

p_sigmaNest(rs,tNest,aNew) $ sum(tNest_f_a(tNest, cap, aNew),1)
  = p_sigmaNest(rs,tNest,aNew) * p_sigmaMult(rs,aNew);

```

```

parameter p_sigmaMult(r,a);
tNest("va") = YES;
tNest("nd") = YES;

--- multiplicative increase in subst. elasticity for nests comprising capital for new vintage is 2.5

p_sigmaMult(rs,aVints) $ xpFlag(rs,aVints) = 2.5;

--- downscale subst. elas of old vintage such that in average over the two vintages, the given subst.
    elasticity is maintained

p_sigmaNest(rs,tNest,aOld) $ sum(tNest_f_a(tNest, cap, aOld), 1)
    = p_sigmaNest(rs,tNest,aOld) * (1 - (1-p_distrVint(rs,aOld)) * p_sigmaMult(rs,aOld));
p_sigmaNest(rs,tNest,aNew) $ sum(tNest_f_a(tNest, cap, aNew), 1)
    = p_sigmaNest(rs,tNest,aNew) * p_sigmaMult(rs,aNew);

```

### 2.8.3 Dynamic updating of share parameters

In recursive-dynamic mode the code considers that a part of the old vintage is removed in each period, due to depreciation while last year's new vintage becomes part of the old vintage. For instance, expected output (stored temporarily on the field .m) is defined as:

```

xp.m(rs,aOld,tSim) $ xp.1(rs,aOld,tSim-1)
    = sum(aVint_aVint(aNew,aOld), xp.1(rs,aNew,tSim-1) * p_nDeprYears(tSim-1))
      + xp.1(rs,aOld,tSim-1)
        * sum(r_r(rs,rsNat), (1 - p_depr(rsNat,tSim-1) * p_nDeprYears(tSim-1)));

```

Similar statements define the expected demand for factors, intermediates and CES nests in the production function. In parallel, the code defines expected prices for the old vintage in the next period as weighted average between last period's old and new vintages, considering as above depreciation. For instance, the unit cost average for the new vintages is defined as:

```

px.m(rs,aOld,tSim) $ xp.m(rs,aOld,tSim)
    = [sum(aVint_aVint(aNew,aOld), xp.1(rs,aNew,tSim-1) * px.1(rs,aNew,tSim-1) * p_nDeprYears(tSim-1))
      + xp.1(rs,aOld,tSim-1) * px.1(rs,aOld,tSim-1)
        * sum(r_r(rs,rsNat), (1 - p_depr(rsNat,tSim-1) * p_nDeprYears(tSim-1)))] / xp.m(rs,aOld,tSim);

```

Based on these update demands for production factors and their prices, the share parameters are updated for the old vintage are updated, for instance for input demands:

```

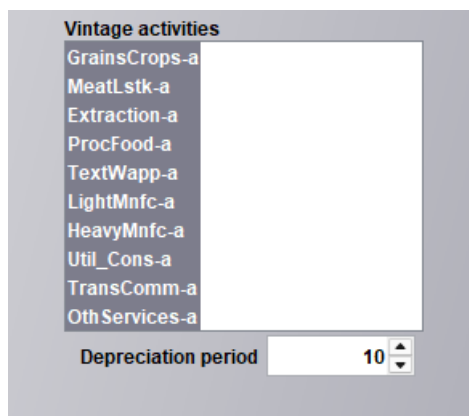
*
* --- redefine share parameters, such that the old vintage would demand at prices and shifters of tSim-1 the
* aggregate quantities of the new and old vintages
*
p_ioNest(rs,i,aOld,iotNest,tSim) $ (xa.m(rs,i,aOld,tSim) $ sum(tNest_i_a(iotNest,i,aOld), p_ioNest(rs,i,aOld,iotNest,tSim) $ (xtNest.m(rs,iotNest,aOld,tSim) > 0)))
    = sum(tNest_i_a(iotNest,i,aOld), (xa.m(rs,i,aOld,tSim) * tNestShare(rs,iotNest,i,aOld) / xtNest.m(rs,iotNest,aOld,tSim))
      * 1/lambdaio(rs,i,aOld,tSim-1) * (([p_sigmaNest(rs,iotNest,aOld)-1] $ (not addCes(rs,iotNest,aOld))
        - p_sigmaNest(rs,iotNest,aOld) $ addCes(rs,iotNest,aOld)))
      * (pa.m(rs,i,aOld,tSim) / ptNest.m(rs,iotNest,aOld,tSim)) * p_sigmaNest(rs,iotNest,aOld));

```

This implies that the technology of the old vintage changes over time. Fast growing sectors with an above-average share of the new vintages will show larger changes in their average technologies.

### 2.8.4 GUI

The GUI allows the user which production activities are broken down two vintages and to set the depreciation period in case of a comparative-static analysis:



myGTAP module”. This module can depict endogenous saving rates for households and other mechanism often found in single country CGE models such as government debt dynamics.

An additional equation not defined in the GTAP standard model defines a private consumption index *ycf*:

```
--- private consumption quantity index, not part of GTAP standard
e ycf(rsNat,ts) $ sum( i, xaiFlag(rsNat,i,"hhsld")) ..
ycf(rsNat,ts)*pcons(rsNat,"hhsld",ts)s(yc,'rsNat,"hhsld",ts') =e= ycf(rsNat,"hhsld",ts)s(yc,'rsNat,"hhsld",ts') ;
```

## 2.9 Household consumption

The **consumer price indices** *pcons* for each household *h* are defined from the budget share *xcshr* and the Armington prices defined in the macro *m\_pa*, equation (38) in VDM 2018. The global *%rsh%* is equal to the active model regions *rsNat* for the standard case. A special extension of myGTAP combined with sub-national units can introduce households at the sub-national level. In this case, *%rsh%* comprises both nations and sub-national units.

```
---- Consumer expenditure deflator (approx.) -- PHHLDINDEX (1005)
e pcons(rsh%,h,ts) $ (sum( i, xaiFlag(rsh%,i,h)) and ( (pcons.range(rsh%,h,ts) ne 0) or (not rres(rsh%)) or (lcu.range(rsh%,ts) ne 0) ) )
pcons(rsh%,h,ts) =e= sum( i $ xaiFlag(rsh%,i,h), xcshr(rsh%,i,h,ts)*m_pa(rsh%,i,h,ts)/pa.l(rsh%,i,h,"%t0%"));
```

The Armington demands for household consumption can be defined either by a CDE demand system as used in the GTAP Standard model, a MAIDADS or a LES demand system as found in many other CGE models. Note that the MAIDADS system can be turned into LES system if the marginal budget shares and the commitment terms are not rendered a function of utility, but fixed. Equally, the LES system collapses to a CD system if the commitments are removed, such the model can host four different demand systems for household consumption. Related benchmarking code is provided and applied according to the user’s choice.

### 2.9.1 CDE case

In the standard GTAP model, a constant difference in elasticity (CDE) indirect demand system is used. The demand system comprises three vectors with product specific parameters such that it can be

benchmarked to given product quantities, prices and income and still target both income and own price elasticities (see Ho et al. 2020<sup>13</sup> for a comparison of different demand systems used in CGE models.).

The equations can be found “*model\dem\_cde.gms*”. The **final demand quantity**  $xa$  (the Armington demand) for each household  $\%rsh\%$  and product  $i$  are defined from the budget shares  $xcshr$  and the private consumption expenditures  $yc$ , see equation  $e\_xac$ , equation (37) in VDM 2018:

```
*
* --- Household demands for goods and services, equation (37) in VDM 2018
*
e_xac(%rsh%,iIn(i),h,ts) $ xaiFlag(%rsh%,i,h) ..
  xa(%rsh%,i,h,ts)s(xa,%rsh%,i,h,ts)*m_pa(%rsh%,i,h,ts)
  =e= xcshr(%rsh%,i,h,ts)s(xa,%rsh%,i,h,ts)*yc(%rsh%,h,ts);
```

The **budget shares**  $xcshr$  for single product as far as they are not part of a CES sub-demand nest are defined in the equation  $e\_xcshr$ , derived from unscaled budget shares  $zcons$ , scaled again by unity based on their sum  $zConsSum$ , equation (36) in VDM 2018. This equation is active for product demands driven directly by the CDE function, excluding cases where they are part of CES demand nests.

```
*
* --- Budget shares, equation (36) in VDM 2018
*
e_xcshr(%rsh%,iIn(i),h,ts) $ (xaiFlag(%rsh%,i,h) $ (not sum(dNest_i_fd(dNest,i,h),1))) ..
  xcshr(%rsh%,i,h,ts)s(xcshr,%rsh%,i,h,ts) =e= (zcons(%rsh%,i,h,ts)/zConsSum(%rsh%,h,ts))s(xcshr,%rsh%,i,h,ts);
```

The unscaled shares  $zcons$  as defined in the equation  $e\_zcons$  depend on utility  $uh$  (i.e. indirectly on expenditures) and the product prices defined in the macro  $m\_pa$  relative to income  $yc$  per capita, defined from population  $pop$  for that household and three given parameter vectors  $alphaa$ ,  $bh$  and  $eh$ . That is an unnumbered equation in VDM 2018, noted before equation (36):

```
*
* --- auxiliary variables in CDE demand system, not numbered equation in VDM 2018
* (above equation (36) in VDM 2018)
*
e_zcons(%rsh%,i,h,ts) $ (xaiFlag(%rsh%,i,h) $ p_alphaa(%rsh%,i,h,"top",ts)) ..
  zcons(%rsh%,i,h,ts)s(zCons,%rsh%,i,h,ts) =e= p_alphaa(%rsh%,i,h,"top",ts)*bh(%rsh%,i,ts)

*
* --- bh measures substitution effect, driven by price change
* relative to income change
*
* m_pa(%rsh%,i,h,ts)**bh(%rsh%,i,ts)
*
* ((yc(%rsh%,h,ts)/pop(%rsh%,h,ts)*1000)**(-bh(%rsh%,i,ts)))
*
* --- expansion effect, driven by utility (= income)
*
* (uh(%rsh%,h,ts)**(eh(%rsh%,i,ts)*bh(%rsh%,i,ts)))s(zCons,%rsh%,i,h,ts);
```

The sum of these unscaled shares must also consider the case of sub-nests in demand, defined in the equation  $e\_zConsSum$ , part of equation (36) in VDM 2018:

<sup>13</sup> Ho, M., Britz, W., Delzeit, R., Leblanc, F., Roson, R., Schuenemann, F., & Weitzel, M. (2020). Modelling consumption and constructing long-term baselines in final demand. *Journal of Global Economic Analysis*, 5(1), 63-108

```

*
* ---- definition of sum of Z, part of equation (36) in VDM 2018
*
e_zConsSum(%rsh%,h,ts) $ sum(j,xaiFlag(%rsh%,j,h)) ..

zConsSum(%rsh%,h,ts)s(zConsSum,%rsh%,h,ts')
=e= [ sum(j $ (xaiFlag(%rsh%,j,h) $ (not sum(dNest_i_fd(dNest,j,h),1))), zcons(%rsh%,j,h,ts))
+ sum(dNest_n_fd("top",dNest,h) $ p_alphaDN(%rsh%,dNest,h,"top",ts), zconsDN(%rsh%,dNest,h,ts))
]s(zConsSum,%rsh%,h,ts');

```

The **utility level**  $u$  is indirectly defined by the following equation in the equation  $e\_uh$ , equation (32) in VDM 2018:

```

*
* --- Household utility (per capita), equation (32) in 2018
*
e_uh(%rsh%,h,ts) $ sum(iIn(i),xaiFlag(%rsh%,i,h)) ..

1 =e= sum(i $ (xaiFlag(%rsh%,i,h) $ p_alphaa(%rsh%,i,h,"top",ts)), zcons(%rsh%,i,h,ts)/bh(%rsh%,i,ts))
*
* --- add sub-nests under final demand function, not part of standard GTAP model
*
+ sum(dNest_n_fd("top",dNest,h) $ p_alphaDN(%rsh%,dNest,h,"top",ts), zconsDN(%rsh%,dNest,h,ts)/bhDN(%rsh%,dNest,ts));

```

Finally, the **elasticity of private expenditure versus private utility**  $phip$  is defined in the equation  $e\_phiP$ . It updates the private consumption share  $betap$  in the income distribution based on the regional household approach:

```

*
* --- Elast. of priv. exp. wrt to priv. utility, equation (27) in VDM 2018
*
e_phiP(rsNat,ts) $ ((phip.range(rsNat,ts) ne 0)) ..

phiP(rsNat,ts) =e= sum( (r_r(%rsh%,rsNat),i,h) $ xaiFlag(%rsh%,i,h), xcshr(%rsh%,i,h,ts)*eh(%rsh%,i,ts)*pop(%rsh%,h,ts))
/sum((r_r(%rsh%,rsNat),h) $ hFlag(%rsh%,h),pop(%rsh%,h,ts));

```

The demand for CES sub-nests directly under the CDE demand function ( $dnest\_n\_fd("top",dNest,h)$ ) follows the same structure as for single commodity, i.e. the budget shares  $xcShrDN$  of such nests are defined from the related  $zConsDN$  variables:

```

*
* --- same as above for CES-subnests, not part of standard GTAP model
*
e_xcshrDN(%rsh%,dNest,h,ts) $ sum(dNest_n_fd("top",dNest,h), p_alphaDN(%rsh%,dNest,h,"top",ts)) ..

xcshrDN(%rsh%,dNest,h,ts)s(xcshrDN,%rsh%,dNest,h,ts') =e= (zconsDN(%rsh%,dNest,h,ts)/zConsSum(%rsh%,h,ts))s(xcshrDN,%rsh%,dNest,h,ts');

```

And the related demand quantity  $xdNest$  is derived from total consumption expenditure  $yc$  and the budget shares  $xcShrDN$ :

```

e_xacDN(%rsh%,dNest,h,ts) $ sum(dNest_n_fd("top",dNest,h), p_alphaDN(%rsh%,dNest,h,"top",ts)) ..

xdNest(%rsh%,dNest,h,ts)*pdNest(%rsh%,dNest,h,ts)s(xdNest,%rsh%,dNest,h,ts')
=e= xcshrDN(%rsh%,dNest,h,ts)*yc(%rsh%,h,ts)s(xdNest,%rsh%,dNest,h,ts');

```

The equivalent variation  $ev$  is defined as follows:

```

*
* --- Implicite definition of ev
*
e_ev(rsNat,h,ts) $ (hFlag(rsNat,h) $ (ev.range(rsNat,h,ts) ne 0) $ (card(iIn) gt 1)) ..

sum(i$ (xaiFlag(rsNat,i,h) $ p_alphaa(rsNat,i,h,"top",ts)),
p_alphaa(rsNat,i,h,"top",ts) * (uh(rsNat,h,ts) ** (bh(rsNat,i,ts) * eh(rsNat,i,ts)))
* (pa.1(rsNat,i,h,"%t0%") * (Pop(rsNat,h,ts)/1000) / (ev(rsNat,h,ts))) ** bh(rsNat,i,ts))

+ sum(dNest_n_fd("top",dNest,h) $ p_alphaDN(rsNat,dNest,h,"top",ts),
p_alphaDN(rsNat,dNest,h,"top",ts) * (uh(rsNat,h,ts) ** (bhDn(rsNat,dNest,ts) * ehDn(rsNat,dNest,ts)))
* (pdNest.1(rsNat,dNest,h,"%t0%") * (Pop(rsNat,h,ts)/1000) / ev(rsNat,h,ts)) ** bhDn(rsNat,dNest,ts))

=e= 1 ;

```

## 2.9.2 LES or CD case

The equations for the LES or CD case are found in the file “*model\dem\_les.gms*”. The LES case is not part of GTAP Standard model.

The **Armington demands**  $xa$  defined by the equation  $e\_xacLes$  in the LES case reflect the constant term  $gammaLES$ , often termed commitment, and a share  $alphaLES$  on non-committed income  $ycNonCom$  divided by the Armington price defined in the macro  $m\_pa$ . Note that the commitment terms are defined on per capita basis (cnst USD dollar per capita) to avoid that they need to updated in case of population changes. To define the demand at household level, they are therefore multiplied with the population of the household and divided by a scaling factor of 1,000:

```
e_xacLes(%rsh%,iIn(i),h,ts) $ p_hasXa(%rsh%,i,h) ..
    xa(%rsh%,i,h,ts)s(xa,'%rsh%,i,h,ts')
    =e= (
    $$iftheni.baseline "%fixCalories%"=="on"
        gammaLES(%rsh%,i,h,ts)
    $$else.baseline
        gammaLESM(%rsh%,i,h,ts)
    $$endif.baseline
    /1000*pop(%rsh%,h,ts) + alphaLES(%rsh%,i,h,ts)/m_pa(%rsh%,i,h,ts)
    * ycNonCom(%rsh%,h,ts)s(xa,'%rsh%,i,h,ts');
```

The reader should note that the system collapses to a CD demand system if the constant terms are zero. The  $gammaLes$  is an additional variable used in the context of G-RDEM if calories are driven by given per capita-demand and the parameters of the MAIDADS demand system are endogenously updated.

This equation replaces equation (37) in VDM 2018, CDE case:

$$XA_{r,i,h} = \gamma_{r,i,h}^{LES} pop_{r,h} + \alpha_{r,i,h}^{LES} ycNonCom_{r,h} \quad (37^*)$$

The same functional relation is also used to define sub-nests demands directly under the LES type demand system `sum(dNest_n_fd("top",dNest,h), p_hasXA(%rsh%,dNest,h) )` as defined in the equation  $e\_xdNestLes$ :

```
--- demand for sub-nests under top nest
e_xdNestLes(%rsh%,dNest,h,ts) $ sum(dNest_n_fd("top",dNest,h), p_hasXA(%rsh%,dNest,h) ) ..
    xdNest(%rsh%,dNest,h,ts)s(xdNest,'%rsh%,dNest,h,ts')
    =e= (
    $$iftheni.baseline "%fixCalories%"=="on"
        gammaLES(%rsh%,dNest,h,ts)
    $$else.baseline
        gammaLESM(%rsh%,dNest,h,ts)
    $$endif.baseline
    /1000*pop(%rsh%,h,ts) + alphaLES(%rsh%,dNest,h,ts)
    /pdNest(%rsh%,dNest,h,ts)
    * ycNonCom(%rsh%,h,ts)s(xdNest,'%rsh%,dNest,h,ts');
```

**Non-committed income**  $ycNonCom$  as defined in the equation  $e\_ycNonCom$  reflect total private consumption expenditure  $yc$  minus the value of the commitments, i.e. the gamma parameters multiplied with the Armington prices defined in the macro  $m\_pa$ , respectively the commitment terms for sub-nest with their prices  $pdNest$ , considering the conversion from per capita to total household level:

```

e_yCNonCom(%rsh%,h,ts) $ sum(i,xaiFlag(%rsh%,i,h)) ..

ycNonCom(%rsh%,h,ts) s(yc,%rsh%,h,ts)
=E= (yc(%rsh%,h,ts)

    --- note: commitment terms are defined in a per capita basis

$iftheni.baseline "%fixCalories%"=="on"
- sum(i $ p_hasXa(%rsh%,i,h), gammaLES(%rsh%,i,h,ts) *pop(%rsh%,h,ts)/1000 * m_pa(%rsh%,i,h,ts))
- sum(dNest_n_fd("top",dNest,h) $ p_hasXa(%rsh%,dNest,h), gammaLES(%rsh%,dNest,h,ts) *pop(%rsh%,h,ts)/1000 * pdNest(%rsh%,dNest,h,ts)) s(yc,%rsh%,h,ts);
$else.baseline
- sum(i $ p_hasXa(%rsh%,i,h), gammaLESM(%rsh%,i,h,ts) *pop(%rsh%,h,ts)/1000 * m_pa(%rsh%,i,h,ts))
- sum(dNest_n_fd("top",dNest,h) $ p_hasXa(%rsh%,dNest,h), gammaLESM(%rsh%,dNest,h,ts) *pop(%rsh%,h,ts)/1000 * pdNest(%rsh%,dNest,h,ts)) s(yc,%rsh%,h,ts);
$endif.baseline

```

In mathematical notation:

$$ycNonCom_{r,h} = \left( yc_{r,h} - \sum_i \gamma_{r,i,h}^{LES} pop_{r,h} PA_{r,i,h} - \sum_{dNest \in top} \gamma_{r,dNest,h}^{LES} pop_{r,h} PDNest_{r,i,h} \right)$$

In case of the CD demand system, the gamma parameters are all zero such that  $ycNonCom$  is equal to total consumption expenditure  $yc$ .

The **budget shares**  $xcshr$  are defined from the Armington demands, prices and expenditures in the equation  $e\_xcshrLESeq$ :

```

*
* --- definition of budget shares
*
p_xcshrLES(%rsh%,iIn(i),h,ts(t)) $ xaiFlag(%rsh%,i,h) ..

xcshr(%rsh%,i,h,ts) *yc(%rsh%,h,ts) s(xa,%rsh%,i,h,ts) *1s(pa,%rsh%,i,h,ts)
=E= xa(%rsh%,i,h,ts) *m_pa(%rsh%,i,h,ts) s(xa,%rsh%,i,h,ts) *1s(pa,%rsh%,i,h,ts)

```

This equation replaces equation (36) in VDM 2018 for the CDE case:

$$s_{r,i,h}^p yc_{r,h} = XA_{r,i,h} PA_{r,i,h} \quad (36^*)$$

Finally, the utility for the private households  $uh$  is defined in the equation  $e\_uhLES$ :

```

*
* --- definition of utility from private household, LES
*
e_uhLES(%rsh%,h,ts) $ (sum(i,xaiFlag(%rsh%,i,h)) $ (p_demSystem eq a_LES)) ..

uh(%rsh%,h,ts) s(uh,%rsh%,h,ts)
=E= ycNonCom(%rsh%,h,ts)/pop(%rsh%,h,ts) * 1000 s(uh,%rsh%,h,ts)
* prod(i $ (xaiFlag(%rsh%,i,h) $ alphaLES.1(%rsh%,i,h,ts)),
(alphaLES(%rsh%,i,h,ts)/m_pa(%rsh%,i,h,ts)) ** alphaLES(%rsh%,i,h,ts))
* prod(dNest_n_fd("top",dNest,h) $ alphaLES.1(%rsh%,dNest,h,ts),
(alphaLES(%rsh%,dNest,h,ts)/pdNest(%rsh%,dNest,h,ts)) ** alphaLES(%rsh%,dNest,h,ts))

* 1/uhLES(%rsh%,h,ts);

```

Note that the CD case is comprised in the equations above if the commitment terms are set to zero. The equation above replaces equation (32) from VDM 2018 for the CDE case:

$$U_r^h = \frac{1}{UBas_r^h} \prod_i \left( XA_{r,i,h} - \gamma_{r,i,h}^{LES} pop_{r,h} \right)^{\alpha_{r,i,h}^{LES}} \prod_{tNest \in top} \left( XdNest_{r,tNest,h} - \gamma_{r,tNest,h}^{LES} pop_{r,h} \right)^{\alpha_{r,tNest,h}^{LES}} \quad (32^*)$$

The elasticity of private expenditure with regard to private utility is defined in the following equations:



```

e_phiPLES(rsNat,ts) $ (p_hip.range(rsNat,ts) ne 0) ..

phiP(rsNat,ts) =e=

sum( r_r(rs,rsNat) $ sum(h $ hFlag(rs,h),1),
*
* --- population weighted 1/[ du/dm] from indirect utility function
*

sum(h $ hFlag(rs,h), pop(rs,h,ts)
* prod(i $ (p_hasXa(rs,i,h) $ alphaLES.l(rs,i,h,"%t0%")),
(m_pa(rs,i,h,ts)
/(alphaLES(rs,i,h,ts))**alphaLES(rs,i,h,ts))
* prod(dNest_n_fd("top",dNest,h) $ (p_hasXa(rs,dNest,h) $ alphaLES.l(rs,dNest,h,"%t0%")),
(pdNest(rs,dNest,h,ts)/alphaLES(rs,dNest,h,ts))**alphaLES(rs,dNest,h,ts)))
/sum(h $ hFlag(rs,h), pop.l(rs,h,"%t0%"))
* prod(i $ (p_hasXa(rs,i,h) $ alphaLES.l(rs,i,h,"%t0%")),
(pa.l(rs,i,h,"%t0%")/alphaLES.l(rs,i,h,"%t0%"))**alphaLES.l(rs,i,h,"%t0%"))
* prod(dNest_n_fd("top",dNest,h) $ (p_hasXa(rs,dNest,h) $ alphaLES.l(rs,dNest,h,"%t0%")),
(pdNest.l(rs,dNest,h,"%t0%")/alphaLES.l(rs,dNest,h,"%t0%"))**alphaLES.l(rs,dNest,h,"%t0%"))
*
* --- conversion to elasticity
*
* sum(h $ hFlag(rs,h), pop(rs,h,ts) * uh(rs,h,ts))/sum(h $ hFlag(rs,h), pop(rs,h,ts) * yc(rs,h,ts))
*
* --- and correct for benchmark relation and weight with population
*
* sum(h $ hFlag(rs,h), pop.l(rs,h,"%t0%") * yc.l(rs,h,"%t0%"))
/sum(h $ hFlag(rs,h), pop.l(rs,h,"%t0%") * uh.l(rs,h,"%t0%"))
)
/sum((r_r(rs,rsNat)) $ sum(h,hFlag(rs,h),1);

```

The econometrically estimated AIDADS system as part of G-RDEM modules is detailed in section “An AIDADS demand system with detail for food consumption”, pages 216 ff. Whereas the marginal budget share parameter *alphaLes* and the per capita commitment terms *gammaLes* are fixed for the LES case, they depend on the utility level in the MAIADS system.

## 2.10 Government consumption

The reader is first reminded that there is no separate government account under the regional household approach and hence no direct link between tax income and government spent. That can be changed by using the “Capital vintages

### 2.10.1 Summary

Capital vintages dis-aggregate a production activity into two variants: one working with new capital generated in the same period from investments and an old one working with capital generated in earlier period. The equations in the model relating to the capital vintages are described in the section “Production block”.

### 2.10.2 Intialization

The intialization of the capital vintages is performed in ‘cal/cal\_vint.gms’. Based on the user selection on the GUI (see section below), the set *aVint(a)* comprises the data base activities dis-aggregated into an old and new vintage, showns on the sets *aOld* and *aNew* and their union set *aVints*. Activities outputting products in the SPE and Melitz model extensions are removed from the set *aVint* as defined selection by the user, warnings are added to the listing as shown below for the case of the Melitz extension:

```

$ifthen,mel,defined iMel
set remVintA(a);
remVintA(aVint) $ sum((rnat,aVint(aVint,aVints),iMel),sam0(rnat,aVint,iMel)) = YES;

if (card(remVintA),
$setinclude 'util/final_log' "Warning in %system.fn%, line: %system.incline%: using vintage activities for Melitz products not possible. These activities are set back to standard "
display "Warning in %system.fn%, line: %system.incline%: sing vintage activities for Melitz products not possible. . Such activities are removed ".remVintA;
vintageA_sel(remVintA) = no;
);
aVint(a) $ sum((rnat,iMel),sam0(rnat,a,iMel)) = no;

```

The remaining data base activities *aVint* are dropped from the SAM and from other data referring to production activities based on the set *aDrop*:

```

*
* --- list of activities to drop from SAM and other data parameters
*
aDrop(aVint) = yes;
*

```

While all other activities (not a new or old capital vintages or an aggregate thereof) are kept according to the set *aKeep*:

```

*
* --- list of activities to keep in SAM and other data parameters
*
aKeep(a)      = yes;
aKeep(aDrop)  = no;
aKeep(aNew)   = no;
aKeep(aOld)   = no;
*

```

The distribution of the data to the old and new variants at the benchmark is based on the default depreciation rate of 4% per annum times the number of depreciation years. This number is equal to unity for recursive-dynamic analysis and chosen by the user for comparative-static one. The distribution shares are set accordingly:

```

--- distribution shares for old and new vintage

p_distrVint(rs,aNew)          = p_depr(rs,"%t0%")+0.04*(p_nDeprYears("%t0%")-1);

p_distrVint(rs,aOld)          = 1-sum(a_avint(a,aNew) $ a_aVint(a,aOld),p_distrVint(rs,aNew));

```

These shares distribute the given SAM (and other data) to the old and new vintages while the original activities are removed, as shown for the case of the SAM:

```

--- distributes outputs to old and new vintages with distribution keys

SAM0(rs,aVints,is) = sum(a_aVint(a,aVints),SAM0(rs,a,is)*p_distrVint(rs,aVints));
SAM0(rs,aDrop,is)  = 0;

```

Similar statements distribute the data referring to the domestic and imported demand, for factor taxes and subsidies and for production taxes. New and old capital are treated differently: old vintages receive all capital minus the new capital allocated to the new vintage. Activities not split up into the old and new vintage receive both old and new capital and a CES nests that allows for substitution between the two.

The substitution elasticities for the new vintage are set 2.5 higher higher compared to the parameters in a model without the vintage extension for technology nests comprising old or new capital as a production factor. The substitution elasticities for the old vintages are downscaled accordingly to keep the average reaction of the model to changes in price relations:

```

:

--- downscale subst. elas of old vintage such that in average over the two vintages, the given subst.
    elasticity is maintained

p_sigmaNest(rs,tNest,aOld) $ sum(tNest_f_a(tNest,cap,aOld),1)
    = p_sigmaNest(rs,tNest,aOld) * (1 - (1-p_distrVint(rs,aOld))*p_sigmaMult(rs,aOld));

p_sigmaNest(rs,tNest,aNew) $ sum(tNest_f_a(tNest,cap,aNew),1)
    = p_sigmaNest(rs,tNest,aNew) * p_sigmaMult(rs,aNew);

```

```

parameter p_sigmaMult(r,a);
tNest("va") = YES;
tNest("nd") = YES;

--- multiplicative increase in subst. elasticity for nests comprising capital for new vintage is 2.5

p_sigmaMult(rs,aVints) $ xpFlag(rs,aVints) = 2.5;

--- downscale subst. elas of old vintage such that in average over the two vintages, the given subst.
    elasticity is maintained

p_sigmaNest(rs,tNest,aOld) $ sum(tNest_f_a(tNest, cap, aOld), 1)
    = p_sigmaNest(rs,tNest,aOld) * (1 - (1-p_distrVint(rs,aOld)) * p_sigmaMult(rs,aOld));
p_sigmaNest(rs,tNest,aNew) $ sum(tNest_f_a(tNest, cap, aNew), 1)
    = p_sigmaNest(rs,tNest,aNew) * p_sigmaMult(rs,aNew);

```

### 2.10.3 Dynamic updating of share parameters

In recursive-dynamic mode the code considers that a part of the old vintage is removed in each period, due to depreciation while last year's new vintage becomes part of the old vintage. For instance, expected output (stored temporarily on the field .m) is defined as:

```

xp.m(rs,aOld,tSim) $ xp.1(rs,aOld,tSim-1)
    = sum(aVint_aVint(aNew,aOld), xp.1(rs,aNew,tSim-1) * p_nDeprYears(tSim-1))
      + xp.1(rs,aOld,tSim-1)
        * sum(r_r(rs,rsNat), (1 - p_depr(rsNat,tSim-1) * p_nDeprYears(tSim-1)));

```

Similar statements define the expected demand for factors, intermediates and CES nests in the production function. In parallel, the code defines expected prices for the old vintage in the next period as weighted average between last period's old and new vintages, considering as above depreciation. For instance, the unit cost average for the new vintages is defined as:

```

px.m(rs,aOld,tSim) $ xp.m(rs,aOld,tSim)
    = [sum(aVint_aVint(aNew,aOld), xp.1(rs,aNew,tSim-1) * px.1(rs,aNew,tSim-1) * p_nDeprYears(tSim-1))
      + xp.1(rs,aOld,tSim-1) * px.1(rs,aOld,tSim-1)
        * sum(r_r(rs,rsNat), (1 - p_depr(rsNat,tSim-1) * p_nDeprYears(tSim-1)))] / xp.m(rs,aOld,tSim);

```

Based on these update demands for production factors and their prices, the share parameters are updated for the old vintage are updated, for instance for input demands:

```

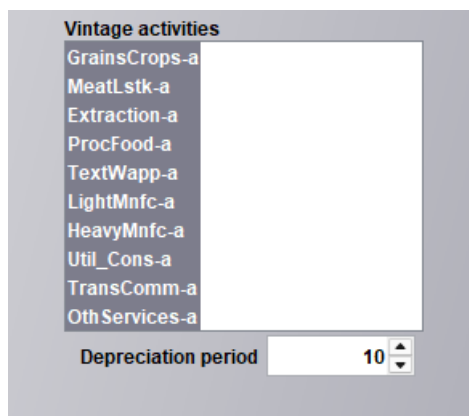
*
* --- redefine share parameters, such that the old vintage would demand at prices and shifters of tSim-1 the
* aggregate quantities of the new and old vintages
*
p_ioNest(rs,i,aOld,iotNest,tSim) $ (xa.m(rs,i,aOld,tSim) $ sum(tNest_i_a(iotNest,i,aOld), p_ioNest(rs,i,aOld,iotNest,tSim) $ (xtNest.m(rs,iotNest,aOld,tSim) > 0)))
    = sum(tNest_i_a(iotNest,i,aOld), (xa.m(rs,i,aOld,tSim) * tNestShare(rs,iotNest,i,aOld) / xtNest.m(rs,iotNest,aOld,tSim))
      * 1/lambdaio(rs,i,aOld,tSim-1) * (([p_sigmaNest(rs,iotNest,aOld)-1] $ (not addCes(rs,iotNest,aOld))
        - p_sigmaNest(rs,iotNest,aOld) $ addCes(rs,iotNest,aOld) ))
      * (pa.m(rs,i,aOld,tSim) / ptNest.m(rs,iotNest,aOld,tSim)) * p_sigmaNest(rs,iotNest,aOld));

```

This implies that the technology of the old vintage changes over time. Fast growing sectors with an above-average share of the new vintages will show larger changes in their average technologies.

### 2.10.4 GUI

The GUI allows the user which production activities are broken down two vintages and to set the depreciation period in case of a comparative-static analysis:



myGTAP module”.

Government consumption expenditure  $y_g$  under the regional household approach is a share  $\beta_{tag}$  of regional income  $regy$ , corrected for the endogenous share of private spent captured by  $\phi_{iRegy}$  (see above for savings), equation (39) in VDM 2018:

```

--- Total government consumption expenditure, equation (30) in VDM 2018
s_yg(rsNat,ts) ..
[ xg(rsNat,ts) $ (p_sigmaU(rsNat) ne 1) + yg(rsNat,ts) $ (p_sigmaU(rsNat) eq 1)]s(xg,'rsNat,ts')
=e= [betaG(rsNat,ts)*phiRegy(rsNat,ts) * regy(rsNat,ts)s(xg,'rsNat,ts')] $ (p_sigmaU(rsNat) eq 1)
+ [betaG(rsNat,ts)*phiRegy(rsNat,ts) * (regy(rsNat,ts)/pu(rsNat,ts))s(xg,'rsNat,ts')
* pg(rsNat,ts)**(-p_sigmaU(rsNat)))] $ (p_sigmaU(rsNat) ne 1) ;

```

The physical demand aggregate  $x_g$  defined in equation  $e_{xg}$  is derived from the government price index  $pg$  and the expenditures  $y_g$ . It allows to consider a case with targeted government consumption expenditures at sub-region level:

```

--- Nominal government expenditures, equation (41) in VDM 2018
s_xg(rsNat,ts) $ p_axg(rsNat,ts) ..
pg(rsNat,ts)*xg(rsNat,ts)s(xg,'rsNat,ts') =e= yg(rsNat,ts)s(xg,'rsNat,ts') - sum((r_r(subr,rsNat),i,gov),p_subrDem(subr,i,gov,ts))s(yi,'rsNat,ts');

```

The equation  $e_{pg}$ , equation (40) in VDM 2018, is expanded in CGEBox to account for demand nests. The related price index  $pg$  is defined under the assumption of a CES / CD / Leontief demand system for the final demand of government using the typical dual price aggregator based on the government specific Armington prices defined by the marco  $m_{pa}$ , a preference shifter variable  $\lambda_{daa}$  and the share parameters  $\alpha_{haa}$  for government. Which case is used is defined by the related substitution elasticity  $\sigma_{mag}$ . Note that the equation also considers the case that the top-level government demand function comprises CES sub-nests that aggregate products or other product groups to product groups:

--- Government expenditure price deflator, equation (40) in VDM 2018

```

e_xg(rsNat,ts) $ p_axg(rsNat,ts) ..
pg(rsNat,ts) =E=
    --- CD case with p_sigmag==1
    ( [ prod((i,gov) $ p_alphaa(rsNat,i,gov,"top",ts),
        (m_pa(rsNat,i,gov,ts)/(p_alphaa(rsNat,i,gov,"top",ts))*lambdaa(rsNat,i,gov,ts)))**p_alphaa(rsNat,i,gov,"top",ts))
    -- CES-subnests, not part of Standard GTAP model
    * prod(dNest_n_fd("top",dNest,gov) $ p_alphadN(rsNat,dNest,gov,"top",ts),
        (pdNest(rsNat,dNest,gov,ts)/(p_alphadN(rsNat,dNest,gov,"top",ts))*lambdaadN(rsNat,dNest,gov,"top",ts)))**p_alphadN(rsNat,dNest,gov,"top",ts))
    ]*p_axg(rsNat,ts)**(-1) ) $ (p_sigmag(rsNat) eq 1)

    --- CES or Leontief case with p_sigmag<>1
    +
    ([ sum((i,gov) $ p_alphaa(rsNat,i,gov,"top",ts),
        p_alphaa(rsNat,i,gov,"top",ts)
        * (m_pa(rsNat,i,gov,ts)/lambdaa(rsNat,i,gov,ts)))** (1-p_sigmag(rsNat)))
    -- CES-subnests, not part of Standard GTAP model
    + sum(dNest_n_fd("top",dNest,gov) $ p_alphadN(rsNat,dNest,gov,"top",ts),
        p_alphadN(rsNat,dNest,gov,"top",ts)* (pdNest(rsNat,dNest,gov,ts)/lambdaadN(rsNat,dNest,gov,"top",ts)))** (1-p_sigmag(rsNat)))
    ]** (1/(1-p_sigmag(rsNat))) * p_axg(rsNat,ts)** (-1)
    ) $ (p_sigmag(rsNat) ne 1);

```

The macro  $m\_pa$  is usually defined as follows. If both a share parameter for imports  $\alpha_{pham}$  and for domestic sales  $\alpha_{phad}$  is given, it uses the definition of the Armington price  $m\_pade$  given below. Otherwise, i.e. if solely demand for the domestic or the imported origin is found, it uses directly either the macro for the domestic price  $m\_pdp$  or for the import prices  $m\_pmp$ . Finally, if neither of the two share parameters is given, the Armington price is used directly, a case relevant when the Melitz module is active. This module will delete the share parameters for products under its control.

```

$macro m_pa(r,i,aa,t) {
    m_pade(r,i,aa,t) $ (alpha(r,i,aa,t) $ alpha(r,i,aa,t) ) \
    + (m_pdp(r,i,aa,t)*alpha(r,i,aa,t)) $ ((not alpha(r,i,aa,t)) $ alpha(r,i,aa,t)) \
    + (m_pmp(r,i,aa,t)*alpha(r,i,aa,t)) $ ((not alpha(r,i,aa,t)) $ alpha(r,i,aa,t)) \
    $if1 "%modulesGTAP_MELITZ%"="on" + pa(r,i,aa,t) $ ( not (alpha(r,i,aa,t)+alpha(r,i,aa,t)) ) \
}

```

The macro  $m\_pade$  defines the dual price aggregator, equation (47) in VDM 2018:

```

$macro m_pade(r,i,aa,t) { (
    (alpha(r,i,aa,t)*m_pdp(r,i,aa,t)**(1-sigma(r,i))) \
    + (alpha(r,i,aa,t)*m_pmp(r,i,aa,t)**(1-sigma(r,i))) )** (1/(1-sigma(r,i))) }

```

The total physical government demand  $xg$  is distributed in the equation  $e\_xag$ , equation (39) in VDM 2018, to demands for individual products based on given share parameters  $\alpha_{phaa}$ , Armington prices captured by the  $m\_pa$  macro and the price index  $pg$ , considering the substitution elasticity  $\sigma_{mag}$ :

--- CES expenditure function for government, equation (39) in VDM 2018

```

e_xag(rsNat,iIn(i),gov,ts) $ (xaiFlag(rsNat,i,gov) $ p_alphaa(rsNat,i,gov,"top",ts)) ..
xa(rsNat,i,gov,ts)s(xa,'rsNat,i,gov,ts')
=e= p_alphaa(rsNat,i,gov,"top",ts)*xg(rsNat,ts)s(xa,'rsNat,i,gov,ts')
* [ ((lambdaa(rsNat,i,gov,ts)*pg(rsNat,ts)/m_pa(rsNat,i,gov,ts))**p_sigmag(rsNat)/lambdaa(rsNat,i,gov,ts) ) $ (p_sigmag(rsNat) ne 0)
+ lambdaa(rsNat,i,gov,ts) $ (p_sigmag(rsNat) eq 0) ]
+ sum(r_r(subr,rsNat),p_subrDem(subr,i,gov,ts)/m_pa(rsNat,i,gov,ts))s(xa,'rsNat,i,gov,ts');

```

Note that that in the case of demand nests for government consumption (not part of the GTAP standard model), additional equations are used as described in the section “Demand sub-nests”. The demand nest equation for the government is defined as:

$$\begin{aligned}
 XdNest_{r,dNest,gov} &= \alpha_{r,dNest,gov}^{top} XG_r \left( \frac{PG_r}{PdNest_{r,dNest,gov}} \right)^{\sigma^g} \\
 &+ \alpha_{r,dNest,gov}^{dNest1} XdNest_{r,gov} \left( \frac{PdNest_{r,dNest1,gov}}{PdNest_{r,dNest,gov}} \right)^{\sigma^g}
 \end{aligned}$$

## 2.11 Investments and savings

**Gross investment expenditures**  $y_i$  are composed of the value of depreciation  $valDep$  and of regional  $rsav$  and foreign savings  $valSavf$  as defined in the equation  $e_{y_i}$ , equation (85) in VDM 2018. Additional terms not found in the standard GTAP model can be introduced. So-called adaptive investments (to Climate Change) that don't contribute to the capital accumulation  $p_{govAdaptInv}$  and  $p_{prvAdaptInv}$  can be added, they require the *myGTAP* extension. Similarly, the costs  $valMaccs$  related to technologies reducing emission factors are deducted. The model can also introduce a simplified version of an enterprise account where a part of factor earnings defines expenditures  $invFromCapInc$  and the remainder defines factor income.

```

yi(rsNat,ts)s(yi,'rsNat,ts') =e= (
    ---- value of depreciation
    valDep(rsNat,ts)
    ---- savings of the regional household
    + rsav(rsNat,ts)
    + walras(ts) $ rres(rsNat)
    ---- direct investments of firm
    + invFromCapInc(rsNat,ts)
    ---- exogenous adaptive investments (require myGTAP)
    + p_govAdaptInv(rsNat,ts)
    + p_prvAdaptInv(rsNat,ts)

    $Sifi "%modelType%"=="EMP"

    $Sifi "%MACCsProcessEmis%"=="on" - valMaccs(rsNat,ts)
    ---- Value of foreign savings in local currency
    + valsavf(rsNat,ts)s(yi,'rsNat,ts') ;

```

The price index of investments  $p_i$  is defined by the equation  $e_{p_i}$ , equation (43) in VDM 2018, expanded for demand nests. For a detailed explanation of this equation, refer to the explanation for government above:

```

--- Investment expenditure price deflator, equation (43) in VDM 2018
e_pi(rsNat,ts) ..
pi(rsNat,ts) =e=

--- CD case with p_sigmai==1
( { prod(i,inv) $ p_alphaa(rsNat,i,inv,"top",ts),
  (m_pa(rsNat,i,inv,ts)/(lambdai(rsNat,i,inv,ts)*p_alphaa(rsNat,i,inv,"top",ts)))**p_alphaa(rsNat,i,inv,"top",ts))

  -- CES-subnests, not part of Standard GTAP model

  * prod(dNest_n_fd("top",dNest,inv) $ p_alphaDN(rsNat,dNest,inv,"top",ts),
    (pdNest(rsNat,dNest,inv,ts)/(p_alphaDN(rsNat,dNest,inv,"top",ts)*lambdadn(rsNat,dNest,inv,"top",ts)))**p_alphaDN(rsNat,dNest,inv,"top",ts))
  ]*p_axi(rsNat,ts)**(-1) ) $ (p_sigmai(rsNat) eq 1)

--- CES or Leontief case with p_sigmai>1
+
([ sum((i,inv) $ p_alphaa(rsNat,i,inv,"top",ts),
  p_alphaa(rsNat,i,inv,"top",ts)*(m_pa(rsNat,i,inv,ts)/lambdai(rsNat,i,inv,ts))**((1-p_sigmai(rsNat))))

  -- CES-subnests, not part of Standard GTAP model

  + sum(dNest_n_fd("top",dNest,inv) $ p_alphaDN(rsNat,dNest,inv,"top",ts),
    p_alphaDN(rsNat,dNest,inv,"top",ts)*(pdNest(rsNat,dNest,inv,ts)/lambdadn(rsNat,dNest,inv,"top",ts))**((1-p_sigmai(rsNat))))

  ]**((1/(1-p_sigmai(rsNat))))*p_axi(rsNat,ts)**(-1)
  ) $ (p_sigmai(rsNat) ne 1)
;

```

The **total physical investment demand**  $x_i$  is derived from the price index  $p_i$  defined above and total investment expenditure  $y_i$ , as defined in equation  $e_{x_i}$ , equation (44) in VDM 2018, considering the case of exogenously defined investment expenditure at regional level.

```

--- Investment volume, equation (44) in VDM 2018
e_xi(rsNat,ts) ..
yi(rsNat,ts)s(yi,'rsNat,ts') =E= pi(rsNat,ts)*xi(rsNat,ts)s(yi,'rsNat,ts')
    + sum((r_r(subr,rsNat),i,inv),p_subrDem(subr,i,inv,ts))s(yi,'rsNat,ts');

```

**Product specific investment demand**  $x_a$  as depicted in the equation  $e_{x_{ai}}$ , equation (42) in VDM 2018. It reflects the share parameters  $\alpha_{aaa}$ , the substitution elasticity  $\sigma_{mai}$  and the shifter variable  $\lambda_{bdai}$ , and adds demand from exogenous investment demand expenditures at sub-national level.

```

--- CES expenditure function, equation (42) in VDM 2018
e_xai(rsNat,iIn(i),inv,ts) $ (xaiFlag(rsNat,i,inv) $ p_alphaa(rsNat,i,inv,"top",ts)) ..
xa(rsNat,i,inv,ts)s(xa,'rsNat,i,inv,ts') =e=
p_alphaa(rsNat,i,inv,"top",ts)*xi(rsNat,ts)s(xa,'rsNat,i,inv,ts')
+ [ ((lambdaa(rsNat,i,inv,ts)*pi(rsNat,ts)/m_pa(rsNat,i,inv,ts))**p_sigmai(rsNat)/lambdaa(rsNat,i,inv,ts) ) $ (p_sigmai(rsNat) ne 0)
+ lambdaa(rsNat,i,inv,ts) $ (p_sigmai(rsNat) eq 0) ]
+ sum(r_r(subr,rsNat),p_subrDem(subr,i,inv,ts)/m_pa(rsNat,i,inv,ts))s(xa,'rsNat,i,inv,ts');

```

The **value of depreciation** *valDep* is a given share *p\_depr* of the capital stock *kstock* which jointly define physical depreciation, multiplied with the average price of savings *pi*, equation (85) in VDM 2018:

```

--- Nominal depreciation, part of equation (85) in VDM 2018
e_valDep(rsNat,ts) ..
valDep(rsNat,ts)s(valDep,'rsNat,ts')
=E= pi(rsNat,ts)*p_depr(rsNat,ts)*kstock(rsNat,ts)s(valDep,'rsNat,ts');

```

**Regional savings** *rsav* are a given share of regional income *betas*, corrected for expansion effects *phiRegy*, equation (31) in VDM 2018:

```

--- Total nominal saving, equation (31) in VDM 2018
e_rsav(rsNat,ts) ..
[ xsav(rsNat,ts) $ (p_sigmaU(rsNat) ne 1) + rsav(rsNat,ts) $ (p_sigmaU(rsNat) eq 1) ]s(xsav,'rsNat,ts')
=E= [ betaS(rsNat,ts) * phiRegy(rsNat,ts) * regy(rsNat,ts)s(xsav,'rsNat,ts') ] $ (p_sigmaU(rsNat) eq 1)
+ [ betaS(rsNat,ts) * phiRegY(rsNat,ts) * (regy(rsNat,ts)/pu(rsNat,ts))s(xsav,'rsNat,ts')
* pSave(rsNat,ts)**(-p_sigmaU(rsNat))] $ (p_sigmaU(rsNat) ne 1) ;

```

The **physical amount of regional savings** *xsav* is based on the savings expenditures *rsav* and the average price of savings *psave*, equation (34) in VDM 2018:

```

--- Physical amount of savings, equation (34) in VDM 2018
e_xsav(rsNat,ts) ..
xsav(rsNat,ts)*psave(rsNat,ts)s(xsav,'rsNat,ts') =e= rsav(rsNat,ts)s(xsav,'rsNat,ts');

```

The **regional value of foreign savings** *valSavf* is defined from the value in foreign currency *savf* and the exchange rate *lcu*, part of equation (85) in VDM 2018:

```

--- Value of foreign savings in local currency, part of equation (85) in VDM 2018
e_valSavf(rsNat,ts) ..
valSavf(rsNat,ts)s(valSavf,'rsNat,ts')
=E= savf(rsNat,ts)*lcu(rsNat,ts)s(valSavf,'rsNat,ts');

```

The **foreign savings** in foreign currency *savf* can be driven by different mechanism. We start by discussing the so-called global bank mechanism which uses expected returns investments in each model region to distribute global savings. The scalar *p\_rorFlag* determines which of the mechanism is active. The different options are defined as acronyms: *a\_equalReturnToInv* depicts the so-called global bank approach discussed next, *a\_fixedAlloactionOfInv* keeps the share of investments expenditures of each model region at the global sum constant and *a\_fixedForeignSavings* treats them as exogenous.



### 2.11.1 Global bank

The global bank mechanism distributes total global savings such that expected returns to net investments are equal across regions. The different steps in the allocation procedure are described by the following variables and equations.

**Regional physical net investments**  $netInv$  are the difference between gross investment demand  $xi$  and physical depreciation derived from the capital stock  $kStock$  and the depreciation rate  $p\_depr$ , part of equation (86) in VDM 2018. Adaptative investments are subtracted along with exogenous investments  $p\_subrDem$  at sub-regional level.

```

--- Physical net investment in each region, part of equation (86) in VDM 2018
netInv(rsNat,ts) ..
netInv(rsNat,ts)s(netInv,'rsNat,ts') =E= [ xi(rsNat,ts)+ sum((r_r(subr,rsNat),i,inv) $ p_subrDem(subr,i,inv,ts),p_subrDem(subr,i,inv,ts)/m_pa(rsNat,i,inv,ts))
- p_govAdaptInv(rsNat,ts)
- p_prvAdaptInv(rsNat,ts)
- p_depr(rsNat,ts)*kstock(rsNat,ts)]s(netInv,'rsNat,ts');

```

The **beginning of period capital stock**  $kStock$  is defined in the equation  $e\_kStock$ , equation (75) in VDM 2018. It converts with the factor  $p\_krat$  the use of capital  $xft$  into the aggregate capital stock, where different capital types are defined by the set  $cap$ :

```

--- Aggregate capital stock, equation (75) in VDM 2018
e_kstock(rsNat,ts) $ ( (kstock.range(rsNat,ts) ne 0) ) ..
kstock(rsNat,ts)s(kapend,'rsNat,ts') =e= sum(cap, xft(rsNat,cap,ts)/p_krat(rsNat,ts))s(kapend,'rsNat,ts');

```

The **end of period capital stock**  $kapEnd$  is derived by deducting depreciation, i.e. the depreciating rate  $p\_depr$  times the number of depreciation year  $p\_nDeprYears$ , and adding physical gross investment  $xi$  plus exogenous sub-regional investment  $p\_subrDem$  times the number of depreciations years, equation (76) in VDM 2018:

```

e_kapEnd(rsNat,ts) ..
kapEnd(rsNat,ts)s(kapend,'rsNat,ts')
=e= [
(1 - p_depr(rsNat,ts)*p_nDeprYears(ts))*kstock(rsNat,ts)
+ (xi(rsNat,ts)+ sum((r_r(subr,rsNat),i,inv) $ p_subrDem(subr,i,inv,ts),p_subrDem(subr,i,inv,ts)/m_pa(rsNat,i,inv,ts)))*p_nDeprYears(ts)
]s(kapend,'rsNat,ts');

```

The **average returns of returns to capital after taxes**  $arent$  is defined in the equation  $e\_arent$ , equation (77) in VDM 2018, from capital prices  $pft$ , considering direct taxes  $kappaf$  and their potential tax shifter  $kappaShft$ :

```

--- Aggregate capital rate of return after tax, equation (77) in VDM 2018
(Note: here, we use always the average across sectors as tax rates are uniform)
e_arent(rsNat,ts) ..
arent(rsNat,ts)
$scale% /arent.1(rsNat,"%t0%")
=e= [sum((cap) $ xftFlag(rsNat,cap),
(1-kappaf(rsNat,cap,ts)-kappaShft(rsNat,ts))*pft(rsNat,cap,ts)*xft(rsNat,cap,ts))/(kstock(rsNat,ts))]
$scale% /arent.1(rsNat,"%t0%")
;

```

The **net rate of return to capital**  $rorc$  as defined in the equation  $e\_rorc$ , equation (77) in VDM 2018, corrects these gross returns factors  $arent$  for the depreciation rate  $p\_fdepr$ :



```
--- Rate of return to capital, equation (78) in VDM 2018
```

```
e_rorc(rsNat,ts) ..  
  
    rorc(rsNat,ts)  
    %scale% /rorc.1(rsNat,"%t0%")  
    =e= (arent(rsNat,ts)/pi(rsNat,ts) - p_fdepr(rsNat,ts))  
    %scale% /rorc.1(rsNat,"%t0%")  
    ;
```

The **expected net returns to capital** *rore* take the change in end of period capital stock *kapEnd* to beginning of period stock *kStock* exponent the elasticity ( $-p\_RorFlex$ ) into account, considering the number of depreciation years *p\_NDeprYears* multiplied with the net rate of return to capital *rorc*, defined in the two equations *e\_rorePart1* and *e\_ror*, equation (79) in VDM 2018,

```
--- Intermediate term in equation (79) of VDM 2018, considering multiple years of simulation
```

```
e_rorePart1(rsNat,ts) ..  
  
    rorePart1(rsNat,ts)s(rorePart1,'rsNat,ts')  
    =E= (kapEnd(rsNat,ts)/kstock(rsNat,ts))**(-p_RorFlex(rsNat,ts)-p_nDeprYears(ts)+1)s(rorePart1,'rsNat,ts') ;
```

```
--- Expected rate of return, equation (79) in VDM 2018
```

```
e_rore(rsNat,ts) ..  
  
    rore(rsNat,ts)  
    %scale% /rore.1(rsNat,"%t0%")  
    =e= rorc(rsNat,ts) * rorePart1(rsNat,ts)  
    %scale% /rore.1(rsNat,"%t0%")  
    ;
```

As shown above, these two equations are used jointly to define this relation to avoid numerical problems in the solver.

The global bank mechanism aims at equalizing the expected net returns *rore* to savings across regions by changing the distribution of foreign savings *savf* to the different regions. The mechanism, as detailed above, considers returns to the (future) capital stock, corrected for taxation of capital returns and depreciation. Equally, investors dampen their expectation according to *p\_rorFlex* for future increases of the capital relative to the current one. This reflects that an increase in the capital stock, all things equal, will decrease capital returns as other factors will get relatively scarcer. Foreign savings are part of the balance of payments as is the B.o.T. Due to  $B.o.P.=0$ , a change in the foreign savings will also require changes in other position of the B.o.P. In the standard GTAP model, these is solely the B.o.T. A redistribution of foreign savings implies in this case that regions have to change the differences in the value of aggregate exports and imports. This needs to be taken into account for shocks that impact the expected return to investments, such as changes in direct taxes in capital income.

At the benchmark, a so-called risk parameter *p\_risk* ensures that the average global return *rorg* are lined up with the expected returns in each region *rore*. Assume now that the price of mobile capital in a region after direct taxes increases, e.g. by tax reform. This will increase the expected returns in this region and thus attract foreign saving. Increasing the foreign savings in a region will in turn increase total savings and thus investments *xi*. That will change the end of period capital *kapEnd* which will affect the relation between end and beginning stock and thus decrease the expected rate *rore*. Equally, an increase in investment *xi* will likely increase their price index *pi*.

The **value of global net investment** *gblValNetInv* as defined in the equation *e\_gblValNetInv* is derived from the regional net investments *netInv*, their regional prices *pi* and the exchange rate *lcu*, part of equation (82) in VDM 2018:

```
--- value of global net investments in current prices, part of equation (82) in VDM 2018
```

```
e_gblValNetInv(ts) ..
    gblValNetInv(ts)s(gblValNetInv,'ts')
    =E= sum(rNat, (pi(rNat,ts)/lcu(rNat,ts) *netInv(rNat,ts)) $ rsNat(rNat)
    + (pi.l(rNat,ts)/lcu.l(rNat,ts) *netInv.l(rNat,ts)) $ (not rsNat(rNat)))
    s(gblValNetInv,'ts');
```

Note the distinction between regions in the current solve *rsNat* or not. For the latter, investment prices and quantities are fixed by using the level field of the variables. This reflects the case where the model is run in single country mode or during pre-solves where a sub-model is solved for one region, only.

The **total global net investments** *xigl* are defined in the equation *e\_xigl* by up adding the regional net investments *netInv*, equation (86) in VDM 2018. Again, for nations not in the current solve, the net investment is not an endogenous variable as its level field is used.

```
--- Global net investment, equation (86) in VDM 2018
```

```
e_xigl(ts) ..
    xigl(ts)s(gblValNetInv,'ts')
    =e= sum(rNat,
    netInv(rNat,ts) $ rsNat(rNat)
    ---add values of countries not in current solve, but in the global framework
    + netInv.l(rNat,ts) $ (not rsNat(rNat))
    )s(gblValNetInv,'ts');
```

The **average global expected return to capital** *rorg* is the value – net investment *netinv* times saving prices *pi* – weighted average of the regional expected returns *rore* as defined in the two equations *e\_rorg* and *e\_gblValNetInv1*. This equation is for information, only, with the global bank mechanism and solely used if the allocation of net investments is fixed.

```
--- average global expected return, equation (81) in VDM 2018
in case of fixed allocation of investments: average of regional expected rates
rore and rorg should with fixed allocation not influence the solution behavior
```

```
e_rorg(ts) $ (p_RoRFlag eq a_fixedAllocationOfInv) ..
    rorg(ts)s(rorg,'ts') =e= (gblValNetInv1(ts)/gblValNetInv(ts))s(rorg,'ts');
```

```
--- value of global net investments in current prices, part of equation (82) in VDM 2018
```

```
e_gblValNetInv(ts) ..
    gblValNetInv(ts)s(gblValNetInv,'ts')
    =E= sum(rNat, (pi(rNat,ts)/lcu(rNat,ts) *netInv(rNat,ts)) $ rsNat(rNat)
    + (pi.l(rNat,ts)/lcu.l(rNat,ts) *netInv.l(rNat,ts)) $ (not rsNat(rNat)))
    s(gblValNetInv,'ts');
```

The **distribution of the net investments in case of the global bank mechanism** (*RoRFlag eq a\_equalReturnToInv*) is steered by the first part of the following *e\_savf*, equation (80) in VDM 2018, which requires for each region that its risk adjusted expected returns *rore* is equal to the global average one *rorg*:

```

--- Determination of foreign savings, equation (80) in VDM 2019
e_savf(rsNat,ts) $( ( (p_RoRFlag eq a_equalReturnToInv)
or ((p_RoRFlag eq a_fixedAllocationOfInv) $ (not rres(rsNat))))
$ (savf.range(rsNat,ts) ne 0))..
0 =e=
--- equal relative changes in expected returns to investments
across regions, rorg in the global currency
(rore(rsNat,ts)*p_risk(rsNat,ts) - rorg(ts) ) $ (p_RoRFlag eq a_equalReturnToInv)
--- fixed allocation of global investment based on parameter chiInv
+ [(netInv(rsNat,ts) - chiInv(rsNat,ts)*xigbl(ts))s(netInv,'rsNat,ts')] $ (p_RoRFlag eq a_fixedAllocationOfInv);

```

### 2.11.2 Fixed allocation of savings

The case of fixed allocation shares of savings is depicted in the second block of the equation  $e\_savf$  if  $p\_RoRFlag$  is set equal to “ $a\_fixedAllocationOfInv$ ” and is based on parameters  $chiInv$  which reflect the benchmark distribution. Note that the mechanism refers to “capShrFix” in VDM 2018:

```

*
* --- Determination of foreign savings, equation (80) in VDM 2019
*
e_savf(rsNat,ts) $( ( (p_RoRFlag eq a_equalReturnToInv)
or ((p_RoRFlag eq a_fixedAllocationOfInv) $ (not rres(rsNat))))
$ (savf.range(rsNat,ts) ne 0))..
0 =e=
*
* --- equal relative changes in expected returns to investments
* across regions, rorg in the global currency
*
(rore(rsNat,ts)*p_risk(rsNat,ts) - rorg(ts) ) $ (p_RoRFlag eq a_equalReturnToInv)
*
* --- fixed allocation of global investment based on parameter chiInv
*
+ [(netInv(rsNat,ts) - chiInv(rsNat,ts)*xigbl(ts))s(netInv,'rsNat,ts')] $ (p_RoRFlag eq a_fixedAllocationOfInv);
*

```

Note that the residual region  $rres$  is excluded from the mechanism. Its foreign saving is defined by the constraint that the global capital account balance must be equal to zero.

### 2.11.3 Capital account and balance of payments

The global capital account balance  $e\_capAcct$ , equation (84) in VDM 2018, ensures that the global sum of the foreign savings  $savf$  is zero:

```

--- Capital account balance: sum of foreign savings over regions is zero,
defines implicitly regional income of residual region, equation (84) in VDM 2018

e_capAcct(ts) $( ( (sum(rsNat $ (savf.range(rsNat,ts) ne 0),1) gt 1)
or ( (p_RoRFlag eq a_fixedForeignSavings) and sum(rres(rsNat),1)
and ( (pnun.range(ts) gt 0) or (not p_singleCountry) ))) ) ..

0 =e= sum(rNat, savf(rNat,ts) $ rsNat(rNat) + savf.l(rNat,ts) $ (not rsNat(rNat))
+ finc(rNat,ts) $ rsNat(rNat) + finc.l(rNat,ts) $ (not rsNat(rNat))
- p_capTrans(rNat,ts)*pnun(ts))
%scale% /sum(rNat,savf.scale(rNat,ts))
;

```

The equation is only active if (1) there are at least two countries in the current solve where foreign savings are not fixed or (2) the residual region is in the current solve and the foreign savings for the other regions are fixed and the global model is used (*not singleCountry*) or the numeraire is not fixed.

The value of exports at f.o.b. minus imports at c.i.f. prices is defined by the equation  $e\_valFobCif$ : (this is an intermediate equation not found in van der Mensbrugghe 2018):

```
e_valFobCif(rsNat,ts) $ (card(iIn) gt 1) ..
    valFobCif(rsNat,ts)s(regy,'rsNat,ts') =E=
    [ sum( (i,rNat1) $ m_xwFlag(rsNat,i,rNat1), %pofob(rsNat,i,rNat1,ts)*m_xws(rsNat,i,rNat1,ts))
      - sum( (i,rNat1) $ m_xwFlag(rNat1,i,rsNat), %pmcif(rNat1,i,rsNat,ts)*xw(rNat1,i,rsNat,ts) ) ]s(regy,'rsNat,ts');
```

To define the B.o.T., the exports of the transport services for the global transport margins *valMarg* must be considered:

```
e_valMarg(rsNat(rsNat),ts) ..
    valMarg(rsNat,ts)s(valMarg,'rsNat,ts')
    =E= sum( (m,tmg) $ xaFlag(rsNat,m,tmg), xa(rsNat,m,tmg,ts)*m_pa(rsNat,m,tmg,ts)*lcu(rsNat,ts) )s(valMarg,'rsNat,ts');
```

The balance of payments equation *e\_bop* is only a check for the correct setup of the model:

```
e_bop(rsNat(rsNat),ts) ..
    (valsavf(rsNat,ts) + bopSlack(rsNat,ts)*regy(rsNat,ts) - p_capTrans(rsNat,ts)*pnum(ts) + finc(rsNat,ts)
    $$iftheni.co2 %ModulesCO2_Emissions%==on
    + sum(ets_rnat(ets,rsNat) $ p_permitCo2Eq(rsNat,ts), emisEtsTrd.1(rsNat,ts)*emisP(rsNat,ts) $ (card(rsNat) eq 1)
    + emisEtsTrd(rsNat,ts) *emisP(rsNat,ts) $ (card(rsNat) gt 1))
    $$endif.co2
    )s(regy,'rsNat,ts')
    =E= [ - valFobCif(rsNat,ts)
          - valMarg(rsNat,ts)
          ]s(regy,'rsNat,ts');
```

The *bopslack* variable should be zero as the elements of the B.o.P. should add up to zero, including the B.o.T. as an element of the B.o.P. In the GTAP standard model, this is reflected by two definitional equations. First, expenditure for total final demand consists of regional income *regy* (household and government demand plus the share of gross savings financed from regional income *rsav*) and foreign savings *fsav* (= the sole element of the B.o.P besides the B.o.T. in the GTAP standard model), where the foreign savings plus depreciation provide the residual income (positive or negative) besides regional savings to finances final investment demand. Second, *regy* income comprises additional transactions from a trade surplus (or deficit), i.e. the B.O.T. The equation above comprises besides the value of the foreign savings and the B.O.T. (= *valFobCif* + *valMarg*) further transactions between the country and the rest-of-the-world as part of the B.o.P. depicted in optional modules.

## 2.12 Demand sub-nests

### 2.12.1 Motivation

Computable General Equilibrium (CGE) models use regular demand systems to depict final demand for goods and services, typically differentiated by demand from (multiple) household(s), government and investors. Regular means that the Marshallian demand equations are derived from a utility maximization problem. For government and investors, often a Cobb-Douglas or Leontief functional form is employed, while for households, Cobb-Douglas (CD) preferences or Linear Expenditure Systems (LES) are frequently used. Due to often relatively small constant (or so-called commitment) terms in the LES systems, the behavior of the CD and LES systems under price and income changes is often not much different. When the constant terms approach zero, a LES demand system exhibits income and own-price elasticities close to unity and cross-price elasticities close to zero, see also Ho et al. 2020. The resulting linear Engel curves, starting at or close to the origin, don't capture structural changes on the demand side, such as regarding Engel's law. Here, either ad-hoc changes in parameters during simulation are introduced or a change in the functional form is needed.

The GTAP Standard model uses a CDE (Constant Difference in Elasticity) demand system (Hanoch 1975) of which the parameters are derived from an AIDADS demand system estimation by Hertel and Reimer 2004. Its three vectors of parameters allow, considering that one vector is needed for

benchmarking, to calibrate to both own and income elasticities. However, the fixed expansion elasticities of the CDE demand system are also not well suited to depict long-term dynamics in budget shares.

CGEBox (Britz and Van der Mensbrugghe 2016) comprises code for benchmarking of and simulations with the different demand systems (CDE, LES with CD as a special case) for private households mentioned above. Introduced in the context of its long-run component G-RDEM (Britz and Roson 2019, Roson and Britz 2021), CGEBox makes a step further and comprises as an option an empirically estimated MAIDADS demand system (Britz 2021) which exhibits exponential Engel Curves. The MAIDADS system is the recommended option for long-run analysis in CGEBox and can be combined with an option ensuring plausible development of calorie intakes (Britz 2020).

None of the demand system above is however flexible enough to allow benchmarking to cross-price elasticities. Using second-order flexible demand systems such as the Generalized Leontief Demand System (cf. Diewert and Wales 1989) employed in the Partial Equilibrium model CAPRI (Britz and Witzke 2014) is unusual in CGE analysis. Instead, following the structure of the production functions, CGE models employ nested demand systems where the lower nests are of the CES (Constant-Elasticity of Substitution) type. This implies expenditures elasticity of unity for commodity or services bundled in one nest, such that their income elasticities depend on the aggregate commodity bundle they are linked to in the top-level demand system.

### **2.12.2 Implementation in model**

Demand sub-nests aggregate individual Armington demands in final demand by private households, government or savings to utility aggregates based on a CES function. Such sub-nests are not part of the GTAP Standard Model. The resulting nests can be either part of the top-level demand function or linked into other demand nests. That mechanism increases the flexibility of depicting substitution relations between individual products in final demand beyond what the top-level final demand system offers, similar to the flexible nesting in the production function. The approach allows to drive the total demand for a product by a agent by shares on different nests, similar to sourcing a product from different production activities. It makes little sense to run shocks with a CGE model comprising products that are close substitutes without CES sub-nests if these shocks lead to differentiated impacts on the product prices, as cross-price effects cannot be controlled by the usual final demand systems as discussed next.

For the CD demand system, all cross-price effects are zero. The LES demand system, cross-price elasticities  $\varepsilon_{i,j}$  are equal to minus the marginal budget share over the current budget share for product  $i$ , times the value of the commitment of product  $j$  on total expenditures. Accordingly, they are likely quite small for two reasons. First, commitments terms tend to be small compared to benchmark demand. And second, with a detailed product list, the value of such a single commitment relative to total expenditure is likely quite small. Even more important is the fact that the LES demand system has two vectors of product specific parameters, only. It is common to keep the marginal budget share parameters unchanged as they relate to the income elasticities and to use the commitments terms for calibration. It is hence impossible to even control for own price effects.

The case of the CDE demand system as found in the GTAP Standard model is more evolved. It comprises three vectors of product specific parameters. This allows to target own and income elasticities, and use the remaining vector for benchmarking. But this also implies that individual cross-price effects cannot be controlled. In the CDE calibration as used in CGEBox, the substitution

elasticities depend on the so-called substitution parameters  $\alpha$  that are not changed during benchmarking. They are part of the parameters delivered with the GTAP data base and used by the GTAP center to target own price elasticities. The Allen-Uzawa elasticities of substitution relate to these parameters as follows (Chen 2017<sup>14</sup>):

$$\sigma_{ij} = \alpha_i + \alpha_j - \sum_k \theta_k \alpha_k + \frac{\alpha_i}{\theta_i} \text{ if } i = j$$

Where  $\theta_i$  depicts the budget shares. The uncompensated demand elasticities  $\sigma_{ij}^m$  can be derived from these substitution elasticities by subtracting the income elasticity  $\eta_i$  and multiplying with the budget shares:

$$\sigma_{ij}^m = \theta_j (\sigma_{ij} - \eta_i)$$

Applying these formulae using the substitution parameters and budget shares at the benchmark suggests that the cross-price elasticities from the CDE demand system are all extremely small.

In summary, rather independent of the top-demand system used, none offers degrees of freedom to target individual cross-price effects and all tend to have quite small ones. This motivates why it is useful to complement the top-level demand system with CES sub-nests.

The **average price for a demand nest**  $pdNest$  for a such a nest  $dNest$  and the final demand agent  $fdn$  is defined by dual price aggregators which distinguish the CD from the CES/Leontief case depending on the substitution elasticity  $p\_sigmaFDNest$ . In both cases, the price index reflects the contribution of individual products  $i$  based on their share parameter  $alphaa$  and the contribution of sub-nests based on their share parameter  $p\_alphaDN$ .

```

pNest(rs,dnest,fdn,ts) $ sum(allDNests,p_alphaDN(rs,dNest,fdn,allDNests,ts)) ..

pdNest(rs,dNest,fdn,ts) $ (not addCES(rs,dNest,fdn))
+ pdNest(rs,dNest,fdn,ts) $ addCES(rs,dNest,fdn)
==>

--- CES or Leontief case with p_sigmaFDNest<>1
+ {
  { sum(dNest_i_fd(dNest,i,fdn))
    %scale%
    ,p_alphaa(rs,i,fdn,dNest,ts) * n_pa(rs,i,fdn,ts) ** (1 $ (not addCES(rs,dNest,fdn)) - p_sigmaFDNest(rs,dNest,fdn))
    / (lambdaa(rs,i,fdn,ts) ** (1-p_sigmaFDNest(rs,dNest,fdn)) $ (not addCES(rs,dNest,fdn)) + p_sigmaFDNest(rs,dNest,fdn) $ addCES(rs,dNest,fdn)))
  }
+ sum(dNest_n_fd(dNest,dNest1,fdn))
  %scale%
  ,p_alphaDN(rs,dNest1,fdn,dNest,ts) * pdNest(rs,dNest1,fdn,ts) ** (1 $ (not addCES(rs,dNest,fdn)) - p_sigmaFDNest(rs,dNest,fdn))
  / (lambdaDN(rs,dNest1,fdn,dNest,ts) ** (1-p_sigmaFDNest(rs,dNest,fdn)) $ (not addCES(rs,dNest,fdn)) + p_sigmaFDNest(rs,dNest,fdn) $ addCES(rs,dNest,fdn)))
} ** (1/(1 $ (not addCES(rs,dNest,fdn)) - p_sigmaFDNest(rs,dNest,fdn)))
} $ (((p_sigmaFDNest(rs,dNest,fdn) ne 1) or addCES(rs,dNest,fdn))
  %scale% $ ( not singleDNest(rs,dNest,fdn))
)

```

<sup>14</sup> Chen, Y. H. H. (2017). The calibration and performance of a non-homothetic cde demand system for cge models. *Journal of Global Economic Analysis*, 2(1), 166-214

```

--- CD case with p_sigmaFDNest==1
+ [
    ---- contribution of individual commodity demands to price aggregator
    prod(dNest_i_fd(dNest,i,fdn)
    %scale% $ p_alphaa(rs,i,fdn,dNest,ts)
    , (m_pa(rs,i,fdn,ts)/(p_alphaa(rs,i,fdn,dNest,ts)*lambdaa(rs,i,fdn,ts)))*p_alphaa(rs,i,fdn,dNest,ts))

    ---- contribution of sub-nests under the current nest to price aggregator
    * prod(dNest_n_fd(dNest,dNest1,fdn)
    %scale% $ p_alphadN(rs,dNest1,fdn,dNest,ts)
    , (pdNest(rs,dNest1,fdn,ts)/(p_alphadN(rs,dNest1,fdn,dNest,ts)*lambdaDn(rs,dNest1,fdn,dNest,ts)))*p_alphadN(rs,dNest1,fdn,dNest,ts))

    * p_axdNestCD(rs,dNest,fdn,ts)

] $ (((p_sigmaFDNest(rs,dNest,fdn) eq 1) and (not addCES(rs,dNest,fdn)))
    %scale% $ ( not singleDNest(rs,dNest,fdn))
)

--- case where only one product or nest is left
%scale% + [ sum(dNest_i_fd(dNest,i,fdn) $ p_alphaa(rs,i,fdn,dNest,ts), m_pa(rs,i,fdn,ts)/lambdaa(rs,i,fdn,ts))
%scale% + sum(dNest_n_fd(dNest,dNest1,fdn) $ p_alphadN(rs,dNest1,fdn,dNest,ts), pdNest(rs,dNest1,fdn,ts)/lambdaDn(rs,dNest1,fdn,dNest,ts))
%scale% ] $ singleDNest(rs,dNest,fdn)
;
    
```

In mathematical notation:

$$\begin{aligned}
 PdNest_{r,dNest,fdn} &= \left[ \sum_{i \in dNest} \alpha_{r,i,fdn}^{dNest} \left( \frac{PA_{r,i,fdn}}{\lambda_{r,i,fdn}} \right)^{\sigma_{fdn}^{dNest}} \right. \\
 &\quad \left. + \sum_{dNest1 \in dNest} \alpha_{r,dNest1,fdn}^{dNest} \left( \frac{PdNest_{r,dNest1,fdn}}{\lambda_{r,dnest,fdn}} \right)^{\sigma_{fdn}^{dNest}} \right]^{\frac{1}{(1-\sigma_{fdn}^{dNest})}}
 \end{aligned}$$

With the list of final demands  $fdn = \{gov, inv, h\}$

The notation already underlines that such nests can comprise other nests. Note that the composite price definition in equation  $e\_pdNest$  considers the case of an additive CES function. This is not shown in the mathematical notation.

The **demand for a sub-nest**  $xdNest$  depends on the total demand top ( $inv$  or  $gov$ ) or is driven by the demand of a sub-nest  $dNest1$ . For household demand, equations for nests as part of the top-level demand function (CDE or LES type) has been discussed in previous sections.

```

e_xdNest(rs,dnest,fdn,ts) $ (sum(alldNestS,p_alphaDN(rs,dNest,fdn,alldNests,ts)) and (not sum(dNest_n_fd("top",dnest,h) $ sameas(fdn,h,1)))) ..
xdNest(rs,dNest,fdn,ts)s(xdNest,'rs,dNest,fdn,ts')
=e=
[
--- driven by top nest demand in case of gov/inv (CES or CD)
Final household demand has separate equation to reflec the more complex CDE case
+ sum(dNest_n_fd("top",dnest,"gov")
  %scale% $ p_alphaDN(rs,dNest,"gov","top",ts)
  , p_alphaDN(rs,dNest,"gov","top",ts) * xg(rs,ts)
  * lambdaDN(rs,dNest,"gov","top",ts)**(p_sigmag(rs)-1)
  * (pg(rs,ts)/pdNest(rs,dNest,"gov",ts))**p_sigmag(rs)) $ sameas(fdn,"gov")
+ sum(dNest_n_fd("top",dnest,"inv")
  %scale% $ p_alphaDN(rs,dNest,"inv","top",ts)
  , p_alphaDN(rs,dNest,"inv","top",ts) * xi(rs,ts)
  * lambdaDN(rs,dNest,"inv","top",ts)**(p_sigmai(rs)-1)
  * (pi(rs,ts)/pdNest(rs,dNest,"inv",ts))**p_sigmai(rs)) $ sameas(fdn,"inv")
-- driven by another nest
+ sum(dNest_n_fd(dNest1,dNest,fdn)
  %scale% $ p_alphaDN(rs,dNest,fdn,dNest1,ts)
  , p_alphaDN(rs,dNest,fdn,dNest1,ts) * xdNest(rs,dNest1,fdn,ts)
  --- preference shifter
  * lambdaDN(rs,dNest,fdn,dNest1,ts)**((p_sigmaFDNest(rs,dNest,fdn)-1)
  $ (not addCes(rs,dNest,fdn)) - p_sigmaFDNest(rs,dNest,fdn) $ addCes(rs,dNest,fdn)))
  --- price effects
  * ([pdNest(rs,dNest1,fdn,ts) $ (not addCes(rs,dNest1,fdn)) + pddNest(rs,dNest1,fdn,ts) $ addCes(rs,dNest1,fdn)]
  /pdNest(rs,dNest,fdn,ts))**p_sigmaFDNest(rs,dNest1,fdn))
]s(xdNest,'rs,dNest,fdn,ts');

```

The **Armington demands**  $xa$  driven by a sub-nest are defined in the equation  $e\_xdNest$  and use the usual CES-structure, i.e. the share parameter  $\alpha$ , the sub-nest total demand  $xdNest$  and the price relation exponent the substitution elasticity  $p\_sigmaFDNest$  as well a preference shifter  $\lambda$ . Note that the demand can be driven by multiple nests.

```

--- individual agent s Armington demands under sub-nests
e_xaDNest(rs,i,fdn,ts) $ (xaiFlag(rs,i,fdn) $ sum(dNest_i_fd(dNest,i,fdn),1) $ iIn(i)) ..
xa(rs,i,fdn,ts)s(xa,'rs,i,fdn,ts') =e=
sum(dNest_i_fd(dNest,i,fdn),
  p_alpha(rs,i,fdn,dNest,ts)*xdNest(rs,dnest,fdn,ts)s(xa,'rs,i,fdn,ts')
  * [
    --- CES case with price effect
    {
      lambda(rs,i,fdn,ts)**(( [ p_sigmaFDNest(rs,dnest,fdn)-1] $ (not addCes(rs,dNest,fdn))
      - p_sigmaFDNest(rs,dnest,fdn) $ addCes(rs,dNest,fdn) ))
      * ([ pdNest(rs,dNest,fdn,ts) $ (not addCes(rs,dNest,fdn))
      + pddNest(rs,dnest,fdn,ts) $ addCes(rs,dNest,fdn) ]/m_pa(rs,i,fdn,ts)) **p_sigmaFDNest(rs,dNest,fdn)
    }
    $ p_sigmaFDNest(rs,dNest,fdn)
    --- Leontief case with fixed composition, no price effect
  + lambda(rs,i,fdn,ts) $ (p_sigmaFDNest(rs,dNest,fdn) eq 0)
  ])
+ sum(r_r(subr,rs),p_subrDem(subr,i,fdn,ts)/m_pa(rs,i,fdn,ts))s(xa,'rs,i,fdn,ts');

```

### 2.12.3 Benchmarking of demand sub-nests

In order to benchmark the structure above, the analyst has to define the underlying nesting structures and choose appropriate substitution elasticities. The nesting structures are based on set-definitions, accompanied with the chosen elasticities, such as in:

```

dNest("non-electric") = YES;
dNest_i_fd("non-electric",coal,fdn) = YES;
dNest_n_fd("non-electric","non-coal",fdn) = YES;
p_sigmaFDNest(r,"non-electric",fdn) = 0.50;

dNest("non-coal") = YES;
dNest_i_fd("non-coal",nonCoal,fdn) = YES;
p_sigmaFDNest(r,"non-coal",fdn) = 1.00;

```

as part of the CES-nesting structure borrowed from GTAP-Power (Peters 2016).



The demand nesting allows to distribute individual products or nests to multiple nests in shares, such as in:

```
dNest("housing") = YES;
dNest_i_fd("housing","dwe-c",h) = YES;
dNest_n_fd("housing","energy",h) = YES;
dNest_n_fd("top","housing",h) = YES;
dNestShare(r,"housing","energy",h) = 0.5;
dNest("housing") = YES;
```

The benchmarking (*cal/cal\_dnest.gms*) of the nested structures consists to three major steps:

1. Checking that the nesting structures introduced by the analyst are consistent
2. Deriving the share parameters for the sub-nests
3. Modification of the top-level demand system parameters

First some symbols are defined. Nests for the international trade margin demand agent *trdmg* are refused as they are not supported by the code.

```
alias(*,uDom);
set wrongName;

set dNesting(*,*,fd);
parameter dNestingErrors(r,*,*,fd);

allDNests(dNest) = YES; allDNests("top") = YES;

set nestsToDo,curNest(*);
alias(nestsToDo,nestsToDo1);

set topdNest(*,i,fd);
*
* --- these nests are not supported
*
if(sum(dnest_i_fd(dNest,i,"trdmg"),1),
    abort "Demand sub-nests for 'trdmg' are not supported",dnest_i_fd);

if(sum(dnest_n_fd(dNest,dNest1,"trdmg"),1),
    abort "Demand sub-nests for 'trdmg' are not supported",dnest_n_fd);

if(sum(dnest_n_fd("top",dNest1,"trdmg"),1),
    abort "Demand sub-nests for 'trdmg' are not supported",dnest_n_fd);
```

The reminder of the code is executed only if nests are present:

```
if (card(dNest),
```

Otherwise, all demand is assigned to the top nest with shares of unity:

```
dNestShare(rs,"top",i,fdn) $ xa.1(rs,i,fdn,"%t0%") = 1;
dNest_i_fd("top",i,fdn) = yes;
```

### **Checks during benchmarking**

The first part of the checks uses the universal domain to make sure that only element names found in the dynamic set *allDNests* are used. The use of a dynamic set allows to add flexibly nests to the model at run time. Using run time statements to define set membership instead of fixed static set is important to reflect to let the code consider the detail of different data bases. For instance, different electricity types are only present if the GTAP-Power Data Base is used in appropriate differentiation.

The tests are coded as follows:

```

:
: *
: *   --- check for typos, as dest is not domain checked to allow for flexible includes
: *
:
: wrongName(uDom) $ sum(dNest_i_fd(uDom,i,fdn) $ (not allDNests(uDom)),1) = YES;
: if ( card(wrongName),
:   abort "Typo in dNest_i_fd on first position, not a final demand sub-nest",dNest_i_fd,dNest,wrongName);
:
: wrongName(uDom) $ sum(dNest_n_fd(dNest1,uDom,fdn) $ (not allDNests(uDom)),1) = YES;
: if ( card(wrongName),
:   abort "Typo in dNest_n_fd on second position, not a final demand sub-nest",dNest_n_fd,wrongName);
:
: wrongName(uDom) $ sum(dNest_n_fd(uDom,dNest,fdn) $ (not allDNests(uDom)),1) = YES;
: if ( card(wrongName),
:   abort "Typo in dNest_n_fd on first position, not a final demand sub-nest",dNest_n_fd,wrongName);
:
:

```

Next, we define a parameter which shows the nesting structure:

```

:
:   ---
:
:   --- build tNesting structure
:
: dNesting("top",i,fdn) $ (not sum(dNest_i_fd(dNest,i,fdn),1)) = YES;
: dNesting(allDNests,dNest,fdn) $ dNest_n_fd(allDNests,dNest,fdn) = YES;
: dNesting(allDNests,i,fdn) $ dNest_i_fd(allDNests,i,fdn) = YES;

```

And assign shares of unity if not defined in the code:

```

:
:   --- set nest shares to unity if not set by user
:
: dNest_i_fd("top",i,fdn) $ (not sum(dNest_i_fd(dNest,i,fdn),1)) = YES;
: dNestShare(rs,allDNests,i,fdn) $ (dNest_i_fd(allDNests,i,fdn) $ (not dNestShare(rs,allDNests,i,fdn) $ (xFlag(rs,i,fdn) or xFlag(rs,fdn,i))) = 1;
: dNestShare(rs,allDNests,dNest,fdn) $ (dNest_n_fd(allDNests,dNest,fdn) $ (not dNestShare(rs,allDNests,dNest,fdn))) = 1;
:

```

The following two checks ensure that the distribution shares add up to unity. They will hence also capture cases where nests are not correctly integrated in the overall structure and are dangling:

```

:
:   --- Check that each commodity is fully distributed (shares add up to unity)
:
: dNestingErrors(rs,dNest,i,fdn) = sum(dNesting(allDNests,i,fdn),dNestShare(rs,allDNests,i,fdn));
: dNestingErrors(rs,dNest,i,fdn) $ (dNestingErrors(rs,dNest,i,fdn) eq 1) = 0;
: abort $ card(dNestingErrors)
: "Error: Final demand for a commodity in final demand sub-nests is not correctly distributed (not at all or shares don't add up to unity)",
:   dNestingErrors,dNesting;
:
:   --- Check that each nest is fully distributed (shares add up to unity)
:
: dNestingErrors(rs,dNest,dNest1,fdn) = sum(dNesting(allDNests,dNest1,fdn),dNestShare(rs,allDNests,dNest1,fdn));
: dNestingErrors(rs,dNest,dNest1,fdn) $ (dNestingErrors(rs,dNest,dNest1,fdn) eq 1) = 0;
: abort $ card(dNestingErrors)
: "Error: Final demand for a sub-nests in final demand sub-nests is not correctly distributed (not at all or shares don't add up to unity)",
:   dNestingErrors,dNesting;

```

### Sequential assignment of share parameters during benchmarking

The share parameters in the nest must be defined bottom-up, from the lowest nests which can comprise single products, only, to the top-level demand system. In the beginning, no nest has already properly defined price and quantity indices:

```

nestsToDo(dNest) = YES;
option kill=xdNest.1;

```

We iterate until the set of not yet defined nests is empty and define in each iteration first the current nests to work on:

```

-
While( card(nestsToDo),

    --- we can calculate all nests which do not require information on nests
    not yet calculated

    option kill=curNest;
    curNest(nestsToDo) $ (not sum(dNest_n_fd(nestsToDo,nestsToDo1,fdn),1)) = YES;

```

At the end of this while statement, the now properly defined nests are taken from the set of nests to work on:

```
*
*      --- and delete from lists of nests to process
*
      nestsToDo(curNest) = NO;
```

As Armington prices are defined as unity at the benchmark, the resulting CES price indices for the demand nests are unity as well:

```
*
      pdNest.l(rs,curNest,fdn,t0) = 1;
```

The quantity indices for the nests consider the demand quantities  $xa$  of single products  $i$  mapped to them and demand quantities of demand sub-nests  $xdNest$ , in both cases considering the distribution shares  $dNestShare$ :

```
--- calculate quantity index

xdNest.l(rs,curNest,fdn,t0)
= [      sum(dNest_i_fd(curNest,i,fdn),
            xa.l(rs,i,fdn,t0)*pa.l(rs,i,fdn,t0)*dNestShare(rs,curNest,i,fdn))

      + sum(dNest_n_fd(curNest,dNestl,fdn),
            xdNest.l(rs,dNestl,fdn,t0)*pdNest.l(rs,dNestl,fdn,t0)*dNestShare(rs,curNest,dNestl,fdn))

      ] / pdNest.l(rs,curNest,fdn,t0);

pdNest.l(rs,curNest,fdn,t0) $ (not xdNest.l(rs,curNest,fdn,t0)) = 0;
```

The share parameters for single products are defined as usual for the CES functional form as follows:

```
--- calculate cost share parameters for products in current nest

alphaa.fx(rs,i,fdn,curNest,t0) $ sum(dNest_i_fd(curNest,i,fdn), xa.l(rs,i,fdn,t0))
= sum(dNest_i_fd(curNest,i,fdn), xa.l(rs,i,fdn,t0)*dNestShare(rs,curNest,i,fdn)/xdNest.l(rs,curNest,fdn,t0)
      * (pa.l(rs,i,fdn,t0)/pdNest.l(rs,curNest,fdn,t0))**p_sigmaFDNest(rs,curNest,fdn)) ;
```

Similar, for sub-nests:

```
--- calculate cost share parameters for sub-nests in current nest

p_alphaDN(rs,dNest,fdn,curNest,t0) $ sum(dNest_n_fd(curNest,dNest,fdn) $ (xdNest.l(rs,curNest,fdn,t0) and xdNest.l(rs,dNest,fdn,t0)),1)
= sum(dNest_n_fd(curNest,dNest,fdn), xdNest.l(rs,dNest,fdn,t0)*dNestShare(rs,curNest,dNest,fdn)/xdNest.l(rs,curNest,fdn,t0)
      * (pdNest(rs,dNest,fdn,t0)/pdNest.l(rs,curNest,fdn,t0))**p_sigmaFDNest(rs,curNest,fdn)) ;
```

In case of CD relation (substitution elasticity is unity), a multiplier is needed as seen in the model equation above:

```
p_axdNestCD(rs,curNest,fdn,t0) $ (xdNest.l(rs,curNest,fdn,t0) $ (p_sigmaFDNest(rs,curNest,fdn) eq 1))
= pdNest.l(rs,curNest,fdn,t0) /
  [      prod(dNest_i_fd(curNest,i,fdn) $ alphaa(rs,i,fdn,curNest,t0),
              (pa.l(rs,i,fdn,t0)/alphaa(rs,i,fdn,curNest,t0))**alphaa(rs,i,fdn,curNest,t0))

      *      prod(dNest_n_fd(curNest,dNestl,fdn) $ p_alphaDN(rs,dNestl,fdn,curNest,t0),
                  (pdNest(rs,dNestl,fdn,t0)/p_alphaDN(rs,dNestl,fdn,curNest,t0))**p_alphaDN(rs,dNestl,fdn,curNest,t0))

      ];
```

## Modification of the top-level demand system parameters and other final steps

In order to reduce the computational burden, cases where a nest in a region comprises one product or sub-nest, only are treated differently in the code (see above). This is captured by the set *singleDNest* (see also model equation above):

```
singleDNest(rs,dNest,fdn) $ ( ( sum(dNest_n_fd(dNest,dNest1,fdn) $ xdNest.1(rs,dNest1,fdn,"%t0%"),1)
+ sum(dNest_i_fd(dNest,i,fdn) $ xa.1(rs,i,fdn,"%t0%"),1) ) eq 1) $ xdNest.1(rs,dNest,fdn,"%t0%") = YES;

p_alphaDN(rs,dNest1,fdn,dNest,"%t0%") $ sum(dNest_n_fd(dNest,dNest1,fdn) $ (singleDNest(rs,dNest,fdn) $ xdNest.1(rs,dNest1,fdn,"%t0%")),1) = 1;
alphaa.fx(rs,i,fdn,dNest,"%t0%") $ sum(dNest_i_fd(dNest,i,fdn) $ (singleDNest(rs,dNest,fdn) $ xa.1(rs,i,fdn,"%t0%")),1) = 1;

p_sigmaFDNest(rs,dNest,fdn) $ singleDNest(rs,dNest,fdn) = 0;
```

For the CDE demand system which is the default case, the expansion and substitution parameters for sub-nests are defined as follows:

```
option kill=bhdn.1,kill=ehdn.1;

nestsToDo(dNest) = YES;
While( card(nestsToDo),
  option kill=curNest;
  curNest(nestsToDo) $ (not sum(dNest_n_fd(nestsToDo,nestsToDo1,h),1)) = YES;
*
*   --- aggregate parameters, using expenditures as weights
*
  bhdn.fx(rs,curNest,t0)
    $ ( sum(dNest_i_fd(curNest,i,h), xcshr.1(rs,i,h,t0)*yc.1(rs,h,t0)*dNestShare(rs,curNest,i,h))
    + sum(dNest_n_fd(curNest,dNest1,h), xdNest.1(rs,dNest1,h,t0)*dNestShare(rs,curNest,dNest1,h)) )
    = ( sum(dNest_i_fd(curNest,i,h), xcshr.1(rs,i,h,t0)*yc.1(rs,h,t0)*dNestShare(rs,curNest,i,h)*bh.1(rs,i,t0))
    + sum(dNest_n_fd(curNest,dNest1,h), bhdn.1(rs,dNest1,t0)*xdNest.1(rs,dNest1,h,t0)*dNestShare(rs,curNest,dNest1,h)) )
    / ( sum(dNest_i_fd(curNest,i,h), xcshr.1(rs,i,h,t0)*yc.1(rs,h,t0)*dNestShare(rs,curNest,i,h))
    + sum(dNest_n_fd(curNest,dNest1,h), xdNest.1(rs,dNest1,h,t0)*dNestShare(rs,curNest,dNest1,h)) );

  ehdn.fx(rs,curNest,t0)
    $ ( sum(dNest_i_fd(curNest,i,h), xcshr.1(rs,i,h,t0)*yc.1(rs,h,t0)*dNestShare(rs,curNest,i,h))
    + sum(dNest_n_fd(curNest,dNest1,h), xdNest.1(rs,dNest1,h,t0)*dNestShare(rs,curNest,dNest1,h)) )
    = ( sum(dNest_i_fd(curNest,i,h), xcshr.1(rs,i,h,t0)*yc.1(rs,h,t0)*eh.1(rs,i,t0)*dNestShare(rs,curNest,i,h))
    + sum(dNest_n_fd(curNest,dNest1,h), ehdn.1(rs,dNest1,t0)*xdNest.1(rs,dNest1,h,t0)*dNestShare(rs,curNest,dNest1,h)) )
    / ( sum(dNest_i_fd(curNest,i,h), xcshr.1(rs,i,h,t0)*yc.1(rs,h,t0)*dNestShare(rs,curNest,i,h))
    + sum(dNest_n_fd(curNest,dNest1,h), xdNest.1(rs,dNest1,h,t0)*dNestShare(rs,curNest,dNest1,h)) );

  nestsToDo(curNest) = NO;
);
```

Based on this and the given parameters for the single products, the share parameters of the CDE demands system are defined as:

```
xcshrDN.1(rs,dNest,h,t0) $ sum(dNest_n_fd("top",dNest,h), xdNest.1(rs,dNest,h,t0)*dNestShare(rs,"top",dNest,h))
= xdNest(rs,dNest,h,t0)*dNestShare(rs,"top",dNest,h)*pdNest(rs,dNest,h,t0)/yc.1(rs,h,t0);

alphaa.fx(rs,i,h,"top",t0) $ (xaFlag(rs,i,h) and sum(dNest_i_fd("top",i,h),1))
= ((xcshr.1(rs,i,h,t0)*dNestShare(rs,"top",i,h)/bh.1(rs,i,t0))
* ((yc.1(rs,h,t0)/pop.1(rs,h,t0)*1000)/pa.1(rs,i,h,t0))*bh.1(rs,i,t0))
* (uh.1(rs,h,t0)*(-eh.1(rs,i,t0)*bh.1(rs,i,t0)))
/ ( sum(dNest_i_fd("top",j,h), xcshr.1(rs,j,h,t0)*dNestShare(rs,"top",j,h)/bh.1(rs,j,t0))
+ sum(dNest1 $ bhDN.1(rs,dNest1,t0), xcshrDN.1(rs,dNest1,h,t0)/bhDN.1(rs,dNest1,t0)) );

p_alphaDN(rs,dNest,h,"top",t0) $ sum(dNest_n_fd("top",dNest,h), xdNest.1(rs,dNest,h,t0)*dNestShare(rs,"top",dNest,h))
= ((xcshrDN.1(rs,dNest,h,t0)/bhDN.1(rs,dNest,t0))
* ((yc.1(rs,h,t0)/pop.1(rs,h,t0)*1000)/pdNest.1(rs,dNest,h,t0))*bhDN.1(rs,dNest,t0))
* (uh.1(rs,h,t0)*(-ehDN.1(rs,dNest,t0)*bhDN.1(rs,dNest,t0)))
/ ( sum(dNest_i_fd("top",j,h), xcshr.1(rs,j,h,t0)*dNestShare(rs,"top",j,h)/bh.1(rs,j,t0))
+ sum(dNest1 $ bhDN.1(rs,dNest1,t0), xcshrDN.1(rs,dNest1,h,t0)/bhDN.1(rs,dNest1,t0)) );
```

## 2.12.4 References

Britz, W. (2016). CGEBox model documentation. *University Bonn, Institute for Food and Resource Economics*. [http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox\\_GUI.pdf](http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox_GUI.pdf).

Britz, W. (2020): Maintaining plausible calorie intakes, crop yields and crop land expansion in long-run simulations with Computable General Equilibrium Models. Discussion Paper 2020:2, University Bonn, Institute for Food and Resource Economics, <https://ageconsearch.umn.edu/record/302922>.

Britz, W. (2021): Estimating a global MAIDADS demand system considering demography, climate and norms, *Bio-based and Applied Economics* 10(3): 219-238

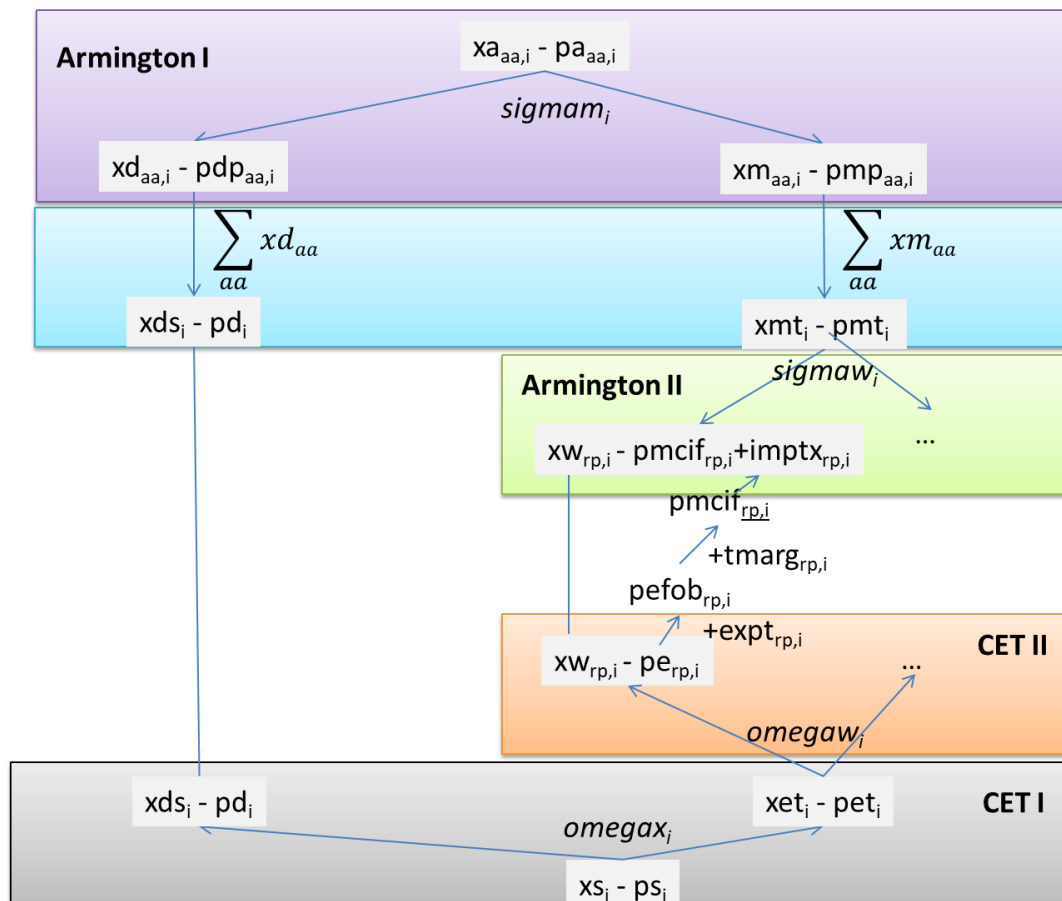
- Britz, W., & Witze, P. (2014). CAPRI model documentation, version 2014. University Bonn.
- Britz, W., Roson, R. (2019): G-RDEM: A GTAP-Based Recursive Dynamic CGE Model for Long-Term Baseline Generation and Analysis, *Journal of Global Economic Analysis* 4(1): 50-96
- Diewert, W. E., & Wales, T. J. (1989). Flexible functional forms and global curvature conditions, NBER Technical Paper Series No. 40, [https://www.nber.org/system/files/working\\_papers/t0040/t0040.pdf](https://www.nber.org/system/files/working_papers/t0040/t0040.pdf)
- Hanoch, G. (1975). Production and demand models with direct or indirect implicit additivity. *Econometrica: Journal of the Econometric Society*, 395-419
- Hertel, Thomas & Reimer, Jeffrey. (2004). Estimation of International Demand Behavior for Use with Input–Output Based Data. *Economic Systems Research*. 16. 347-366. 10.1080/0953531042000304245.
- Ho, M., Britz, W., Delzeit, R., Leblanc, F., Roson, R., Schuenemann, F., Weitzel, M. (2020): Modelling Consumption and Constructing Long-Term Baselines in Final Demand, *Journal of Global Economic Analysis* 5(1): 63-80
- Peters, J. C. (2016). GTAP-E-Power: an electricity-detailed economy-wide model. *Journal of Global Economic Analysis*, 1(2), 156-187
- Roson, R., Britz, W. (2021): Simulating long run structural change with a dynamic general equilibrium model, *Int. J. Computational Economics and Econometrics* 11(4): 368-404

## **2.13 International trade and domestic sales, and related prices**

### **2.13.1 Overview**

The graphic below depicts the main relations in international trade and the distribution of supply in the model's equation structure. The top-level Armington nest in the uppermost box distributes the Armington demand for each agent to demand for the domestic and imported origin. Next, the demands for the imported and domestic origin are aggregated over the agents. The resulting total import demand  $xmt$  is then split up into demands of the different exporters  $xw$ , driven by the c.i.f. price plus import taxes. The difference between the c.i.f. and f.o.b. price are the endogenous transport margins. Taking export taxation or subsidization into account, the export prices  $pe$  in each exporter region are derived.

Distribution of supply  $xs$  to total exports  $xet$  and domestic sales  $xds$  is driven by a transformation nest depicted in the lowest box. The transformation can also be infinite as the default case. Distribution of the total exports  $xet$  to different destination  $xw$  is handled by a second transformation nest, again, with infinite transformation as the default case.



**Figure 3: Overview on distribution of supply and sourcing of demand**

Note that the set of Armington agents depends on the chosen structure of the model:

1. In the default layout, it comprises the list of sectors, one aggregate private household, government and investment demand.
2. Alternative, that differentiation can be completely removed, i.e. the shares in the upper nests are identical for each agent, or can be defined identical across sectors.

The MRIO extension introduces an own set of equations (see sub-section “Model equations”) which allows to dis-aggregate also the lower Armington nests, i.e. the bi-lateral demands, by agent. Furthermore, note that the Melitz and Krugmann extensions use one nest only considering love-of-variety.

### 2.13.2 Individual equations

The **agent specific prices for imports**  $pmp$  as defined in the equation  $e\_pmp$ , equation (46) in VDM 2018, reflects the average import price  $pmt$  and agent specific import taxes  $mintx$  plus tax shifts  $itxshft$  and emission taxes:

```
*
*   --- Price for the imported origin, equation (46) in VDM 2018
*
e_pmp(rs,iIn(i),aa,ts) $ (p_alpham(rs,i,aa,ts) $ (not iMrio(i)) $ (not a(aa) or sum(sameas(aa,aIn),1))) ..
    pmp(rs,i,aa,ts) = m_pmpDef(rs,i,aa,ts);
```

According to the macro *m\_pmpDef* that adds ad-valorem use taxes to the average price of imports defined by the macro *m\_pmt* and per unit emission taxes:

```
--- the macro m_pmp substitutes out the imports' sales prices
$macro m_pmpDef(r,i,aa,t) ( \
    --- average import price ( pmt or pmtMrio )
    m_pmt(r,i,aa,t) * (1 + mintx(r,i,aa,t) + itxshft(r,i,t)) \
    --- add emission taxes for imported use
    $if %ModulesCO2_Emissions==on + emis(r,t) * p_emidCo2eqTaxed(r,i,aa,t) * (1-redemiaCo2eq(r,aa,t) * p_nCo2ShareI(r,i,aa) $ a(aa) + p_emiShift(r,i,aa,t)) \
)
```

Note that the equation is not active if the MRIO extension is switched on for that product (*not iMrio(i)*) as in that case, the basis to derive the agent specific is specific to group of agents and not equal to *pmt*:

```
$iftheni.MRIO "%modulesGTAP_MRIO%"=="on"
    --- we have four MRIO agents (all intermediate demand, private / gov consumption, savings)
    The first part is for products under the normal Armington, we use the standard import price pmt
    The second part selects the matching MRIO import prices
    $macro m_pmt(r,i,aa,t) (pmt(r,i,t) $ (not iMrio(i)) + sum(mrio_A1(mrioA,aa), pmtMrio(r,i,mrioA,t)) $ iMrio(i))
$else.MRIO
    $macro m_pmt(r,i,aa,t) pmt(r,i,t)
$endif.MRIO
```

Equally, the equation requires a non-zero share parameters for imports *p\_alpham*. The Melitz extension sets this share parameter to zero for the products handled in the Melitz model and accordingly removes the equation for these products. Finally, the model will for the default case substitute out the *pmp* prices for intermediate demand (*not a(aa)*).

Similar, the **agent specific prices for domestic origin** *pdp* reflect price in the domestic markets, defined in the equation *e\_pd*, equation (45) in VDM 2018. These are equal to sectoral prices *ps* under infinite transformation or equal to domestic sales prices *pd* in case of non-infinite transformation. Taxes are added as in the case imports above:

```
--- Price for the domestic origin, equation (45) in VDM 2018
e_pdp(rs,iIn(i),aa,ts) $ (p_alphad(rs,i,aa,ts) $ (not a(aa) or sum(sameas(aa,aIn),1))) ..
pdp(rs,i,aa,ts) =e m_pdpDef(rs,i,aa,ts);
```

Similar to above, the macro considers ad-valorem use taxes and per unit emission taxes. The relevant price is defined by the macro *m\_domPrice*:

```
$macro m_pdpDef(r,i,aa,t) ( m_domPrice(r,i,aa,t) * (1 + [dintx(r,i,aa,t) + itxshft(r,i,t)] $ (not ttmg(aa))) \
    --- add emission taxes for domestic use
    $if %ModulesCO2_Emissions==on + emis(r,t) * p_emidCo2eqTaxed(r,i,aa,t) * (1-redemiaCo2eq(r,aa,t) * p_nCo2ShareD(r,i,aa) $ a(aa) + p_emiShift(r,i,aa,t)) \
)
```

The macro *m\_domPrice* considers the case of using the Melitz extension or not. The case of the Melitz extension is discussed in the chapter on this extension. For the standard case, the relevant price is equal to the domestic price *pd* if the CET transformation for product *i* is not infinite and otherwise equal to the supply price *ps*:

```
$iftheni.mel "%melitz%"=="on"
    $macro m_domPrice(r,i,aa,t) [ ( pd(r,i,t) $ (p_omegax(r,i) ne inf) + ps(r,i,t) $ (p_omegax(r,i) eq inf)) $ (not iMel(i)) \
    + ( pd(r,i,t) $ p_melDomUse(r,aa,i) + m_pfirm(r,i,"domes",t) $ (not p_melDomUse(r,aa,i))) $ iMel(i)
$else.mel
    $macro m_domPrice(r,i,aa,t) ( pd(r,i,t) $ (p_omegax(r,i) ne inf) + ps(r,i,t) $ (p_omegax(r,i) eq inf))
$endif.mel
```



Note again that the Melitz extension will delete the *alphad* parameters for the products handled by imperfect competition to replace the equation *e\_pdp* as it uses a different pricing system.

The equations *e\_pdp* and *e\_pmp* are by default not comprised in the model and the prices are replaced instead by the macros shown above.

The **Armington prices** of the different agents *pa* are defined in *e\_pa*, equation (47) in VDM 2018. That dual price aggregator as usually reflects the given shares for domestic *p\_alphad* and imported *p\_alpham* origin and the related prices as defined above as well as the substitution elasticity between imports and domestic origin *p\_sigmam* which is not agent specific, but depends on the region and product:

```
*
* --- Armington price, equation (47) in VDM 2018
*
* Note WB: equation is only used if both domestic and import demand are
* present (check for share parameters), otherwise PA is substituted out with m_pa macro
*
e_pa(rs,iIn(i),aa,ts) $ ( (p_alphad(rs,i,aa,ts) or p_alpham(rs,i,aa,ts)) $ (not a(aa) or sum(sameas(aa,aIn),1))
*
* WB: If additional prices are substituted out (switched on in interface)
* the Armington price aggregator for all agents aa
* is calculated in the macro m_pa and the equation is not active
*
$iftheni.subsPrices "%subsArmPrices%"=="on"
$ (1 eq 2)
$endif.subsPrices
)..
pa(rs,i,aa,ts) =e= mm_paDef(rs,i,aa,ts);
```

As defined in the macro *mm\_padef*:

```
$macro mm_padef(r,i,aa,t) { ( [p_alphad(r,i,aa,t) * pdp(r,i,aa,t) ** (1-p_sigmam(r,i))] $ p_alphad(r,i,aa,t) \
+ [p_alpham(r,i,aa,t) * (pmp(r,i,aa,t)/lambdaxm(r,i,aa,t)) ** (1-p_sigmam(r,i))] $ p_alpham(r,i,aa,t) \
) ** (1/(1-p_sigmam(r,i))) }
```

Note here that the equation is normally substituted out. That is not the case if the standard GTAP layout is used or the substitution is explicitly switched off on the interface.

**Domestic demand** *xd* by the different agents is driven by the share parameter *p\_alphad* and times the total Armington demand defined in the macro *m\_xa*, times the price relation exponent the substitution elasticity. It is defined in the equation *e\_xd*, equation (48) in VDM 2018. The relevant price indices are defined by the macro *m\_pa* in case of the standard CES function and by *mm\_padef* in case of the additive variant:

```
$iftheni.hsubr "%hSubr%"=="on"
e_xd(rs,iIn(i),aaDM,ts) $ ( (p_alphad(rs,i,aaDM,ts) gt eps) $ (xd.range(rs,i,aaDM,ts) ne 0) ) ..
xd(rs,i,aaDM,ts) s(xd,'rs,i,aaDM,ts') =e= p_alphad(rs,i,aaDM,ts) * mm_xa(rs,i,aaDM,ts) s(xd,'rs,i,aaDM,ts')
* ((m_pa(rs,i,aaDM,ts) $ (not addCes(rs,i,"imp")) + mm_padef(rs,i,aaDM,ts) $ addCes(rs,i,"imp"))
/m_pdp(rs,i,aaDM,ts)) ** p_sigmam(rs,i);
e_xn(subr,iIn(i),aa,ts) $ (p_alphan(subr,i,aa,ts) gt eps) ..
xn(subr,i,aa,ts) s(xn,'subr,i,aa,ts') =e= p_alphan(subr,i,aa,ts) * mm_xa(subr,i,aa,ts) s(xn,'subr,i,aa,ts')
* ((m_pa(subr,i,aa,ts) $ (not addCes(subr,i,"imp")) + mm_padef(subr,i,aa,ts) $ addCes(subr,i,"imp"))
/sum(r_r(subr,discr),m_pdp(discr,i,aa,ts))) ** p_sigmam(subr,i);
$else.hsubr
e_xd(rsNat,iIn(i),aaDM,ts) $ ( (p_alphad(rsNat,i,aaDM,ts) gt eps) $ (xd.range(rsNat,i,aaDM,ts) ne 0) ) ..
xd(rsNat,i,aaDM,ts) s(xd,'rsNat,i,aaDM,ts') =e= p_alphad(rsNat,i,aaDM,ts) * mm_xa(rsNat,i,aaDM,ts) s(xd,'rsNat,i,aaDM,ts')
* ((m_pa(rsNat,i,aaDM,ts) $ (not addCes(rsNat,i,"imp")) + mm_padef(rsNat,i,aaDM,ts) $ addCes(rsNat,i,"imp"))
/m_pdp(rsNat,i,aaDM,ts)) ** p_sigmam(rsNat,i);
$endif.hsubr
```

Note that it is possible to define households at the level of sub-regional units (first block) where an additional demand *xn* for the national (as opposed to the regional) origin is introduced.



It is linearly aggregated over agents to **total domestic sales**  $xds$  in the equation  $e_{pd}$ , equation (67) in VDM 2018, potentially taking the demand for the national origin from sub-regional households into account:

```
*
* --- Domestic goods market equilibrium, equation (67) in VDM 2018
*
e_pd(rsNat,iIn(i),ts) $ xdFlag(rsNat,i) ..
    xds(rsNat,i,ts)/xds.scale(rsNat,i,ts)
    =e= sum(aa $ (m_alphad(rsNat,i,aa,ts) gt eps), xd(rsNat,i,aa,ts))/xds.scale(rsNat,i,ts)

    $$iftheni.hsubr %hSubr%="on"
        + sum( (rs,h) $ (r_r(rs,rsNat) $ (p_alphan(rs,i,h,ts) gt eps)),xn(rs,i,h,ts))/xds.scale(rsNat,i,ts)
    $$endif.hsubr
;
```

**Imported demand** by each agent  $xm$  is defined accordingly in the equation  $e_{xm}$ , equation (49) in VDM 2018, potentially considering the case of demand of household at sub-regional level:

```
*
* --- Import demand, equation (49) in VDM 2018
*
$$iftheni.hsubr "%hSubr%"="on"
e_xm(rs,iIn(i),aa,ts) $ ( (p_alphan(rs,i,aa,ts) gt eps) $ (xm.range(rs,i,aa,ts) ne 0) ) ..
    xm(rs,i,aa,ts)/xm.scale(rs,i,aa,ts) =e= p_alphan(rs,i,aa,ts)
    * mm_xa(rs,i,aa,ts)/xm.scale(rs,i,aa,ts)
    * (m_pa(rs,i,aa,ts)/sum(r_r(rs,rsNat),m_pmp(rsNat,i,aa,ts)))*p_sigam(rs,i)
    * lambdaXm(rs,i,aa,ts)**(p_sigam(rs,i)-1)
    ;
$$else.hsubr
e_xm(rsNat,iIn(i),aa,ts) $ ( (p_alphan(rsNat,i,aa,ts) gt eps) $ (xm.range(rsNat,i,aa,ts) ne 0) ) ..
    xm(rsNat,i,aa,ts)/xm.scale(rsNat,i,aa,ts) =e= p_alphan(rsNat,i,aa,ts)
    * mm_xa(rsNat,i,aa,ts)/xm.scale(rsNat,i,aa,ts)
    * (m_pa(rsNat,i,aa,ts)/m_pmp(rsNat,i,aa,ts))*p_sigam(rsNat,i)
    * lambdaXm(rsNat,i,aa,ts)**(p_sigam(rsNat,i)-1)
    ;
$$endif.hsubr
```

**Total import demand**  $xmt$  is defined by the equation  $e_{xmt}$  an adding up over the demand of the individual Armington agents, equation (50) in VDM 2018, potentially considering the case of demand of household at sub-regional level:

```
*
* --- Total import demand, equation (50) in VDM 2018
*
e_xmt(rsNat,iIn(i),ts) $ xmtFlag(rsNat,i) ..
    xmt(rsNat,i,ts)/xmt.scale(rsNat,i,ts) =e= sum(aa $ (p_alphan(rsNat,i,aa,ts) gt eps), xm(rsNat,i,aa,ts))/xmt.scale(rsNat,i,ts)
    $$ifi "%ModulesGTAP_NUTS2%"="ON" +sum( (subr,h) $ (r_r(subr,rsNat) $ (p_alphan(subr,i,h,ts) gt eps)),xm(subr,i,h,ts))/xmt.scale(rsNat,i,ts)
    ;
```

The bilateral **cost, insurance and freight prices**  $pmcif$  are defined by the equation  $e_{pmcif}$  and the macro  $m_{pmcif}$ . As the default, these prices are substituted out from the model. The set  $rrComb$  is used in case of only one region being solved to depict the bi-lateral trade links to include.

```
--- c.i.f. price, normally substituted out, macro defined above, equation (65) in VDM 2018
e_pmcif(rNat1,iIn(i),rNat,ts) $ (m_xwFlag(rNat1,i,rNat) $ rrComb(rNat1,rNat)) ..
    pmcif(rNat1,i,rNat,ts) =E= m_pmcif(rNat1,i,rNat,ts);
```

The macro  $m_{pmcif}$  is defined as follows, equation (65) in VDM 2018:

```
--- costs, insurance and freight price (c.i.f): add transport margins to f.o.b. price,
    equation (65) in VDM 2018, plus equation (60) substituted out
$macro m_pmcif(r,i,rp,t) { [ m_pecob(r,i,rp,t)/lcu(r,t) \
    + sum(m $ p_amgm(m,r,i,rp), p_amgm(m,r,i,rp)*ptmg(m,t)\
        /m_lambdamg(m,r,i,rp,t)) * tmarg(r,i,rp,t) \
    ]*lcu(rp,t) }
```

It converts the bilateral f.o.b. (free on board) price defined in the macro  $m\_pefob$  to international currency (division by  $lcu$ ) and adds the per unit transport margin cost in international currency. These costs are defined based on transport mode  $m$  (water, air, rest) specific shares  $p\_amgm$  on the given transport margin  $tmarg$ , updated with the mode specific average global price for that mode  $ptmg$ . The mode specific costs can be shifted by  $m\_lambdamg$ . The resulting prices – f.o.b plus transport margin – in international currency are finally converted into local currency again by the multiplication by  $lcu$ .

Note that with a dense bi-lateral trade matrix, the number of variables relating to bi-lateral trade relations increases quadratic in the number of regions and linear in the number of sectors. Under full density, using 50 regions and 50 sectors implies hence  $50 \times 50 \times 50 = 125,000$  non-zero elements for each variable defined bi-laterally. That explains why substitutions of the e.g. the f.o.b. and c.i.f. prices and the bi-lateral trade margins can dramatically reduce model size.

Bilateral **free on-board prices**  $pefob$  are defined by the macro  $mm\_pefob$ , equation (64) in VDM 2018. They reflect: bilateral export taxes  $exptx$  and a product specific export tax shifter  $etax$  levied on the relevant price which depends if and how a transformation of output is used as seen below. If no CET approach is used, the supply price  $ps$  is the basis for f.o.b calculation. In case there is only a CET between total exports and domestic sales, but none between bi-lateral export flows, the average price of exports  $pet$  is used, otherwise, bi-lateral export prices  $pe$  define the basis for f.o.b prices.

```
*
* --- free on board price (f.o.b.), equation (64), in VDM 2018
*
*$macro mm_pefob(r,i,rp,t) { \
*
*   --- export taxes
*
*   (1 + exptx(r,i,rp,t) + etax(r,i,t)) \
*
*   --- the following might substitutes out pe depending on the CETs
*   between domestic sales and export resp. between exports
*
*   * ( \
*       --- use bilateral import prices if a CET between export destination is active
*       pe(r,i,rp,t) $ psFlag(r,i,"pe") \
*       --- use the export prices if there is only a CET between total exports and domestic use
*       + pet(r,i,t) $ psFlag(r,i,"pet") \
*       --- otherwise, use supply price
*       + ps(r,i,t) $ psFlag(r,i,"ps")    )}
*
```

Similar to the case of c.i.f. prices, the fob prices  $pefob$  are defined from that macro in the equation  $e\_pefob$ . Again, the default case is that these equations are substituted out from the model.

```
*
* --- f.o.b. price, normally substituted out, macro defined above, equation (64) in VDM 2018
*
*$e pefob(rNat1,iIn(i),rNat,ts) $ (m_xwFlag(rNat1,i,rNat) $ rrComb(rNat1,rNat)) ..
*
* pefob(rNat1,i,rNat,ts) =E= m_pefob(rNat1,i,rNat,ts);
```

The user can define on the interface if these prices are substituted from the model.

The **average price of imports**  $pmt$  is defined in the equation  $e\_pmt$ , equation (52) in VDM 2018, and considers three cases. The first case applies for substitution elasticity different from unity and for shares not considered small. It uses the standard dual price aggregator using the bilateral demand share parameters  $p\_amw$ , the cif price defined by  $\%pmcif\%$  plus bilateral import taxes  $imptx$  plus a product specific import tax shifter  $mtax$  and reflects a preference shifter defined in the macro  $m\_lambdam$ . The second case uses the dual price aggregator for the CD case where the substitution elasticity is unity. The third case reflects small shares which are treated a la Leontief.

```

pmt(rNat,iIn(i),ts) $ ( (pmt.range(rNat,i,ts) ne 0) $ (not p_pmtElas(rNat,i)) $ xmtFlag(rNat,i) $ (not iMrio(i)) $ (not iSpe(i)) ) ..

[ pmt(rNat,i,ts) $ addCes(rnat,i,"imp")
+ pmt(rNat,i,ts) $ (not addCes(rnat,i,"imp")) ]s(pmt,'rNat,i,ts') =e=

--- CES case with p_sigmaw <> 1

[sum(rNat1 $ (xwFlag(rNat1,i,rNat)
$if1 "%hasSmallShare%"=="on" and (not isSmallImpShare(rNat1,i,rNat))
),
p_amw(rNat1,i,rNat,ts)
*((1 + imptx(rNat1,i,rNat,ts) + mtax(rNat,i,ts)) * %pmcif%(rNat1,i,rNat,ts) ]
/m_lambdam(rNat1,i,rNat,ts))**(1 $ (not addCes(rnat,i,"imp")) -p_sigmaw(rNat,i))
)] $ ((p_sigmaw(rNat,i) ne 1) or addCes(rnat,i,"imp")) $ p_axPmtCD(rNat,i,ts)s(pmt,'rNat,i,ts')
)

--- CD case with p_sigmaw == 1

+ [prod(rNat1 $ (xwFlag(rNat1,i,rNat)
$if1 "%hasSmallShare%"=="on" and (not isSmallImpShare(rNat1,i,rNat))
),
((1 + imptx(rNat1,i,rNat,ts) + mtax(rNat,i,ts)) * %pmcif%(rNat1,i,rNat,ts)s(pmt,'rNat,i,ts') ]
/m_lambdam(rNat1,i,rNat,ts)) * p_amw(rNat1,i,rNat,ts)
) * p_axPmtCD(rNat,i,ts)s(pmt,'rNat,i,ts')
)] $ ((p_sigmaw(rNat,i) eq 1) and (not addCes(rnat,i,"imp"))) $ p_axPmtCD(rNat,i,ts)

$iftheni.smallShare "%hasSmallShare%"=="on"
+ sum(rNat1 $ isSmallImpShare(rNat1,i,rNat), p_amw(rNat1,i,rNat,ts)
*((1 + imptx(rNat1,i,rNat,ts) + mtax(rNat,i,ts))
* %pmcif%(rNat1,i,rNat,ts)s(pmt,'rNat,i,ts') ]/m_lambdam(rNat1,i,rNat,ts)))
$endif.smallShare
;

```

The equation also considers the case of an additive CES function for imports. In this case, an additional equation is needed to define the exhaustion price *pmt* that differs from the dual price aggregator *pmtd* defined by the equation above:

```

e_pmt(rNat,iIn(i),ts) $ ( (pmt.range(rNat,i,ts) ne 0) $ (not p_pmtElas(rNat,i)) $ xmtFlag(rNat,i) $ (not iMrio(i)) $ (not iSpe(i)) $ addCes(rnat,i,"imp") ) ..

pmt(rnat,i,ts)s(pmt,'rNat,i,ts')
=E= sum(rnat1 $ p_amw(rNat1,i,rNat,ts), p_amw(rNat1,i,rNat,ts)s(pmt,'rNat,i,ts')
* ((1 + imptx(rNat1,i,rNat,ts) + mtax(rNat,i,ts)) * %pmcif%(rNat1,i,rNat,ts)) ** (1-p_sigmaw(rNat,i))
* m_lambdam(rNat1,i,rNat,ts) ** (p_sigmaw(rNat,i) - 1) ) * pmtd(rNat,i,ts) ** p_sigmaw(rNat,i);

```

Note that the single country case can either use the lower-level Armington / CET equations of the trading partner of the country solved, or use import and export elasticities. To host the latter case, the *e\_pmt* equation will not be introduced in the model if import prices are elasticity driven.

The allocation of total imports *xmt* to the **bi-lateral imports** *xw* is defined in the equation *e\_xw*, equation (52) in VDM, and is based on the share parameters *p\_amw*, the substitution between origins *p\_sigmaw* and the relevant price relation, i.e. the average import price *pmt* divided by the c.i.f. price *%pmcif%* plus bi-lateral import taxes *imptx* plus a potential import tax shifter *mtax*. Preference shifters as defined in the macro *m\_lambdam* can be used as well.

```

--- Bilateral import demand, equation (51) in VDM 2018
Note: the price PM in VDM is always substituted out as pmCif * (1+imptx+mtax), replaces equation (66)

e_xw(rNat1,iIn(i),rNat,ts) $ (xwFlag(rNat1,i,rNat) $ (not iMrio(i)) $ (not iSpe(i)) $ (not armComProCur(i)) $ rrComb(rNat1,rNat)) ..

xw(rNat1,i,rNat,ts)s(xw,'rNat1,i,rnat,ts') =e= p_amw(rNat1,i,rNat,ts)*xmt(rNat,i,ts)s(xw,'rNat1,i,rnat,ts')

* [ (( pmtd(rNat,i,ts) $ addCes(rnat,i,"imp") + pmt(rnat,i,ts) $ (not addCes(rnat,i,"imp")) ) / ((1 + imptx(rNat1,i,rNat,ts) + mtax(rNat,i,ts))

--- Note WB: cif price is defined via macro %pmcif%, normally not a variable in model

* %pmcif%(rNat1,i,rNat,ts) ] ** p_sigmaw(rNat,i)
---- correct shifter additive CES (!)
* m_lambdam(rNat1,i,rNat,ts) ** (p_sigmaw(rNat,i) - 1) )

$iftheni.smallShare "%hasSmallShare%"=="on"
$ (not isSmallImpShare(rNat1,i,rNat))
+ m_lambdam(rNat1,i,rNat,ts) $ isSmallImpShare(rNat1,i,rNat)
$endif.smallShare
;

```

Note that in opposite to VDM 2018, the import price defined in equation (66) is always substituted out. Equally, the equation can consider the case of an additive CES function.

That macro *m\_lambdam* is introduced to avoid that for each bi-lateral trade link in the model, a variable must be fixed to unity if no shifter is present. It thereof introduces unity in the equation – see first line below – if the shifter is not initialized, i.e. equal to zero and not fixed. The shifter variable is used if it either fixed, i.e. the range is zero, or its starting value is not zero.

```
*
* --- same, for bi-lateral Armington preferences shifters
*
* $macro m_lambda(r,i,rp,t) { [ 1 $ ( (lambda.range(r,i,rp,t) ne 0) and (lambda.l(r,i,rp,t) eq 0) ) \
+ lambda(r,i,rp,t) $ ( (lambda.range(r,i,rp,t) eq 0) or (lambda.l(r,i,rp,t) ne 0) ) ] $ card(lambda)\
+ 1 $ (card(lambda) eq 0) }
```

Similar macros are used for other shifter variables as well.

As the model allows non-infinite transformation of outputs, the following equation  $e\_xds$  (implicitly), equation (53) in VDM 2018, defines **domestic sales**  $xds$ . The first case is that of infinite transformation as found in the GTAP standard model where by definition the price of domestic sales  $pd$  is equal to the average supply price  $ps$ . The second case distributes total supply  $xs$  of a product to domestic sales based on the share parameter  $p\_gd$  and the relation between the domestic sales price  $pd$  and average supply price  $ps$  times exponent the transformation elasticity  $p\_omegax$ :

```
-
e_xds(rsNat,iIn(i),ts) $ (xdFlag(rsNat,i) $ (p_omegax(rsNat,i) ne inf) ) ..
0 =e=
*
* --- infinite CET <=> prices are equal, GTAP standard
*
* (pd(rsNat,i,ts) - ps(rsNat,i,ts)) $ (p_omegax(rsNat,i) eq inf)
*
* --- CET mechanism otherwise
*
+ (xds(rsNat,i,ts)/xds.scale(rsNat,i,ts)
- p_gd(rsNat,i,ts)*xs(rsNat,i,ts)/xds.scale(rsNat,i,ts)
* (pd(rsNat,i,ts)/(ps(rsNat,i,ts) $ (not iACET(i)) + pds(rsNat,i,ts) $ iACET(i)))**p_omegax(rsNat,i) $ (p_omegax(rsNat,i) ne inf);
-
```

Note that the equation allows for the use of the additive function form. As in other cases discussed above, this form implies that the exhaustion price  $ps$  differs from the dual price aggregator  $pds$ .

A similar equation defines **total exports**  $xet$  in the equation  $e\_xet$ , equation (54) in VDM 2018. The relevant share parameter is depicted by  $p\_ge$  while the average price of exports is called  $pet$ :

```
*
* --- Total export demand, equation (54) in VDM 2018
*
e_xet(rsNat,iIn(i),ts) $ (xetFlag(rsNat,i) $ (xet.range(rsNat,i,ts) ne 0) ) ..
xet(rsNat,i,ts)/xet.scale(rsNat,i,ts)
=e=
*
* --- case with CET
*
* ((p_ge(rsNat,i,ts)*xs(rsNat,i,ts))/xet.scale(rsNat,i,ts)
* * (pet(rsNat,i,ts)/(ps(rsNat,i,ts) $ (not iACET(i)) + pds(rsNat,i,ts) $ iACET(i)))**p_omegax(rsNat,i) $ (p_omegax(rsNat,i) ne inf)
*
* --- infinite CET: physical aggregation (global model) or elasticity driven (single country option)
*
+ ( sum( rpNat $ m_xwFlag(rsNat,i,rpNat), m_xws(rsNat,i,rpNat,ts)/xet.scale(rsNat,i,ts) $ (not p_petElas(rsNat,i))
+ (p_axPet(rsNat,i,ts)/xet.scale(rsNat,i,ts) * pet(rsNat,i,ts)**p_petElas(rsNat,i) $ p_petElas(rsNat,i)
) $ (p_omegax(rsNat,i) eq inf);
```

Note the special case for the single country model where exports are driven by an export elasticity.

The average **supply price**  $ps$  is defined in the equation  $e\_xs$ , equation (55) in VDM 2018. In case of infinite transformation and thus a linear aggregator – the first block – the sum of domestic sales  $xds$  and exports  $xet$  must be equal to physical output  $xs$ . In case of not-infinite transformation, a dual price aggregator is used based on the share parameters  $p\_gd$  and  $p\_ge$ , composites price  $ps$  respectively  $pds$ , the price of exports  $pet$  and the transformation elasticity  $p\_omegax$ :

```

--- Market clearing for domestic supply, equation (55) in VDM 2018

$xs(rsNat,iIn(i),ts) $ (xsFlag(rsNat,i) $ (ps.range(rsNat,i,ts) ne 0))..
0 =e=

--- infinite CET <=> physical balancing

+ (xs(rsNat,i,ts)s(xs,'rsNat,i,ts')
  $if1 "%modulesGTAP_NUTS2%"=="on" + sum((r_r(subr,rsNat),aaDm), p_subrDem(subr,i,aaDm,ts))s(xs,'rsNat,i,ts')
    - (xds(rsNat,i,ts)s(xs,'rsNat,i,ts') ) $ xdfFlag(rsNat,i)
    - (xet(rsNat,i,ts)s(xs,'rsNat,i,ts') ) $ xetFlag(rsNat,i) ) $ (p_omegax(rsNat,i) eq inf)

--- dual price aggregator

+ ( ps(rsNat,i,ts) $ (not iACET(i))
  + pds(rsNat,i,ts) $ iACET(i)
    - ( p_gd(rsNat,i,ts)*pd(rsNat,i,ts)**(1 $ (not iACET(i)) + p_omegax(rsNat,i))
      + p_ge(rsNat,i,ts)*pet(rsNat,i,ts)**(1 $ (not iACET(i)) + p_omegax(rsNat,i))
      )*(1/(1 $ (not iACET(i)) + p_omegax(rsNat,i)) ) ) $ (p_omegax(rsNat,i) ne inf)
;

```

**Bilateral export supply** is by definition equal to bilateral export demand  $xw$ , that equality is used to define indirectly the bilateral export price  $pe$  in case of non-infinite transformation in the first block, as defined in the equation  $e_{pe}$ , equation (56) in VDM 2018. Otherwise, bilateral export prices  $pe$  and the average export prices  $pet$  are by definition equal:

```

--- This equation substitutes out the equilibrium condition for bilateral trade, equation (56) in VDM 2018
(makes it possible to drop equation (68) in VDM 2018)

e_pe(rNat,iIn(i),rNat1,ts) $ (m_xwFlag(rNat,i,rNat1) $ psFlag(rNat,i,"pe") $ rrComb(rNat,rNat1)) ..

0 =e= ( m_xws(rNat,i,rNat1,ts)s(xw,'rNat,i,rNat1,ts')
  - (p_gw(rNat,i,rNat1,ts)*xet(rNat,i,ts))s(xw,'rNat,i,rNat1,ts')
    * [ 1
      + { (pe(rNat,i,rNat1,ts)/(pdet(rNat,i,ts) $ iACET(i) + pet(rNat,i,ts) $ (not iACET(i))) )**p_omegaw(rNat,i)
        } $ (not isSmallExpShare(rNat,i,rNat1)) ]
    ) $ (p_omegaw(rNat,i) ne inf)

--- case with law of one price (infinite transformation)

+ ( pe(rNat,i,rNat1,ts) - (pet(rNat,i,ts) $ (not iACET(i)) + pdet(rNat,i,ts) $ iACET(i)) $ (p_omegax(rNat,i) ne inf)
  - ps(rNat,i,ts) $ (p_omegax(rNat,i) eq inf) )
  $ (p_omegaw(rNat,i) eq inf) ;

```

The set `iACET` depicts cases where a volume preserving CET is used.

The **aggregate price of export**  $pet$  defined in the equation  $e_{pet}$ , equation (57) in VDM 2018, is either defined from a dual price aggregator in case of non-infinite transformation or equal to the supply price  $ps$  in case of infinite transformation. Note the inclusion of the special case of small export shares handled via Leontief:

```

*
* --- Aggregate price for exports, equation (57) in VDM 2018
*
e_pet(rNat,iIn(i),ts) $ (xetFlag(rNat,i) $ sum(rsNat,rrComb(rNat,rsNat)) $ (pet.range(rNat,i,ts) ne 0)
  $ (not psFlag(rNat,i,"ps"))) ..

0 =e= (
  pdet(rNat,i,ts) $ iACET(i)
  + pet(rNat,i,ts) $ (not iACET(i))
  - p_axPet(rNat,i,ts)
    * sum(rNat1 $ (m_xwFlag(rNat,i,rNat1) $ (not isSmallExpShare(rNat,i,rNat1))),
      p_gw(rNat,i,rNat1,ts) * pe(rNat,i,rNat1,ts)**(1 $ (not iACET(i)) + p_omegaw(rNat,i))
      )*(1/(1 $ (not iACET(i)) + p_omegaw(rNat,i)))

  $ifthen1.smallShare "%hasSmallShare%"=="on"

*
* --- Leontief treatment of small export shares
*
  + sum(rNat1 $ (m_xwFlag(rNat,i,rNat1) $ isSmallExpShare(rNat,i,rNat1)),
    p_gw(rNat,i,rNat1,ts)*pe(rNat,i,rNat1,ts)) $ card(isSmallExpShare)
  ) $ psFlag(rNat,i,"pe")

+ (xet(rNat,i,ts)/xet.scale(rNat,i,ts)
  - sum( rpNat $ m_xwFlag(rNat,i,rpNat), m_xws(rNat,i,rpNat,ts)/xet.scale(rNat,i,ts)) $ (not p_petElas(rNat,i))
    + (p_axPet(rNat,i,ts)/xet.scale(rNat,i,ts) * pet(rNat,i,ts)**p_petElas(rNat,i)) $ p_petElas(rNat,i)
    ) $ psFlag(rNat,i,"pet");
;

```

In case of an additive ACET, the exhaustion prices is implicitly defined:

```

p_det(rNat,iIn(i),ts) $ (xetFlag(rNat,i) $ sum(rsNat,rrComb(rNat,rsNat)) $ (p_det.range(rNat,i,ts) ne 0) $ iACET(i)) ..
0 =e= sum(rNat1 $ m_xwFlag(rNat,i,rNat1),
          m_xws(rNat,i,rNat1,ts)*(pe(rNat,i,rNat1,ts)-pet(rnat,i,ts)));

```

The **global demand for transport services** *xtmg* of mode *m* is based on a Leontief approach and defined in the equation *e\_xtmg*, summarizing equations (58), (59) and (61) in VDM 2018. It assumes that a global transport service agent buys transport services demands by mode (air, water, rest) from the different model region such as minimize global transport cost at given global demand for international transport services. The given bi-lateral transport margin demand *tmarg* are distributed to the different transport modes *m* based on the share parameter *p\_amgm* multiplied with the bilateral transport flows defined in the macro *m\_xws*, reflecting a potential demand shifter *m\_lambdamg*. Note that substitution is between regions providing shares on international transport by transport mode (see equation *e\_axtmg* below), and not between different modes:

```

--- the following equation summarizes equations (58),(59),(61) in VDM 2018
    and thus save a considerable number of variables and equations in the overall model

e_xtmg(m,ts) $ (xtmg.range(m,ts) ne 0) ..

xtmg(m,ts)s(xtmg,'m,ts')
=e= sum((rNat,i,rNat1) $ p_amgm(m,rNat,i,rNat1),
        ---- equation (58) - bilateral import flows
        times per unit demand for transport services
        p_amgm(m,rNat,i,rNat1)
        * ( m_xws(rNat,i,rNat1,ts) $ rrComb(rNat,rNat1)
            $$onexpand
            + m_xws1(rNat,i,rNat1,ts) $ (not rrComb(rNat,rNat1))
            $$offexpand
        )
        --- equation (59) multiplied with trade margins
        * tmarg(rNat,i,rNat1,ts)
        /m_lambdamg(m,rNat,i,rNat1,ts)
    )s(xtmg,'m,ts') ;

```

The **region-specific demand for each transport mode** *xa*, defined in equation *e\_xatmg*, equation (62) in VDM 2018, is based on a CES demand system which reflects the average global price for each transport mode *ptgm*, the regional specific price defined in the macro *m\_pa* and the substitution elasticity *p\_sigmamg*. Take sea transport as an example. Shipping one unit of a certain good *i* from *rNat* and *rNat1* requires the amount of transport services *tmarg* of which a certain share *p\_amgm* are sea transport services. Adding these amounts up in equation *e\_xtmg* above by considering the current trade flows *m\_xw* defines the global demand for sea transport services. The equation *e\_xatmg* below defines how much each model region contributes to these global services.

```

--- Armington demand for transport services by mode, distributes global demand
    xtmg to region, equation (62) in VDM 2018

e_xatmg(rsNat,m,tmg,ts) $ (xaiFlag(rsNat,m,tmg) ) ..

xa(rsNat,m,tmg,ts)s(xa,'rsNat,m,tmg,ts')
=e= p_alphaa(rsNat,m,tmg,"top",ts)*xtmg(m,ts)s(xa,'rsNat,m,tmg,ts')
    * ( ptmg(m,ts)/(m_pa(rsNat,m,tmg,ts)/lcu(rsNat,ts)))**p_sigmamg(m) ;

```

The global average price for each transport mode *ptgm* is defined in the equation *e\_ptm*, equation (63) in VDM 2018, via a dual price aggregator which distinguishes the CD and CES/Leontief case:



```

*
* --- Average global price of transport mode m, equation (63) in VDM 2018
*
e_ptmg(m,ts) $ (ptmg.range(m,ts) ne 0) ..
    ptmg(m,ts) =e=
*
*   --- case with elasticity <> 1 (CES)
*
    ( (sum( (rNat,tmg) $ alphaa.1(rNat,m,tmg,ts),
        alphaa(rNat,m,tmg,ts)
        * ( (m_pa(rNat,m,tmg,ts)/lcu(rNat,ts)) $ rsNat(rNat)
          + (pa.1(rNat,m,tmg,ts)/lcu.1(rNat,ts)) $ (not rsNat(rNat)))** (1-p_sigmamg(m)) )** (1/(1-p_sigmamg(m))) )/p_axmg(m,ts)
    ) $ (p_sigmamg(m) ne 1)
*
*   --- Cobb-double case with elasticity == 1
*
+ ( sum(tmg, prod(rNat $ alphaa.1(rNat,m,tmg,ts),
    (( (m_pa(rNat,m,tmg,ts)/lcu(rNat,ts)) $ rsNat(rNat)
    + (pa.1(rNat,m,tmg,ts)/lcu.1(rNat,ts)) $ (not rsNat(rNat)))
    /alphaa(rNat,m,tmg,ts))** (alphaa(rNat,m,tmg,ts))) )/p_axmg(m,ts)
) $ (p_sigmamg(m) eq 1)
;

```

### 2.13.3 Melitz model

The methodological details of the Melitz extension are described in a separate document. In here, only the equations are briefly discussed (see *GtapMelitz\GtapMelitz\_model.gms*).

The Armington price of each agent for a product *iMel* included in the Melitz model as shown in equation *e\_pamel* below uses the usual CES dual price aggregator, however, the share parameters *p\_amel* are updated which changes of the number of firms operating on the trade link to yield a different type of price aggregator. The number of firms is defined via the macro *m\_nFirmOp*, while the numbers of firms operating under the benchmark is stored in the parameter *p\_nFormsOp0*:

```

*
* e_pamel(rsNat(rpNat),iMel(i),aa,ts(t)) $ ((m_alpham(rpNat,i,aa,t) or m_alphad(rpNat,i,aa,t))
*   $ (sum(sameas(aIn,aa),1) or (not a(aa)) ) $ iIn(i)) ..
*
* rpNat is the import destination (= Armington agents)
* the rNat1 are the exporters delivering into rpNat, "domes" are domestic sales
*
pa(rpNat,i,aa,t) =E=
    [sum(rNat1 $ p_amel(rNat1,iMel,aa,rpNat,t),
*
*   --- exporters
*
    p_amel(rNat1,i,aa,rpNat,t)*m_nFirmsOp(rNat1,i,rpNat,t)/p_nFirmsOp0(rNat1,i,rpNat,t)
    *m_pAgentm(rNat1,i,aa,rpNat,t)** (1-p_sigmaMel(rpNat,i,rpNat,t))
*
*   --- domestic sales subject to price markups, takes domestic use taxes incl. carbon taxes into account
*
    +[p_amel(rpNat,i,aa,"domes",t)*m_nFirmsOp(rpNat,i,"domes",t)/p_nFirmsOp0(rpNat,i,"domes",t)*(1-p_melDomUse(rpNat,aa,i))
    * (m_pFirm(rpNat,i,"domes",t)*(1 + dintx(rpNat,i,aa,t) + itxshft(rpNat,i,t))
    * $if1 %ModulesCO2_Emissions%==on + emisrp(rpNat,t) * p_emid(rpNat,i,aa,t)
    )** (1-p_sigmaMel(rpNat,i,rpNat,t))] $ (p_amel(rpNat,i,aa,"domes",t) $ p_nFirmsOp0(rpNat,i,"domes",t))
*
*   --- domestic sales not subject to price markups (use of own output by firms),
*   takes domestic use taxes incl. carbon taxes into account
*
    +[p_amel(rpNat,i,aa,"domes",t)*p_melDomUse(rpNat,aa,i)
    * (ps(rpNat,i,t)*(1 + dintx(rpNat,i,aa,t) + itxshft(rpNat,i,t))
    * $if1 %ModulesCO2_Emissions%==on + emisrp(rpNat,t) * p_emid(rpNat,i,aa,t)
    )** (1-p_sigmaMel(rpNat,i,rpNat,t))] $ (p_amel(rpNat,i,aa,"domes",t) $ p_melDomUse(rpNat,aa,i))
*
    ]** (1/(1-p_sigmaMel(rpNat,i,rpNat,t)));
*

```

The Melitz extension allows to treat domestic sales of intermediates to same industry, i.e. the diagonal I/O element, different from the other agents by removing the love of variety effect and thus also the price markup. That has shown to help in many cases where the solver ran into infeasibilities. The share of such demand is depicted by the parameter *p\_melDomUse*. Its base price is the output price *ps* (variables per unit cost plus production taxes) and not the margin and tax inclusive price.

The relevant prices for the agents are depicted in the two macros *m\_pAgentm* for bi-lateral import links and *m\_pAgentd* for the domestic sales case.

The bi-lateral import case uses the c.i.f. price plus import and emission taxes:

```

$$macro m_pAgentm(r,i,aa,rrl,t) \
    [ %pmcif(r,i,rrl,t)*(1 + imptx(r,i,rrl,t) + mtax(rrl,i,t)) \
      * (1 + mintx(rrl,i,aa,t) + itxshft(rrl,i,t)) \
      $$ifi %ModulesCO2_Emissions==on + emisp(rrl,t) * p_emii(rrl,i,aa,t) \
    ]

$$macro m_pAgentd(r,i,aa,t) \
    [ ( m_pFirm(r,i,"domes",t) * (1 - p_melDomUse(r,aa,i)) * m_nFirmsOp(r,i,"Domes",t)/p_nFirmsOp0(r,i,"domes",t) ) $ p_nFirmsOp0(r,i,"domes",t) \
      + ps(r,i,t)*p_melDomUse(r,aa,i)) \
      * (1 + dintx(r,i,aa,t) + itxshft(r,i,t)) \
      $$ifi %ModulesCO2_Emissions==on + emisp(r,t) * p_emid(r,i,aa,t) \
    ]

```

Note that the basis for the f.o.b. price from which the c.i.f. price is derived are the prices charged by the average firm on the trade link defined by the macro *m\_pFirm* for Melitz products *iMel(i)* as shown in the macro below:

```

$$macro m_pegob(r,i,rp,t) [ m_pFirm(r,i,rp,t) * (1 + exptx(r,i,rp,t) + etax(r,i,t)) $ iMel(i) \
    + mm_pegob(r,i,rp,t) $ (not iMel(i))]

$$macro m_pFirm(r,i,rp,t) pFirm(r,i,rp,t)

```

The domestic demand by agent is defined in the *e\_xdMel* equation. As in the standard Armington model, it is a share *p\_amel* of total Armington demand *xa* updated with a price relation exponent *a* substitution elasticity. However, the share parameter is updated by

```

e_xdMel(rs(r),iMel(i),aa,ts(t)) $ ( (m_alphad(r,i,aa,t) gt eps)
    $ (sum(sameas(aIn,aa),1) or (not a(aa)) ) $ iIn(i)) ..

xd(r,i,aa,t)s(xd,'r,i,aa,t') =E=
[ p_amel(r,i,aa,"domes",t)
* --- share parameter is updated with change in firms operating on trade link
* m_nFirmsOp(r,i,"domes",t)/p_nFirmsOp0(r,i,"domes",t) * (1 - p_melDomUse(r,aa,i))

* ( xa(r,i,aa,t) $ ( (m_alphad(r,i,aa,t) gt eps) + (m_alpham(r,i,aa,t) gt eps))
    $$ifi "%hSubr%"=="on" + xa(r,i,aa,t) $ sum(r_r(subr,r) $ (xaFlag(subr,i,aa) $ sum(sameas(h,aa),1)),1)
)

    * (pa(r,i,aa,t)/
      (m_pFirm(r,i,"domes",t)*(1 + dintx(r,i,aa,t) + itxshft(r,i,t))
      $$ifi %ModulesCO2_Emissions==on + emisp(r,t) * p_emid(r,i,aa,t)
      ))*p_sigmaMel(r,i,r) $ p_nFirmsOp0(r,i,"domes",t)

* --- domestic demand not subject to price markups (demand of firms for on output)
* (see next equations)
*
+ xdmFlag(r,i,aa,t) $ p_melDomUse(r,aa,i)
]s(xd,'r,i,aa,t');

```

The value of total import demand *xm* by agent is defined residually as the differences between the value of its total Armington demand and the value of domestic demand:

```

* --- recalculate total import demand of agents xm from exhaustion at given average import prices pmt
* RHS is value of total Armington demand minus values of domestic demand
*
xmMelVal(rsNat(rpNat),iMel(i),aa,ts(t)) $ (m_alpham(rpNat,i,aa,t) $ (sum(sameas(aIn,aa),1) or (not a(aa)) ) $ iIn(i)) ..
xmMelVal(rpNat,i,aa,t)s(xmMelVal,'rpNat,i,aa,t')

=E= [ ( xa(rpNat,i,aa,t) $ ( (m_alphad(rpNat,i,aa,t) gt eps) + (m_alpham(rpNat,i,aa,t) gt eps))
    $$ifi "%hSubr%"=="on" + xa(rpNat,i,aa,t) $ sum(r_r(subr,rpNat) $ (xaFlag(subr,i,aa) $ sum(sameas(h,aa),1)),1)
    )*pa(rpNat,i,aa,t)
    - (xd(rpNat,i,aa,t)*m_pdp(rpNat,i,aa,t)) $ m_alphad(rpNat,i,aa,t) ]s(xmMelVal,'rpNat,i,aa,t');

```

The import demand by agents *xm* is defined from spendings on imports *xmMelVal* and the price of imports in the equation *e\_xmMel*:

```

e_xmMel(rsNat(rpNat),iMel(i),aa,ts(t)) $ (m_alpham(rpNat,i,aa,t) $ (sum(sameas(aIn,aa),1) or (not a(aa)) ) $ iIn(i)) ..
xmMelVal(rpNat,i,aa,t)s(xmMelVal,'rpNat,i,aa,t') =E=
xm(rpNat,i,aa,t)s(xmMelVal,'rpNat,i,aa,t') * pmt(rpNat,i,t) * (1 + mintx(rpNat,i,aa,t) + itxshft(rpNat,i,t));

```

The price index for imports is defined via the usual dual CES price index definition considering that the share parameters are updated with the number of firms being active on the trade link:



```

*
* --- Average import price, take # of firms operating on trade link into account
* macro pmCif considers the Melitz prices
*
e_pmtMel(rsNat(rNat),iMel(i),ts(t)) $ (sum(rNat1,xwFlagMel(rNat1,i,rNat)) $ iIn(i)) ..
*
pmt(rNat,i,t)

=E= sum(rNat1 $ xwFlagMel(rNat1,i,rNat),p_amew(rNat1,i,rNat,t)
      * m_nFirmsOp(rNat1,i,rNat,t)/p_nFirmsOp0(rNat1,i,rNat,t)
      * [%pmCif$(rNat1,i,rNat,t)*(1 + imptx(rNat1,i,rNat,t) + mtax(rNat,i,t))]**(1-p_sigmax(rNat,i))
      )**(1/(1-p_sigmax(rNat,i)));

```

Total imports *xmt* are defined as in the standard model by adding up over agents:

```

--- import aggregate (as in standard model, physical aggregation)

e_xmtMel(rsNat(rNat),iMel(i),ts(t)) $ (sum(aa, m_alpham(rNat,i,aa,t)) $ iIn(i)) ..

xmt(rNat,i,t)s(xmt,'rNat,i,t') =E= sum(aa $ m_alpham(rNat,i,aa,t), xm(rNat,i,aa,t))s(xmt,'rNat,i,t');

*
* --- import aggregate (as in standard model, physical aggregation)
*
e_xmtMel(rsNat(rNat),iMel(i),ts(t)) $ (sum(aa, m_alpham(rNat,i,aa,t)) $ iIn(i)) ..

xmt(rNat,i,t)s(xmt,'rNat,i,t') =E= sum(aa $ m_alpham(rNat,i,aa,t), xm(rNat,i,aa,t))s(xmt,'rNat,i,t');

```

As mentioned already above, the Melitz extension allows to treat domestic sales of intermediates to same industry, i.e. the diagonal I/O element, different from the other agents by removing the love of variety effect and thus also the price markup. That has shown to help in many cases where the solver ran into infeasibilities. In that case, the following simpler equation *e\_xddMel* defines the domestic demand by the same industry for its own output *xdMFlag*:

```

*
* --- domestic sales of intermediates to same industry (=diagonal I/O element),
* does not feature love of variety effect and hence no price markup
*
e_xddMel(rsNat(rpNat),iMel(i),aa,ts(t)) $ (m_alphad(rpNat,i,aa,t) $ p_melDomUse(rpNat,aa,i) $ iIn(i)) ..

xdMFlag(rpNat,i,aa,t)s(xd,'rpNat,i,aa,t') =E= p_amel(rpNat,i,aa,"domes",t)*xa(rpNat,i,aa,t)*p_melDomUse(rpNat,aa,i)
      * (pa(rpNat,i,aa,t)/
      (ps(rpNat,i,t)*(1 + dintx(rpNat,i,aa,t) + itxshft(rpNat,i,t))
      $ $if %ModulesCO2_Emissions%==on + emisrp(rpNat,t) * p_emid(rpNat,i,aa,t)
      ))**p_sigmaMel(rpNat,i,rpNat)s(xd,'rpNat,i,aa,t');

```

The total demand for domestic produce *xds* subject to price markups is defined by adding up over the domestic demands in the *e\_xdMelf* equation by adding the domestic demand quantities by each agent *xd*. The demand for own produce not subject to price margins *xdMFlag* is subtracted.

```

*
* --- domestic sales, corrected by domestic demand of own output by industry
*
e_xdMelf(rsNat(rpNat),iMel(i),"domes",ts(t)) $ (sum(aa, p_amel(rpNat,i,aa,"domes",t)) $ p_nFirmsOp0(rpNat,i,"domes",t) $ iIn(i)) ..

xds(rpNat,i,t)s(xds,'rpNat,i,t')
=E= [sum(aa $ p_amel(rpNat,i,aa,"domes",t), xd(rpNat,i,aa,t)-xdMFlag(rpNat,i,aa,t) $ p_melDomUse(rpNat,aa,i)
]s(xds,'rpNat,i,t') ;

```

The total bi-lateral demand is defined accordingly as an adding up over the agents' bi-lateral demands in the equation *e\_xwmelf*. In the Melitz case, it defines the average output of the firms being active on that link. Alternatively, the equation structure can be used to use one Armington nest, only, which allows quantifying the impact on model solution compared to the usual two-stage Armington system. In that case, the equation drives the bi-lateral demands *xw*.

```

*
* --- bilateral export demand, including exports to the same block
*
* e_xwMel(rNat,iMel(i),rNat1,ts(t)) $ (xwFlagMel(rNat,i,rNat1) $ rrComb(rNat,rNat1) $ iIn(i)) ..
*
*   $$iftheni.dem not "%HETSetup%"=="Only aggregated demand (no heterogenous firms)"
*   qAvFirm(rNat,i,rNat1,t)s(xw,'rNat,i,rNat1,t')
*   $$else.dem
*   xw(rNat,i,rNat1,t)s(xw,'rNat,i,rNat1,t')
*   $$endif.dem
*
*   =E=
*
*   --- armington duals: distribute total quantity according to price relations
*   (prices as in dual price aggregator above), added up over agents
*
*   sum(aa $ p_amel(rNat,i,aa,rNat1,t), p_amel(rNat,i,aa,rNat1,t)/p_nFirmsOp0(rNat,i,rNat1,t)
*   * ( ( xa(rNat1,i,aa,t) $ ( (m_alphad(rNat1,i,aa,t) gt eps) + (m_alpham(rNat1,i,aa,t) gt eps))
*   $$ifi "%hSubr%"=="on" + xa(rNat1,i,aa,t) $ sum(r_r(subr,rNat1) $ (xaFlag(subr,i,aa) $ sum(sameas(h,aa),1)),1)
*   ) $ rsNat(rNat1)
*   + xa.l(rNat1,i,aa,t) $ (not rsNat(rNat1))
*   )
*   * (pa(rNat1,i,aa,t)/m_pAgentm(rNat,i,aa,rNat1,t)) * p_sigmaMel(rNat,i,rNat1)s(xw,'rNat,i,rNat1,t');
*

```

The fix cost price is equal to the price of the fix cost nest “*fCost*” if a separate fix cost nest is active defined in the *e\_ptNest* equation in the core model. If no separate fix cost nest is used, the fix cost price is defined in the equation as equal to unit costs of production:

```

*
* --- no fix cost nest: unit cost price of fix costs equal to producer price
*
* e_FcostP(rs(r),"fCost",a,ts(t)) $ (sum(iMel $ xFlag(r,a,iMel),p_MFirmsEnt0(r,iMel,t))
*   ) ..
*   pTNest(r,"fCost",a,t) =E= px(r,a,t);

```

The price markup on each trade link is indirectly defined in *e\_MKUP* equation by the endogenous average firm productivity *phiFirm* and the price charged by the average firm *pFirm*:

```

*
* --- price markup definition (= unit cost times relative iceberg transport costs,
*   corrected for productivity, and for sigmaMel)
*
* e_MKUP(rNat,iMel(i),rrd,ts(t)) $ ((xwFlagMel(rNat,i,rrd) $ (sum(sameas(rNat1,rrd),rrComb(rNat,rNat1)) or rsNat(rNat))) $ iIn(i))
*   $ (pFirm.range(rNat,i,rrd,t) ne 0) ) ..
*
*   p_tau(rNat,i,rrd,t)*ps(rNat,i,t)*1/(1 - 1/p_sigmaMel(rNat,i,rrd))/p_phiFirm0(rNat,iMel,rrd,t)
*   =E= phiFirm(rNat,i,rrd,t)/p_phiFirm0(rNat,iMel,rrd,t) * pFirm(rNat,i,rrd,t);
*

```

The endogenous number of firms entering the industry *mFirmsEnt* is defined as follows:

```

*
* e_firmsEnt(rsNat(rNat),iMel(i),ts(t)) $ (p_MFirmsEnt0(rNat,iMel,t) $ (MFirmsEnt.range(rNat,i,t) ne 0) $ iIn(i)) ..
*
*   p_delt_fs(rNat,i,t)*vFirmsEnt(rNat,i,t)
*   / (p_MFirmsEnt0(rNat,i,t)*p_delt_fs(rNat,i,t))
*
*   --- nFirmsOp * qAv is replaced by the resulting quantities xw and xds to simplify the equation
*
*   =E= sum(rrd $ xwFlagMel(rNat,i,rrd),
*   (sum(sameas(rNat,rrd), xw(rNat,i,rNat,t)) + xds(rNat,i,t) $ sameas(rrd,"domes"))
*   $$iftheni "%HETSetup%"=="Melitz"
*   * pFirm(rNat,i,rrd,t)/pFirm.l(rNat,i,rrd,"%t0%") * (p_sigmaMel(rNat,i,rrd)-1)/(p_a*p_sigmaMel(rNat,i,rrd))
*   $$else
*   * (pFirm(rNat,i,rrd,t)/pFirm.l(rNat,i,rrd,"%t0%")-p_tau(rNat,iMel,rrd,t)*ps(rNat,i,t))
*   $$endif
*
*   )/(p_MFirmsEnt0(rNat,i,t)*p_delt_fs(rNat,i,t));

```

It is based on the total revenues *vFirmsEnt* defined in the following equation:

```

*
* e_vFirmsEnt(rsNat(rNat),iMel(i),ts(t)) $ (p_MFirmsEnt0(rNat,iMel,t) $ (MFirmsEnt.range(rNat,i,t) ne 0) $ iIn(i)) ..
*
*   vFirmsEnt(rNat,i,t)/p_MFirmsEnt0(rNat,i,t)
*   =E= MFirmsEnt(rNat,i,t)*sum(a $ xFlag(rNat,a,i), pTNest(rNat,"fCost",a,t))/p_MFirmsEnt0(rNat,i,t);

```

Note the difference between the Melitz set-up where fix costs and mark-ups are trade link specific and the Krugman model where only industry wide fix costs are present.

In the Krugman model, only the total industry size *MFirmsEnt* changes and updates also the number of firms being active on each trade link *nFirmsOp* as defined by *e\_krug* equation:

```
e_krug(rsNat(rNat),iMel(i),rrd,ts(t)) $ xwFlagMel(rNat,i,rrd) ..
NfirmsOp(rNat,i,rrd,t)/p_nFirmsOp0(rNat,i,rrd,t) =E= MFirmsEnt(rNat,i,t)/p_MFirmsEnt0(rNat,i,t);
```

The number of firms operating on a trade link in the Melitz case is defined by the zero-cutoff profit equation *e\_ZCP*. The LHS defines the per unit fixed cost. The RHS ensures that these fix costs are exhausted by the number of firms being active, reflecting the productivity distribution, considering the fixed cost share *p\_fc*:

```
*
* --- zero cutoff profit, determines number of firms operating nFirmsOP on link r->rrd
*
e_ZCP(rNat,iMel(i),rrd,ts(t)) $ (xwFlagMel(rNat,i,rrd) $ (nFirmsOp.range(rNat,i,rrd,t) ne 0)
$ (sum(sameas(rNat1,rrd),rrComb(rNat,rNat1)) or rsNat(rNat)) $ iIn(i) ) ..
*
* --- fixed costs (= unit costs times fix cost share)
*
sum(a $ xFlag(rNat,a,i), pTnest(rNat,"fCost",a,t))
=E= QAvFirm(rNat,i,rrd,t)/p_fc(rNat,i,rrd,t)
* pFirm(rNat,i,rrd,t)/pFirm.1(rNat,i,rrd,"%t0%") * (p_a+1-p_sigmaMel(rNat,i,rrd))/(p_a*p_sigmaMel(rNat,i,rrd));
*
```

The pareto productivity *phiFirm* which enters the price markup equation is defined in the equation *e\_PAR*.

```
*
* --- pareto productivity
*
e_PAR(rNat,iMel(i),rrd,ts(t)) $ (xwFlagMel(rNat,i,rrd) $ (sum(sameas(rNat1,rrd),rrComb(rNat,rNat1)) or rsNat(rNat)) $ iIn(i) ) ..
phiFirm(rNat,i,rrd,t)s(phiFirm,'rNat,iMel,rrd,t')
=E= p_b s(phiFirm,'rNat,iMel,rrd,t') * (p_a/(p_a+1-p_sigmaMel(rNat,i,rrd)))+(1/(p_sigmaMel(rNat,i,rrd)-1))
* [MFirmsEnt(rNat,i,t)/NfirmsOp(rNat,i,rrd,t)]^(1/p_a);
```

Variable cost exhaustion is ensured by the *e\_MKT* equation:

```
*
* --- input market, ie. input demand from GTAP CES nests equals to variable costs Melitz model
*
e_MKT(rsNat(rNat),iMel(i),ts(t)) $ (p_MFirmsEnt0(rNat,iMel,t) $ iIn(i) ) ..
sum(a $ xFlag(rNat,a,i), m_xp(rNat,a,t))
%scale% /sum(a $ xFlag(rNat,a,i),xp.scale(rNat,a,ts))
=E=
*
* --- icebeg cost of trade p_tau times output per firm QAvFirm times # of operating firm,
* corrected for average productivity
{ sum(rrd $ xwFlagMel(rNat,i,rrd),
[sum(sameas(rpNat,rrd), xw(rNat,i,rpNat,t))+xds(rNat,i,t) $ sameas(rrd,"domes")])
*p_tau(rNat,i,rrd,t)
$if "HETSetup"=="Melitz" /phiFirm(rNat,i,rrd,t)
)
*
-- costs of intermediate input use by same industry
+ sum(aa $ (p_amel(rNat,i,aa,"domes",t) $ p_melDomUse(rNat,aa,i)), xdMFlag(rNat,i,aa,t))
$if "%fcost_nest%" == "off" + sum(a $ xFlag(rNat,a,i), xtNest(rNat,"fCost",a,t))
}
%scale% /sum(a $ xFlag(rNat,a,i),xp.scale(rNat,a,ts))
;
```

Note that the total output comprises fix cost input demand if no separate fix costs nest is active. Therefore, these costs are added for that case, see the line starting with *\$\$ifi*.

Fix costs are defined as follows in the *e\_fCost* equation. They consist of industry wide fix costs, i.e. the endogenous number of firms in the industry *MFirmEnt* times the given fix costs per firm *p\_delt\_fs*, plus the given fix cost on each trade link *p\_fc* times endogenous number of the firms operating on the link defined by the *m\_NFirmsOp* macro:

```

*
* --- total fix cost demand
*
e_fCost(rsNat(rNat),"fCost",a,ts(t)) $ (sum(iMel $ xFlag(rNat,a,iMel),p_MFirmsEnt0(rNat,iMel,t)) $ aIn(a)) ..
    xTNest(rNat,"fCost",a,t)s(xTNest,'rNat,"fCost",a,t') =E=
*
*      --- fix costs times # of existing firms (= industry level, Krugmann)
*
    sum(i $ xFlag(rNat,a,i), p_delt_fs(rNat,i,t)*MFirmsEnt(rNat,i,t) +
*
*      --- # of firms operating on the different markets times
*      times given bilateral fix cost p_fc (Melitz)
*      sum(rrd $ xwFlagMel(rNat,i,rrd), m_NFirmsOp(rNat,i,rrd,t)*p_fc(rNat,i,rrd,t))
*      )s(xTNest,'rNat,"fCost",a,t');
*

```

In case of sub-national units, the following equation define the average fix cost price and the sourcing of fix cost from sub-regions:

```

*
* --- CET-aggregator of fix cost across regions
*
e_fCostR(rs(subr),"fCost",a,ts(t)) $ (sum(r_r(subr,disr),iMel) $ xFlag(subr,a,iMel),p_MFirmsEnt0(disr,iMel,t)) $ aIn(a)) ..
    xTNest(subr,"fCost",a,t)s(xTNest,'subr,"fCost",a,t')
    =E= sum(r_r(subr,disr) $ p_gff(subr,a,disr,t),
    xTNest(disr,"fCost",a,t) * p_gff(subr,a,disr,t)
    * (ptNest(disr,"fCost",a,t)/ptNest(subr,"fCost",a,t))**p_sigmaASubr(disr,a))s(xTNest,'subr,"fCost",a,t');
*
* --- matching dual price aggregator
*
e_fCostN(disr,"fCost",a,ts(t)) $ (sum(iMel $ xFlag(disr,a,iMel),p_MFirmsEnt0(disr,iMel,t)) $ aIn(a) $ rs(disr)) ..
    ptNest(disr,"fCost",a,t) =E= sum(r_r(subr,disr) $ p_gff(subr,a,disr,t),
    p_gff(subr,a,disr,t)*ptNest(subr,"fCost",a,t)**(1-p_sigmaASubr(disr,a))**(1/(1-p_sigmaASubr(disr,a))));
*

```

There is a block of equations which defines variables used elsewhere in the model, but not defined by the equation in the Melitz model:

```

*
* --- the following equations calculate elements which are removed from the normal GTAP Model
*      by using only one Armington nest and aggregating over agents in the Melitz model,
*
* --- definition of total exports from average output per firm and # of firms
*      operating on the trade link rp->r
*
e_xwMel(rNat1,iMel(i),rNat,ts(t)) $ ( xwFlagMel(rNat1,i,rNat) $ rrComb(rNat1,rNat) $ iIn(i)) ..
    xw(rNat1,i,rNat,t)s(xw,'rNat1,i,rNat,t') =E= m_xw(rNat1,i,rNat,t)s(xw,'rNat1,i,rNat,t');
*
e_xdsMel(rsNat(rNat1),iMel(i),ts(t)) $ (xwFlagMel(rNat1,i,"domes") $ iIn(i)) ..
    xds(rNat1,i,t)s(xds,'rNat1,i,t') =E= m_xds(rNat1,i,t)s(xds,'rNat1,i,t');
*
* --- total export demand, aggregated from bi-lateral export quantities
*
e_xetMel(rsNat(rNat),iMel(i),ts(t)) $ (sum(rNat1, xwFlagMel(rNat,i,rNat1)) $ iIn(i)) ..
    xet(rNat,i,t)s(xet,'rNat,i,t') =E= sum(rNat1 $ xwFlagMel(rNat,i,rNat1), xw(rNat,i,rNat1,t))s(xet,'rNat,i,t') ;
*

```

The average price of imports is implicitly defined:

```

*
* --- implicitly define average price of exports from exhaustion
*
e_petMel(rsNat(rNat),iMel(i),ts(t)) $ (sum(rNat1,xwFlagMel(rNat,i,rNat1)) $ iIn(i)) ..
*
*      --- value of exports
*
    xet(rNat,i,t)*pet(rNat,i,t)s(xet,'rNat,i,t')
*
    =E=
*
*      --- value of total imports, defined from value of each import flow
*      (xm comprises the transport margins and the original border protection,
*
    sum(rNat1 $ xwFlagMel(rNat,i,rNat1), xw(rNat,i,rNat1,t)*m_pFirm(rNat,i,rNat1,t))s(xet,'rNat,i,t') ;
*

```

In order to check for the correct functioning of model, a revenue check equation *e\_revCheck* is introduced:

```

* --- check revenue exhaustion (variable and fix costs are recovered by sales revenues)
e_revCheck(rsNat(rNat),iMel(i),ts(t)) $ (p_MFirmsEnt0(rNat,iMel,t) $ iIn(i)) ..

revCheck(rNat,i,t)
  %scale% /sum(a $ xFlag(rNat,a,i),xp.scale(rNat,a,t))
  =E=
  (sum(rrd $ xwFlagMel(rNat,i,rrd),
    (sum(sameas(rNat1,rrd), xw(rNat,i,rNat1,t)) + xds(rNat,i,t) $ sameas(rrd,"domes"))*pFirm(rNat,i,rrd,t))
    + sum(aa $ (m_alphad(rNat,i,aa,t) $ p_melDomUse(rNat,aa,i)), xdmFlag(rNat,i,aa,t) $ ps(rNat,i,t))
    -sum(a $ xFlag(rNat,a,i), m_xp(rNat,a,t) $ ps(rNat,i,t))
    -sum(a $ xFlag(rNat,a,i), xtNest(rNat,"fCost",a,t) $ pTNest(rNat,"fCost",a,t) $ (1+prdx(rNat,a,i,t)))
    )
  %scale% /sum(a $ xFlag(rNat,a,i),xp.scale(rNat,a,t))
;

```

The variable *revCheck* on the LHS should yield a zero in the benchmark and under a simulation. The first term in *sum(rrd ...)* adds the total spendings of the agents for sales subject to markups. The next line adds the revenues from own intermediate consumption not subject to markets. From this, total production costs are subtracted: the production tax inclusive per variable per unit cost *ps* times total output *m\_xp* and that value of the fix cost, considering production taxes.

## 2.14 Price indices

The model defines different price indices which can be used as regional (or global) numeraires and/or for reporting purposes.

Average factor prices *pft* and total stock *xft* for non-non-mobile factors *fnm* are jointly defined by the following four equations. The last equation defines the revenues at market prices by adding up the factor demands *xf* of the activities times the activity specific factor prices *pf*. For model regions with sub-regions *disr*, revenues are summed up over the sub-region regions. Economy-wide factor use *xft* is based on physical aggregation of the factor use by activity in equation *e\_xftFnm*. Factor prices *pft* are then defined by dividing total revenues by total factor use in equation *e\_pftFnm*.

```

--- calculate average price of non-mobile factors
e_pftFnm(rs,fnm,ts) $ ( (xftEqFlag(rs,fnm) or disr(rs)) $ sum(aIn $ xftFlag(rs,fnm,aIn), 1) ) ..

pft(rs,fnm,ts) $ (pft,'rs,fnm,ts')
  =E= (vft(rs,fnm,ts)/xft(rs,fnm,ts)) $ (pft,'rs,fnm,ts');

e_xftFnm(rs,fnm,ts) $ (xftEqFlag(rs,fnm) and (subr(rs) or (not sameas(fnm,"newCap"))))
  $ sum(aIn $ xftFlag(rs,fnm,aIn),1) $ (xft.range(rs,fnm,ts) ne 0)
  $ $if %modulesGMIG%==on and (not 1(fnm))
  ) ..

xft(rs,fnm,ts) $ (xft,'rs,fnm,ts')
  =E= [sum(a $ xftFlag(rs,fnm,a), xft(rs,fnm,a,ts)) $ (xft,'rs,fnm,ts')] $ (not disr(rs))
  + [sum(r_r(subr,rs) $ xftFlag(subr,fnm), xft(subr,fnm,ts)) $ (xft,'rs,fnm,ts')] $ disr(rs);

e_vft(rs,f,ts) $ (sum(aIn $ xftFlag(rs,f,aIn),1) and (vft.range(rs,f,ts) ne 0)) ..

vft(rs,f,ts) $ (vft,'rs,f,ts')
  =E= [sum(a $ xftFlag(rs,f,a), xft(rs,f,a,ts) * pf(rs,f,a,ts)) $ (vft,'rs,f,ts')] $ (not disr(rs))
  + [sum(r_r(subr,rs) $ xftFlag(subr,f), vft(subr,f,ts)) $ (vft,'rs,f,ts')] $ disr(rs);

```

**Regional factor price indices** *pfact* are defined in the equation *e\_pfact* based on factor prices *pft* and weights *phif* and are used to define them in the benchmark, whereas in shocks or follow up years a Fisher index is used (equations (93) and (94) in VDM 2018).

```

e_pfact(rsNat(rNat2),ts(tRun)) $ ( (pfact.range(rNat2,tRun) ne 0) or (not rres(rNat2)) or (lcu.range(rNat2,tRun) ne 0) ) ..

pfact(rNat2,tRun) =e= sum(f, p_phif(rNat2,f,tRun)*pft(rNat2,f,tRun)) $ sameas(tRun,"%t0%")
  + [ pfact.1(rNat2,tRun-1) * sqrt( m_qfactr(tRun,tRun-1)/m_qfactr(tRun-1,tRun-1)
    * m_qfactr(tRun,tRun) /m_qfactr(tRun-1,tRun) ) ] $ (not sameas(tRun,"%t0%"));

```

The Fisher-index uses the macro *q\_factr* shown below. Here, *tp* stands for the time point defining the prices and *tq* for the time point defining the quantities. If *tp* or *tq* are equal to the current simulation time point *ts*, the endogenous quantity and prices *xft* and *pft* are used, otherwise the stored levels from previous solution *xft.l* and *pft.l* are entering the expressions.

```

$macro m_qfactr(tp,tq) sum (f $ (xft.l(rNat2,f,tp)), \
  ( (pft(rNat2,f,tp)/lcu(rNat2,tp)) $ (ts(tp) and rsNat(rNat2)) \
    + (pft.l(rNat2,f,tp)/lcu.l(rNat2,tp)) $ ( (not ts(tp)) or (not rsNat(rNat2))) ) \
    * ( xft(rNat2,f,tq) $ (ts(tq) and rsNat(rNat2)) \
      + xft.l(rNat2,f,tq) $ ( (not ts(tq)) or (not rsNat(rNat2))) ) )

```

The **average world price of factors** *pwfact* as defined in the equation *e\_pwfact* uses weights *phifw* for the model regions and reflects the exchange rates *lcu* to aggregate the regional factor prices *pft*. It is global numeraire price in the model, equations (95) and (96) in VDM 2018.

```

--- average world price of factors, equation (95) and (96) in VDM 2018
    fixed to numeraire with the following equation

$macro m_qfactw(tp,tq) sum ( rNat2 $ sum(f,p_phifw(rNat2,f,tp)), m_qfactr(tp,tq))

e_pwfact(ts(tRun)) $ (pwfact.range(tRun) ne 0) ..

pwfact(tRun) =e sum((rNat2,f), p_phifw(rNat2,f,tRun) * ( pft(rNat2,f,tRun) $ rsNat(rNat2)
  + pft.l(rNat2,f,tRun) $ (not rsNat(rNat2)) ) /lcu(rNat2,tRun)) $ sameas(tRun,"%t0%")

+ [ pwfact.l(tRun-1) * sqrt( m_qfactw(tRun,tRun-1)/m_qfactw(tRun-1,tRun-1)
  * m_qfactw(tRun,tRun) /m_qfactw(tRun-1,tRun) ) ] $ (not sameas(tRun,"%t0%"));

```

**Regional producer price indices** *pprod* as defined in the equation *e\_pprod* are based on weights *p\_phii*:

```

*
* --- Regional producer price index,
*
e_pprod(rsNat(rNat),ts(t)) $ ( ((pprod.range(rNat,t) ne 0) or (not rres(rNat)) or (lcu.range(rNat,t) ne 0)) ) ..
  pprod(rNat,t) =e sum(i, p_phii(rNat,i,t)*ps(rNat,i,t)) ;

```

They are for information, only, and not used in other equations.

**Average domestic consumption prices** *pabs* are an average of the Armington prices for the different types of final demand *fd* (final demand prices for households, government, investment and domestic supply of trade margins) and weights *p\_phia*. That is an approximate version of equation (90) in VDM 2018. That price index is not used elsewhere in the model.

```

*
* --- price index of domestic consumption, equation (90) in VDM 2018 (only approximate, not used in model equations)
*
e_pabs(rsNat,ts) ..
  pabs(rsNat,ts) =e sum((i,fd) $ p_phia(rsNat,i,fd,ts), p_phia(rsNat,i,fd,ts)*m_pa(rsNat,i,fd,ts)) ;

```

## 2.15 List of main prices in model

Table 3: Prices in model

| Variable | Content                                        | Indices     |
|----------|------------------------------------------------|-------------|
| px       | Unit costs of production                       | r,a,t       |
| pf       | Activity specific factor price, tax exclusive  | r,f,a,t     |
| pfa      | Activity specific factor price, tax inclusive  | r,f,a,t     |
| ptnest   | Exhaustion price index of technology nest      | r,tNest,a,t |
| pdtNest  | CES price index of technology nest (with ACES) | r,tNest,a,t |
| pfNest   | CET price index of factor supply nest          | r,fnest,t   |

| Variable  | Content                                                       | Indices      |
|-----------|---------------------------------------------------------------|--------------|
| pft       | Aggregate price of factors                                    | r,f,t        |
| p         | Price of output                                               | r,a,i,t      |
| pavgVint  | CES price index of output over vintages                       | r,a,i,t      |
| pavgVintD | Exhaustion price of output over vintages                      | r,a,i,t      |
| pp        | Production tax inclusive price                                | r,a,i,t      |
| ps        | Exhaustion price of domestic supply                           | r,i,t        |
| pds       | CET price index of domestic supply (additive CET)             | r,i,t        |
| pmNest    | Exhaustion price index of make nest                           | r,mNest,i,t  |
| pmNestD   | CES price index of make nest (additive CES)                   | r,mNest,i,t  |
| pe        | Prices for bilateral export supply                            | r,i,rp,t     |
| pet       | Average price of export supply                                | r,i,t        |
| pefob     | Border price of exports (free on board)                       | r,i,rp,t     |
| pmcif     | Border price of imports (cost, insurance, freight)            | r,i,rp,t     |
| pm        | Bilateral price of imports, tax inclusive                     | r,i,rp,t     |
| pmt       | Exhaustion price index of imports                             | r,i,t        |
| pmtD      | CES price of imports (with ACES)                              | r,i,t        |
| pmtMrio   | Price of aggregate imports, by mrio agent                     | r,i,mrioA,t  |
| pd        | Price of domestically produced good                           | r,i,t        |
| pdp       | Purchaser price of domestic good                              | r,i,aa,t     |
| pdNest    | Exhaustion price index of final demand sub-nests              | r,dNest,fd,t |
| pdNest    | CES price aggregator of sub-nest below final demand equations | r,dNest,fd,t |
| pa        | Armington exhaustion price                                    | r,i,aa,t     |
| pad       | CES Armington price index (additive CES)                      | r,i,aa,t     |
| pm        | Bilateral price of imports, tax inclusive                     | r,i,rp,t     |
| pmp       | Public expenditure price deflator                             | r,i,aa,t     |
| pcons     | Consumer price deflator                                       | r,h,t        |
| pi        | Investment expenditure price deflator                         | r,t          |
| pg        | Public expenditure price deflator                             | r,t          |
| pfact     | Factor price index                                            | r,t          |
| pprod     | Producer price index                                          | r,t          |
| pabs      | Price of aggregate domestic absorption                        |              |
| pwfact    | World factor price index                                      | t            |
| ptmg      | Global price index of transport services by mode              | m,t          |

## 2.16 Flexible nesting in the production function

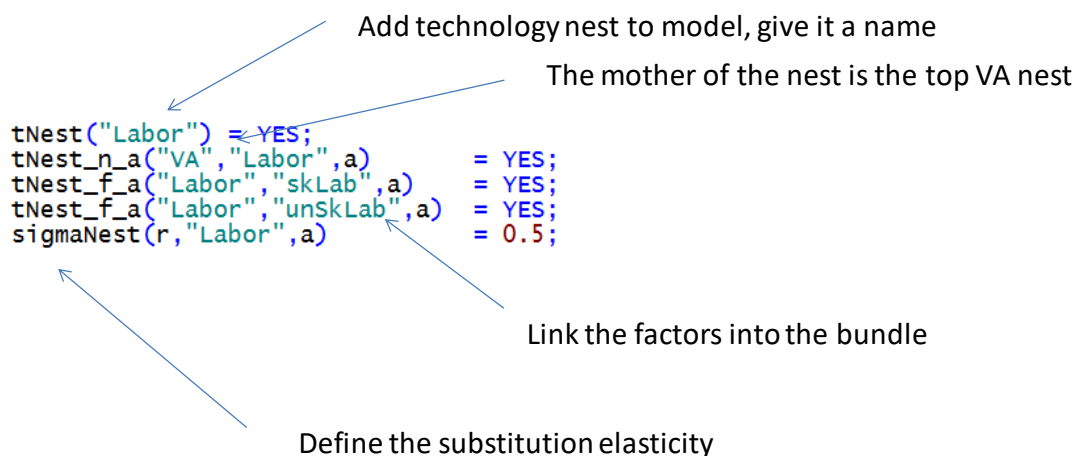
While most CGE models apply the CES functional form to depict the production function, quite some differences exist how input composites are defined. Usually, each CES nest is represented in the model's programming code by its own quantity and price aggregator equation. Adding or changing nests thus requires coding efforts – new variables, equations and parameters need to be defined and code added for benchmarking of the parameters. Often, also the equations relating to the top-level value-added aggregator and/or the intermediate demand nest need to be adjusted if the nesting is changed.

The GAMS code underlying CGEBox applies a different strategy. Here, nested CES structures in the production function are represented by a generic approach where a small number of equations and matching variables handle different nesting structures. In the standard GTAP-implementation, with the exemption of price indices for the value-added and intermediate composite nest, such sub-nests are not present and related equations empty.

This flexible and generic nesting approach is based on sets and cross-sets in GAMS which define the lists of factors, intermediates and sub-nests comprised in a CES composite nest, along with the upper-level nest it belongs too. These nesting definitions enter the equations in the core model mentioned above and matching code dealing with parameter calibration. Hence, the user does not need to introduce additional equations in the code to use that feature – it is sufficient to provide the structure of the nesting to be applied via set definitions and the related substitution elasticities. The code also tests for potential errors such as duplicate assignments or sub-nests not linked into another nests.

The following examples should be sufficient to show the application of that feature and demonstrate its flexibility. Please note that the elasticity parameter in the current model version carry a “*p\_*” in front, in the model code, please use *p\_sigmaNest* where the screen shots still show *sigmaNest*.

1. An example of a sub-nest under the top VA nest which aggregates the two labor categories found in the GTAP 8 data base into an aggregate:



Note: The code will automatically remove the factors linked into nests from the top VA nest



The second example shows how to introduce a CES-composite of intermediates linked into the top ND-nest, as e.g. applied in the GTAP-E nesting, or to allow for high substitution of feed intermediates in animal production:

```

tNest("primEne") = YES;
Set primEne(i) /
    coa-c
    oil-c
    gas-c
/;
alias(primEne, primEne1);
tNest_n_a("ND", "primEne", a)
tNest_i_a("primEne", primEne, a)
sigmaNest(r, "primEne", a)

```

Add technology nest to model, give it a name

Define a helper set for the intermediates (not used elsewhere)

The mother of the nest is the top ND-bundle

```

$ sum( (r, primEne1, t), io(r, primEne1, a, t)) = YES;
$ sum( (r, primEne1, t), io(r, primEne1, a, t)) = YES;
$ sum( (primEne1, t), io(r, primEne1, a, t)) = 0.5;

```

Link the intermediates into the bundle

Define the substitution elasticity

Note: The code will automatically remove intermediates linked into nests from the top ND bundle

- The third example shows how to combine intermediates and factors into a CES composite, which in the example is a sub-nest of the top VA nest:

```

tNest("energy") = YES;
Set scndEne(i) /
    p_c-c
    ely-c
    gdt-c
/;
tNest_n_a("VA", "energy", a) = YES;
tNest_f_a("energy", "capital", a) = YES;
tNest_i_a("energy", scndEne, a) = YES;
sigmaNest(r, "energy", a) = 0.5;

```

Add technology nest to model, give it a name

Define a helper set for the intermediates (not used elsewhere)

The mother of the nest is the top VA-bundle

Link the factors into the bundle

Link the intermediates into the bundle

Define the substitution elasticity

- The last example reproduces the nesting of the GTAP-E model (the composite of skilled and unskilled labor is already shown above):

```

*
* --- top level nest, aggregate of capital and energy
*
tNest("CAP+ENE") = YES;
tNest_n_a("VA","CAP+ENE",a) = YES;
tNest_f_a("CAP+ENE","capital",a) = YES;
tNest_n_a("CAP+ENE","energy",a) = YES;
sigmaNest(r,"CAP+ENE",a) = 0.25;

tNest("energy") = YES;
tNest_i_a("energy","ele-c",a) = YES;
tNest_n_a("energy","non-electric",a) = YES;
sigmaNest(r,"energy",a) = 1.00;

tNest("non-electric") = YES;
tNest_i_a("non-electric","coa-c",a) = YES;
tNest_n_a("non-electric","non-coal",a) = YES;
sigmaNest(r,"non-electric",a) = 0.50;

tNest("non-coal") = YES;
tNest_i_a("non-coal","gas-c",a) = YES;
tNest_i_a("non-coal","oil-c",a) = YES;
tNest_i_a("non-coal","p_c-c",a) = YES;
sigmaNest(r,"non-coal",a) = 1.00;

```

The actual code in use for the GTAP-E is more complex as it reflects a potential aggregation of the detailed products, with statements such as:

```

set nonCoal0 / gas,oil,p_c/;
set nonCoal(i); nonCoal(i) $ sum(mapi(nonCoal0,i),1) = yes;

```

Depending on the resulting sets, sub-nests might be skipped if they collapse into one commodity.

The nesting used in a specific model set-up along with the substitution elasticities are reported in a table in the exploitation tools. Equally, the code reports the quantity and price aggregators for nests and the resulting values and aggregates them up over individual sectors and over regions.

It is important to note that the post-model reporting redefines the top-level VA and ND nests such that they match the usual definition, i.e. an aggregation over primary factors and intermediates, respectively. Information about these nests in the definition used in the model can be retrieved with the separate tables showing all technology nests.

A similar generic approach is implemented for factor supply: nested CET functions can be used to supply primary factors to the production sectors. The top-level nest is labeled with “*xft*”. The following example used in the GTAP-AGR implementation shows that approach:

```

fNest("agr") = yes;
fNest_a_f("agr",agr,agrFNest) = yes;
fNest_n_f("xft","agr",agrFNest) = yes;
p_omegafNest(r,"agr",agrFNest) = inf;

fNest("nonagr") = yes;
fNest_a_f("nonagr",nonagr,agrFNest) = yes;
fNest_n_f("xft","nonagr",agrFNest) = yes;
p_omegafNest(r,"nonagr",agrFNest) = inf;

```

Again, the definition of what activities are agricultural ones is endogenously determined based on the mapping from the GTAP sectors to the sectors used in the benchmark SAM:

```

set agr0(i0) / pdr    "Paddy rice"
               wht    "Wheat"
               gro    "Cereal grains nec"
               v_f    "Vegetables, fruit, nuts"
               osd    "Oil seeds"
               c_b    "Sugar cane, sugar beet"
               pfb    "Plant-based fibers"
               ocr    "Crops nec"
               ctl    "Cattle,sheep,goats,horses"
               oap    "Animal products nec"
               rmk    "Raw milk"
               wol    "Wool, silk-worm cocoons"
/;

```

```

set agr(a); agr(a) $ sum(mapa(agr0,A),1) = YES;

```

```

set nonAgr(a); nonAgr(a) $ (not agr(a)) = YES;

```

Similar to the CES production nests, the post-model processing reports the structure, parameterization and simulation results in the exploitation tools.

Equally, these flexible nesting structures are also applicable for final demand (government, investments and households), as shown here again with a nesting used by GTAP-E:

```

tNest_n_a("energy","non-electric",a)    $ tNest("non-electric")    = YES;
tNest_i_a("energy",coal,a)              $ (not tNest("non-electric")) = YES;
tNest_i_a("energy",nonCoal,a)           $ (not tNest("non-electric")) = YES;

dNest_n_fd("energy","non-electric",fdn) $ dNest("non-electric")    = YES;
dNest_i_fd("energy",coal,fdn)           $ (not dNest("non-electric")) = YES;
dNest_i_fd("energy",nonCoal,fdn)        $ (not dNest("non-electric")) = YES;
dNest_n_fd("top","energy",fdn)          = YES;
p_sigmaFDNest(r,"energy",fdn)           = 1.01;

```

Finally, the model can also aggregate a product outputted by multiple activities (make matrix) over multiple CES nests. This is, for instance, used to yield the same composition of electricity demand across all demanders with the GTAP-Power nesting. It requires that the user aggregates the GTAP-Power data base accordingly to comprise one electricity product, ony:

```

*
*   --- separate nests for base and peak load power generation
*
mNest("ebl") = yes;
mNest("epl") = yes;
*
*   --- Leontief bundle of base and peak load define the electricity product
*
mNest_n("top","ebl",i)    $ sameas(i,"ely-c") = yes;
mNest_n("top","epl",i)    $ sameas(i,"ely-c") = yes;
p_sigmasNest(rNat,"top",i) $ sameas(i,"ely-c") = eps;
*
*   --- Base load is a flexible composite of base load power generation activities
*
mNest_a("ebl",a,i) $ (sum(sameas(ebl_a,a),1) $ sameas(i,"ely-c")) = YES;
p_sigmasNest(rNat,"ebl",i) $ sameas(i,"ely-c") = 1.386;
*
*   --- Peak load is a less flexible composite of peak load power generation activities
*
mNest_a("epl",a,i) $ (sum(sameas(epl_a,a),1) $ sameas(i,"ely-c")) = YES;
p_sigmasNest(rNat,"epl",i) $ sameas(i,"ely-c") = 0.472;

```

An advantage of defining the composition of supply at the level of make nests is that it avoids to define the average price and sourcing instead at the level of each demanders based on nests in the production function and final demands, saving variables and equations.

The code allows for factors or intermediate inputs to be linked to several nests based on share defined at the benchmark.

As mentioned, the flexible nesting requires additional equations in the model. We discuss these here briefly for the production function (the equation was introduced already above), the approach for factor supply, final demand and the make composition is similar. The input composite for a nest consists potentially of factors, intermediate inputs and other nests. The resulting price index for a nest  $ptNest$  is defined as usually in the dual formulation (this is a simplified version of the actual model code that supports also the additive CES):

```
e ptNest(rs,tNest,aIn(a),ts) $ xtFlag(rs,tNest,a) ..
ptNest(rs,tNest,a,ts) =e=
  sum(r_r(rs,rsNat),
  (
  { [
*
* --- aggregate over intermediate inputs assigned to technology nest
*
    sum(tNest_i_a(tNest,i,a) $ p_ioNest(rs,i,a,tNest,ts),
        p_ioNest(rs,i,a,tNest,ts)
        * (m_pa(rsNat,i,a,ts)/lambdaio(rs,i,a,ts))**(1-p_sigmaNest(rs,tNest,a)))
*
* --- aggregate over factors assigned to technology nests
*
    + sum(tNest_f_a(tNest,f,a) $ p_afNest(rs,f,a,tNest,ts),
        p_afNest(rs,f,a,tNest,ts)
        * (m_pfa(rs,f,a,ts)/lambdaf(rs,f,a,ts))**(1-p_sigmaNest(rs,tNest,a)))
*
* --- aggregate over sub technology nests assigned to that technology nests
*
    + sum(tNest_n_a(tNest,tNest1,a) $ p_atNest(rs,tNest1,a,ts),
        p_atNest(rs,tNest1,a,ts)
        * (ptNest(rs,tNest1,a,ts)/lambdaN(rs,tNest1,a,ts))**(1-p_sigmaNest(rs,tNest,a)))
    ]** (1/(1-p_sigmaNest(rs,tNest,a)))
  )
  $ (p_sigmaNest(rs,tNest,a) ne 1)
```

The first summation aggregates the intermediate inputs assigned to the nest, based on the nest-specific input coefficient  $ioNest$ , Hicks non-neutral productivity shifters  $lambdaio$  and the sectoral specific tax inclusive price for the intermediate via the macro  $m\_pa$ . In a similar fashion, the second summation blocks aggregate factor use, while the last block considers sub-nests linked into the current nests. That implies that a multi-level nested structure does not require additional equations in the model.

The Armington intermediate demand  $xa$  for commodity  $i$  by sector  $a$  is shown below. It is driven by the relation between the price  $ptNest$  for the nests it belongs to and the Armington price  $m\_paInt$ , times the total demand for the nest  $xtNest$ , and the benchmark cost shares  $ioNest$ . A technology shifters  $lambdaIo$  is considered as well. The relevant prices are either  $ptNest$  for the standard case or  $pdtNest$  for the additive one.

```
--- Demand for intermediates, equation (6) in VDM 2018, with extension for commodity nests
$ xap(rs,iIn(i),aIn(a),ts) $ (xapFlag(rs,i,a) $ (not disr(rs)) ) ..

xa(rs,i,a,ts)s(xa,'rs,i,a,ts') =e=
+ sum(tNest_i_a(iotNest,i,a) $ p_ioNestT(rs,iotNest,a,i),
--- input coefficient times demand for technology nest
p_ioNestT(rs,iotNest,a,i) * xtNest(rs,iotNest,a,ts)s(xa,'rs,i,a,ts')

--- relative price impact (in case sigmaNest is not 0)
* ( ([ptNest(rs,iotNest,a,ts) $ (not addCES(rs,iotNest,a)) + pdTNest(rs,iotNest,a,ts) $ addCes(rs,iotNest,a)]
/m_paInt(rs,i,a,ts))**p_sigmaNest(rs,iotNest,a)-1] $ (p_sigmaNest(rs,iotNest,a) gt eps) + 1)
* lambdaio(rs,i,a,ts)**( [p_sigmaNest(rs,iotNest,a)-1] $ (Not addCes(rs,iotNest,a))
- p_sigmaNest(rs,iotNest,a) $ addCes(rs,iotNest,a) )
);
```

Not that one intermediate input  $i$  can be demanded by different nests. Similar equations drive the demand for individual factors  $xf$  and the demand for nests  $xtNest$  as shown above in the documentation of the model's equation.

### 3. Modules

The following chapters describe the different modules which are currently integrated into CGEBox. In most cases, the code of a module consists of (potentially) the following separate blocks of code, typically each in its own file:

1. Code declaring and defining additional data and parameters needed for the module
2. Code declaring and defining the necessary equations
3. Code for setting starting values of variables and parameters for the next solution
4. Code for setting bounds of variables
5. Code for defining scaling factors variables
6. Code for defining a start point for the next solution
7. Code used for post-model processing of the results

A module introduces additional or alternative equations into the overall model. Some modules are mutually exclusive, such as different option to define aggregates of Armington agents. For others, settings might need to be adjusted for combined use. For instance, the Armington with commitment terms, the Melitz specification and the spatial arbitrage formulation are mutually exclusive for one international product market, but it is possible to define for each non-overlapping lists of products to introduce these options in parallel into one model configuration.

#### 3.1 Capital vintages

##### 3.1.1 Summary

Capital vintages dis-aggregate a production activity into two variants: one working with new capital generated in the same period from investments and an old one working with capital generated in earlier period. The equations in the model relating to the capital vintages are described in the section “Production block”.

##### 3.1.2 Initialization

The initialization of the capital vintages is performed in ‘*cal/cal\_vint.gms*’. Based on the user selection on the GUI (see section below), the set  $aVint(a)$  comprises the data base activities dis-aggregated into an old and new vintage, showns on the sets  $aOld$  and  $aNew$  and their union set  $aVints$ . Activities

outputting products in the SPE and Melitz model extensions are removed from the set *aVint* as defined selection by the user, warnings are added to the listing as shown below for the case of the Melitz extension:

```
$Sifthen.mel defined iMel
set remVintA(a);
remVintA(aVint) $ sum((rnat,a_vint(aVint,aVint),iMel),sam0(rnat,aVint,iMel)) = YES;

if ( card(remVintA),
    $Substinclude 'util/final_log' "Warning in %system.fn%, line: %system.incline%: using vintage activities for Melitz products not possible. These activities are set back to standard "
    display "Warning in %system.fn%, line: %system.incline%: using vintage activities for Melitz products not possible. . Such activities are removed ",remVintA;
    vintageA_sel(remVintA) = no;
);
aVint(a) $ sum((rnat,iMel),sam0(rnat,a,iMel)) = no;
```

The remaining data base activities *aVint* are dropped from the SAM and from other data referring to production activities based on the set *aDrop*:

```
*
*
* --- list of activities to drop from SAM and other data parameters
*
* aDrop(aVint) = yes;
*
```

While all other activities (not a new or old capital vintages or an aggregate thereof) are kept according to the set *aKeep*:

```
*
* --- list of activities to keep in SAM and other data parameters
*
* aKeep(a) = yes;
* aKeep(aDrop) = no;
* aKeep(aNew) = no;
* aKeep(aOld) = no;
*
```

The distribution of the data to the old and new variants at the benchmark is based on the default depreciation rate of 4% per annum times the number of depreciation years. This number is equal to unity for recursive-dynamic analysis and chosen by the user for comparative-static one. The distribution shares are set accordingly:

```
--- distribution shares for old and new vintage

p_distrVint(rs,aNew) = p_depr(rs,"%t0%")+0.04*(p_nDeprYears("%t0%")-1);

p_distrVint(rs,aOld) = 1-sum(a_avint(a,aNew) $ a_aVint(a,aOld),p_distrVint(rs,aNew));
```

These shares distribute the given SAM (and other data) to the old and new vintages while the original activities are removed, as shown for the case of the SAM:

```
--- distributes outputs to old and new vintages with distribution keys

SAM0(rs,aVints,is) = sum(a_aVint(a,aVints),SAM0(rs,a,is)*p_distrVint(rs,aVints));
SAM0(rs,aDrop,is) = 0;
```

Similar statements distribute the data referring to the domestic and imported demand, for factor taxes and subsidies and for production taxes. New and old capital are treated differently: old vintages receive all capital minus the new capital allocated to the new vintage. Activities not split up into the old and new vintage receive both old and new capital and a CES nests that allows for substitution between the two.

The substitution elasticities for the new vintage are set 2.5 higher higher compared to the parameters in a model without the vintage extension for technology nests comprising old or new capital as a production factor. The substitution elasticities for the old vintages are downscaled accordingly to keep the average reaction of the model to changes in price relations:

:

```

--- downscale subst. elas of old vintage such that in average over the two vintages, the given subst.
    elasticity is maintained

p_sigmaNest(rs,tNest,aOld) $ sum(tNest_f_a(tNest, cap, aOld), 1)
    = p_sigmaNest(rs,tNest,aOld) * (1 - (1-p_distrVint(rs,aOld)) * p_sigmaMult(rs,aOld));

p_sigmaNest(rs,tNest,aNew) $ sum(tNest_f_a(tNest, cap, aNew), 1)
    = p_sigmaNest(rs,tNest,aNew) * p_sigmaMult(rs,aNew);

parameter p_sigmaMult(r,a);
tNest("va") = YES;
tNest("nd") = YES;

--- multiplicative increase in subst. elasticity for nests comprising capital for new vintage is 2.5

p_sigmaMult(rs,aVints) $ xpFlag(rs,aVints) = 2.5;

--- downscale subst. elas of old vintage such that in average over the two vintages, the given subst.
    elasticity is maintained

p_sigmaNest(rs,tNest,aOld) $ sum(tNest_f_a(tNest, cap, aOld), 1)
    = p_sigmaNest(rs,tNest,aOld) * (1 - (1-p_distrVint(rs,aOld)) * p_sigmaMult(rs,aOld));
p_sigmaNest(rs,tNest,aNew) $ sum(tNest_f_a(tNest, cap, aNew), 1)
    = p_sigmaNest(rs,tNest,aNew) * p_sigmaMult(rs,aNew);

```

### 3.1.3 Dynamic updating of share parameters

In recursive-dynamic mode the code considers that a part of the old vintage is removed in each period, due to depreciation while last year's new vintage becomes part of the old vintage. For instance, expected output (stored temporarily on the field .m) is defined as:

```

xp.m(rs,aOld,tSim) $ xp.l(rs,aOld,tSim-1)
    = sum(aVint_aVint(aNew,aOld), xp.l(rs,aNew,tSim-1) * p_nDeprYears(tSim-1))
      + xp.l(rs,aOld,tSim-1)
      * sum(r_r(rs,rsNat), (1 - p_depr(rsNat,tSim-1) * p_nDeprYears(tSim-1)));

```

Similar statements define the expected demand for factors, intermediates and CES nests in the production function. In parallel, the code defines expected prices for the old vintage in the next period as weighted average between last period's old and new vintages, considering as above depreciation. For instance, the unit cost average for the new vintages is defined as:

```

px.m(rs,aOld,tSim) $ xp.m(rs,aOld,tSim)
    = [sum(aVint_aVint(aNew,aOld), xp.l(rs,aNew,tSim-1) * px.l(rs,aNew,tSim-1) * p_nDeprYears(tSim-1))
      + xp.l(rs,aOld,tSim-1) * px.l(rs,aOld,tSim-1)
      * sum(r_r(rs,rsNat), (1 - p_depr(rsNat,tSim-1) * p_nDeprYears(tSim-1)))] / xp.m(rs,aOld,tSim);

```

Based on these update demands for production factors and their prices, the share parameters are updated for the old vintage are updated, for instance for input demands:

```

*
* --- redefine share parameters, such that the old vintage would demand at prices and shifters of tSim-1 the
* aggregate quantities of the new and old vintages
*
p_ioNest(rs,i,aOld,iotNest,tSim) $ (xa.m(rs,i,aOld,tSim) $ sum(tNest_i_a(iotNest,i,aOld), p_ioNest(rs,i,aOld,iotNest,tSim) $ (xtNest.m(rs,iotNest,aOld,tSim) gt 0)))
    = sum(tNest_i_a(iotNest,i,aOld), (xa.m(rs,i,aOld,tSim) * tNestShare(rs,iotNest,i,aOld) / xtNest.m(rs,iotNest,aOld,tSim))
      * 1/lambdaio(rs,i,aOld,tSim-1) * ((p_sigmaNest(rs,iotNest,aOld)-1) $ (not addCes(rs,iotNest,aOld))
      - p_sigmaNest(rs,iotNest,aOld) $ addCes(rs,iotNest,aOld)))
      * (pa.m(rs,i,aOld,tSim) / ptNest.m(rs,iotNest,aOld,tSim)) * p_sigmaNest(rs,iotNest,aOld));

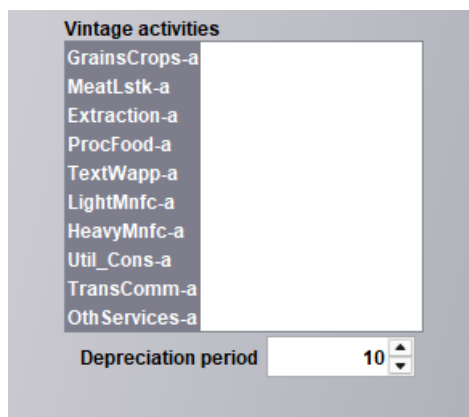
```

This implies that the technology of the old vintage changes over time. Fast growing sectors with an above-average share of the new vintages will show larger changes in their average technologies.



### 3.1.4 GUI

The GUI allows the user which production activities are broken down two vintages and to set the depreciation period in case of a comparative-static analysis:



## 3.2 myGTAP module

### 3.2.1 Summary

The *myGTAP* module allows representing factor earnings, income taxes paid, consumption and saving decisions and additional variables such as inter-household transfers for one or multiple private households in the model, replacing the regional household approach of the GTAP standard model. This also implies a separate government account, on demand with a government bond market. This also implies a separate government account, on demand with a government bond market. The user supplies the necessary data such the list of households and their specific factor stock shares via an additional file in a specific format as detailed below; they are not part of the standard GTAP data base. The additional data are supplied during the benchmark step and during the construction of the data base. This implies additional computational load during simulation but allows to re-define the model structure without the need to re-build the data base. This is especially valuable for structural sensitivity where the *myGTAP* module is switched on or off or different configuration of the *myGTAP* module are compared. The module is not thought as a replica of the *myGTAP* model by Minor and Walmsley 2013<sup>15</sup>.

The GUI allows selecting such a file along with closures rules related to different variables found in the module.

### 3.2.2 Motivation

Most CGEs application at single country level will make use of a household survey to provide distributional impact analyses beyond economic totals. The regional household approach of the standard GTAP model is not suitable for that type of analysis as it allocates the economy-wide sum of factor and tax income based on modified CD function to the different final demand categories. This decouples revenues and expenditures of the accounts for the individual agents by introducing a central utility maximizer which decides about how much income is consumed by these agents. Accordingly, the *myGTAP* extension (Minor and Walmsley 2013) replaces the regional household of the standard

---

<sup>15</sup> Minor, Peter and Walmsley, Terrie (2013): MyGTAP: A Program for Customizing and Extending the GTAP Database for Multiple Households, Split Factors, Remittances, Foreign Aid and Transfers, GTAP Working Paper No. 79, [https://www.gtap.agecon.purdue.edu/resources/res\\_display.asp?RecordID=4321](https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=4321)



GTAP model by separated accounts as usually found in the single country CGE models. We provide in here a module which features of *myGTAP* to ease the integration of key data from household surveys in the CGEBox framework. The module also covers features such as remittances, foreign aids, transfers and government debt dynamics.

As in general with the CGEBox framework, the *myGTAP* extension can be combined in a modular fashion with the different other extensions of CGEBox such as flexible nestings or the Melitz model implementation. Furthermore, the framework can be applied as a single country model or a global one, in comparative –static or recursive–dynamic fashion. One key aspect of the *myGTAP* implementation is the possibility to provide detailed household data for one or some countries only while using an aggregate private household for the remaining ones, an approach also chosen for the GEMPACK implementation.

### 3.2.3 Implementation

#### Model equations

The government account as well the potentially multiple private households employ each a CD function to decide about the consumption and savings share in their income. Alternatively, utility can be fixed and saving rates adjust endogenously – a closure often found for the government account. The income for the government account encompasses all tax flows (including transfers to/from households and direct taxes), plus foreign aid exchanges.

The income of the different private households is defined as factor income minus direct taxes, remittance and foreign capital income exchanges, net transfers between households and from government. All international transfers are in international currency.

In the *myGTAP* module, each private household  $h$  supplies its own per capita factor stocks  $x_{fth}$  to generate factor income  $f_{actyH}$  net of depreciation. Total household income  $h_{inc}$ , see equation (1), may include beside the factor income  $f_{actyH}$  additionally transfers from government  $transg$ , remittances  $remih$ , foreign aid to households  $fyiH$ , transfers from other households  $transh$  in the same region, and considers paying direct taxes  $dirTax$ , remittances  $remoH$ , foreign aid  $fyoH$  and transfers to other households  $trnh$ , as well as a share  $p_{shrDep}$  of debt servings  $p_{captrans}$ :

```

hInc(rs(r),h,ts(t)) $ (hFlag(r,h) $ (Hinc.range(r,h,t) ne 0)) ..
hInc(r,h,t)s(hinc,'r,h,t') =E=[
    --- factor income (defined above)
    factYH(r,h,t) $ (factYH.range(r,h,t) ne 0)
    --- debt servings in G-DEM linked to global bank, distributed with fixed share
    - sum(r_r(r,Rnat),p_capTrans(rNat,t)*pnum(t))*p_shrDep(rs,h,t)
    --- private adaptive investments
    - sum(r_r(r,Rnat),p_prvAdaptInv(rNat,t))*p_shrDep(rs,h,t)
    --- government transfers (normally fixed in real terms), multiplied with gov cons. index
    + trng(r,h,t)
    --- remittances received, converted from international to local currency
    + remiH(r,h,t)*sum(r_r(r,Rnat),lcu(rNat,t))
    --- share on foreign captial income (currently not use in default model)
    + fyiH(r,h,t)*fyiHShrCor(t)*sum(r_r(r,Rnat),lcu(rNat,t))
    --- private transfers received (normally fixed in real terms)
    update with consumer price index of sending household
    + sum(h0, trnh(r,h0,h,t)*pcons(r,h0,t))

```

(1)

```

--- direct taxes paid (as defined above)
$$iftheni.hSubr %hSubr%==on
  -sum(f $ (sum(a, p_gfh(r,h,f,a,t)) + sum(fNest,p_gfh(r,h,fNest,t)) + p_gfh(r,h,f,"natPool",t)),dirTax(r,f,h,t))
$$else.hSubr
  -sum(f $ sum(r_r(rp,r), sum(a, p_gfh(rp,h,f,a,t)) + sum(fNest,p_gfh(rp,h,f,fNest,t)) + p_gfh(rp,h,f,"natPool",t)),dirTax(r,f,h,t))
$$endif.hSubr

--- remittances sent, converted from international currency
-remoH(r,h,t)*sum(r_r(r,rNat),lcu(rNat,t))

--- foreign capital income sent, converted from international currency
- fyoH(r,h,t)*sum(r_r(r,rNat),lcu(rNat,t))

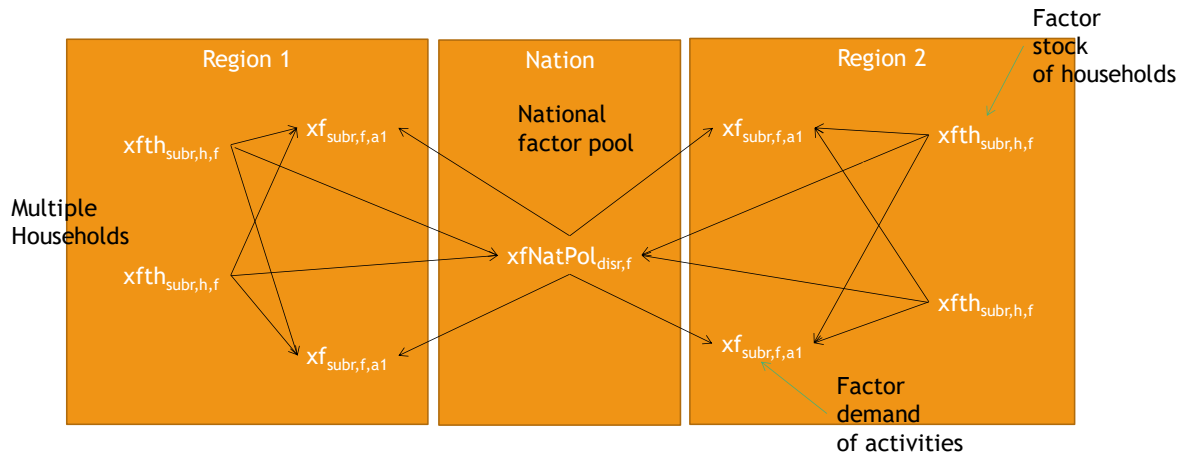
- sum(r_r(r,rNat) $ (factYH.range(r,h,t) ne 0),factYCorr(rNat,t)
  * sum(sameas(t,tRun),factyh.l(r,h,tRun-1) + factyh.l(r,h,t) $ sameas("%t0%",t)))
/sum(rp,h0) $ (r_r(rp,rNat) $ hFlag(rp,h0) $ (factYH.range(r,h0,t) ne 0)),
  sum(sameas(t,tRun),factyh.l(r,h0,tRun-1) + factyh.l(r,h0,t) $ sameas("%t0%",t))))

--- private transfer sent (normally fixed in real terms), at own consumer price index
- sum(h0, trnh(r,h,h0,t)*pcons(r,h,t))
+ p_payPastGovBonds(r,h,t)
]s(hinc,'r,h,t');

```

The elements *trng*, *trnh*, *remih*, *remoh*, *fyoH*, *fyiH* require additional data not available in the GTAP Data Base. They are therefore set to zero if not explicitly introduced by the user. However, entries for *trng* might be generated during benchmarking if data processing leads to negative household incomes or other undesired outcomes. Otherwise, they also default to zero.

In order to ensure homogeneity, some of these flows are multiplied with price indices, such as *pcons* for household transfers. To consider the case of endogenous exchange rates *lcu*, all exchanges with the rest-of-the-world are converted into the international currency.



**Figure 4: Overview on factor supply in myGTAP module with sub-national households**

The myGTAP module also allows to link households to subsistence production in a sub-region as depicted in the graphic about. In that case, besides the direct allocation of household resources to production activities in the same region, it can allocate resources to a national pool from which they are distributed to the activities in the different regions. Specifically, the factor stock  $xfh$  of a household  $h$  in sub-region  $subr$  is distributed based on a CET function to production activities in the same region (first line in equation (2), the relation is depicted by the left and right block in Figure 4 and to a national pool (second line in equation (2) and middle part of in Figure 4). That leads to the following composite price index  $pftH$  for the average returns to a factor  $f$  supplied by household  $h$ :

```

*
* --- average price received by household for supplying a factor to regional activities and the pool
*
* pftH(rs(r),h,f,ts(t)) $ ((sum(a, p_gfh(r,h,f,a,t)) + sum(fNest $ p_gfh(r,h,f,fNest,t), p_gfh(r,h,f,fNest,t)) + p_gfh(r,h,f,"natPool",t))) $ pftH.range(r,h,f,t) ..
*
* pftH(r,h,f,t)
* =E= (
*
* --- case with zero transformation elasticity, i.e. fixed shares
*
* sum(a $ p_gfh(r,h,f,a,t), p_gfh(r,h,f,a,t) * pf(r,f,a,t))
* + sum(fNest $ p_gfh(r,h,f,fNest,t), p_gfh(r,h,f,fNest,t) * pfNest(r,f,fNest,t))
* + p_gfh(r,h,f,"natPool",t) * sum(r_r(r,rNat),pftNatPool(rNat,f,t))) $ ( (p_omegafH(r,f) eq eps) $ p_omegafH(r,f))
*
* --- sluggish supply via CET
*
* + [ ( sum(a $ p_gfh(r,h,f,a,t), p_gfh(r,h,f,a,t) * pf(r,f,a,t)**(1+p_omegafH(r,f)))
* + sum(fNest $ p_gfh(r,h,f,fNest,t), p_gfh(r,h,f,fNest,t) * pfNest(r,f,fNest,t)**(1+p_omegafH(r,f)))
* + p_gfh(r,h,f,"natPool",t) * sum(r_r(r,rNat),pftNatPool(rNat,f,t)**(1+p_omegafH(r,f))) )
* * sum(r_r(r,rNat),pftNatPool(rNat,f,t))**((1+p_omegafH(r,f)) / (1+p_omegafH(r,f))) ]
* $ ( (p_omegafH(r,f) gt eps) and (p_omegafH(r,f) ne inf))
*
* --- Land under AEE
*
* + pft(r,f,t) $ ( (not p_omegafH(r,f)) or (p_omegafH(r,f) eq inf));

```

(2)

Where  $p\_gfh$  are the CET distribution parameters,  $pf$  are factor prices received for factor  $f$  employed in activity  $a$ , and  $pftNatPool$  is the price of the national factor pool. The cross-set  $r\_r$  depicts the relation between region  $r$  and the nation  $rNat$  it belongs to. If no sub-regions are present, it comprises a diagonal relation between the nations themselves.

Note: The \$ not  $p\_omegafH$  condition is used in conjunction with GTAP-AEZ. In that case, the average price of land in that region  $pft$  is paid to the households and fixed shares are used.

The par capita factor stock  $xftH$  can be updated by factor supply activities  $p\_etaf$  or changes in the xft, for instance by the GTAP-AEZ module:

```

*
* -----household stocks per capita might follow region wide price dependent change
*
* in total factor stock (equation only active if the stock is not fixed)
*
* e_xftH(rs(r),h,f,ts(t)) $ ( (xftH.range(r,h,f,t) ne 0) $ p_xftHBase(r,h,f) $ ((not sum(r_r(r,rNat),hFlag(dixr,h) $ (p_omegafH(dixr,f) ne 0)),1)) or dixr(r)) ) ..
*
* xftH(r,h,f,t)/p_xftHBase(r,h,f) =E=
*
* ---- AEE specific: the regional land base is endogenous and shifts the endowments of the households
*
* [ xft(r,f,t)/p_xftBase(r,f) ] $ p_xftBase(r,f)
*
* ---- Other cases: base might be depending on factor supply elasticity
*
* + [ ((pftH(r,h,f,t)/sum(r_r(rs,rNat),pgdmp(rsNat,ts)))*p_etaf(r,f)) $ p_etaf(r,f) + 1 $ (not p_etaf(r,f)) ] $ (not p_xftBase(r,f));

```

The factor demand of the activities  $xf$  is hence sourced by households in the same region  $r$  and a national pool of nation  $rNat$ :

(3)

```

xftReq(rs(r),f,a,ts(t)) $ (xftFlag(rs,f,a) $ p_omegafH(r,f) $ (pf.range(r,f,a,t) ne 0) $ (not sameas(f,"oldCap")))
$ sum(tNest_f_a(tNest,f,a), p_afNest(r,f,a,tNest,t))
$ (not sum((h,fNest_a_f(fNest,a,f)),p_gfh(r,h,f,fNest,t))) ) ..

[sum(aVint_aVint(a,a0) $ xftFlag(rs,f,a0),xft(r,f,a0,t))s(xf,'r,f,a,t')] $ ((p_omegafH(rs,f) ne inf) and (not aNew(a)))
+ (pf(r,f,a,t)s(pf,'r,f,a,t')) $ ((p_omegafH(rs,f) eq inf) or aNew(a))
=
[
---- households supplying factors to the regional activities in which it is located
(the Leontief case with fixed shares is written separately)

sum(h $ p_gfh(r,h,f,a,t),
xftH(r,h,f,t)s(xf,'r,f,a,t')*pop(r,h,t)/1000 * sum(aVint_aVint(a,a0),p_gfh(r,h,f,a0,t))* (1+p_xfthShift(r,h,f,t))
* (pf(r,f,a,t)/pftH(r,h,f,t))*p_omegafH(r,f)) $ (p_omegafH(r,f) ne eps)

--- this is a copy from the core model
+ ( sum(aVint_aVint(a,a0),p_gf(rs,f,a0,ts))s(xf,'r,f,a,t')
* ( ncpVusIn[ (pf(rs,f,a,ts)/sum(r_r(rs,rNat),pgdpmp(rsNat,ts)))*5,
(pf(rs,f,a,ts)/sum(r_r(rs,rNat),pgdpmp(rsNat,ts)))*p_etaff(rs,f,a),1.E-6]
$ (p_etaff(rs,f,a) and (not sameas(f,"oldCap")))
+ 1 $ (not p_etaff(rs,f,a))
) $ (p_omegafH(r,f) eq eps)

---- the national pool supply to the regional activities

+ sum(r_r(rs,rNat) $ (p_gfNatPool(rs,f,a,t) $ (p_omegafNatPool(rNat,f) eq eps)),
xftNatPool(rNat,f,t)s(xf,'r,f,a,t') * sum(aVint_aVint(a,a0),p_gfNatPool(rs,f,a0,t)))

+ sum(r_r(rs,rNat) $ p_gfNatPool(r,f,a,t),
xftNatPool(rNat,f,t)s(xf,'r,f,a,t') * sum(aVint_aVint(a,a0),p_gfNatPool(rs,f,a0,t))
* (pf(rs,f,a,t)/pftNatPool(rNat,f,t))*p_omegafNatPool(rNat,f)) $ sum(r_r(rs,rNat) $ (p_omegafNatPool(rNat,f) ne eps)

] $ ((p_omegafH(rs,f) ne inf) and (not aNew(a)))

+ sum(aVint_aVint(a,a0),pf(rs,f,a0,ts)s(pf,'r,f,a,t')) $ ((p_omegafH(rs,f) ne inf) and aNew(a))

+ (pft(r,f,t)s(pf,'r,f,a,t')) $ (p_omegafH(rs,f) eq inf);

```

The first expression in (3) above distributes a share of the per-capita factor stock  $xftH$  of the different households  $h$  in the same region to the activity  $a$ , depending on the relation between factor returns of that activity  $pf$  and the average returns at household level  $pftH$ , as defined above in equation (2). The second expression distributes a share of the endogenous national pool  $xftNatPool$ , using the relation to the average pool price  $pftNatPool$ . The third expression relates to the supply of immobile factor based on a factor supply elasticity  $p_etaff$  and is a copy from the main model's equation. The two remaining lines refer to the capital vintage mechanism where factor prices (besides capital) are assumed as equal and the case of infinite factor supply where all households face the same factor prices.

As households supply a part of their factor stocks to a national pool (think about capital or migrant labor), the national pool  $xftNatPool$  is sourced by all households in this nation. That is captured by summing over both  $r_r(r,rNat)$  to cover all regions  $r$  belonging to the nation  $rNat$  and over households  $h$  in equation (4) below. The share of the households' factor stocks supplied to the national factor pool depends on the relation between the pool price  $pftNatPool$  and the average returns received by each household  $pftH$ :

```

*
* --- National pool (e.g. migrant labour): sourced by stocks of the different households (in different regions)
*
* Note: the factor stock of the household xftH are expressed at per capita basis (therefore the multiplication with pop)
*
e_xftNatPool(rsNat(rNat),f,ts(t)) $ (sum( (r_r(r,rNat),h) $ p_omegafH(r,f), p_gfh(r,h,f,"natPool",t))) ..
xftNatPool(rNat,f,t)s(xftNatPool,'rNat,f,t')
= sum( (r_r(r,rNat),h) $ (p_gfh(r,h,f,"natPool",t) $ p_omegafH(r,f)),
[ xftH(r,h,f,t)s(xftNatPool,'rNat,f,t') * pop(r,h,t)/1000*(1+p_xfthShift(r,h,f,t)) * p_gfh(r,h,f,"natPool",t) ] $ (p_omegafH(r,f) eq eps)
+ [ xftH(r,h,f,t)s(xftNatPool,'rNat,f,t') * pop(r,h,t)/1000*(1+p_xfthShift(r,h,f,t))
* p_gfh(r,h,f,"natPool",t) * (pftNatPool(rNat,f,t)/pftH(r,h,f,t))*p_omegafH(r,f) ] $ (p_omegafH(r,f) ne eps));

```

(4)

The average pool price  $pftNatPool$  is defined from the factor returns  $pf$  of the activities and the share parameters  $p_gfNatPool$ :

```

* --- average price received for supplying primary factor from the national pool to the activities
* in the subnational regions
*
e_pftNatPool(rsNat(rNat),f,ts(t)) $ (sum( (r_r(rNat),h) $ p_omegafH(r,f),p_gfh(r,h,f,"natPool",t)) ) ..
pftNatPool(rNat,f,t)
=E= sum( (r_r(rNat),a) $ p_gfNatPool(r,f,a,t), p_gfNatPool(r,f,a,t)*pf(r,f,a,t)) $ (p_omegafNatPool(rNat,f) eq eps)
+ [sum( (r_r(rNat),a) $ p_gfNatPool(r,f,a,t),
p_gfNatPool(r,f,a,t)
* pf(r,f,a,t)**(1 + p_omegafNatPool(rNat,f)) ) ** (1/(1+p_omegafNatPool(rNat,f)))) $ (p_omegafNatPool(rNat,f) ne eps);

```

(5)

Note the different share parameters used:  $p\_gfh$  refers to the shares supplied by households,  $p\_gfNatPool$  to shares on the national pool.

Demand for factor supply nests is defined in equation  $e\_xfNestH$  below, using entries in the share parameter  $p\_gfh$  relation to factor nests.

```

*
* --- Regional use factor use of each activity
*
e_xfNestH(rs(r),f,fNest,ts(t)) $ (sum(h,p_gfh(r,h,f,fNest,t)) $ p_omegafH(r,f)
$ (xfNest.range(r,f,fNest,t) ne 0)) ..

xfNest(r,f,fNest,t) s(xfNest,'r,f,fNest,t')

=E=
[
sum(h $ p_gfh(r,h,f,fNest,t),
xftH(r,h,f,t)*pop(r,h,t)/1000*(1+p_xfthShift(r,h,f,t)) * p_gfh(r,h,f,fNest,t))
$ (p_omegafH(r,f) eq eps)

+
sum(h $ p_gfh(r,h,f,fNest,t),
xftH(r,h,f,t)*pop(r,h,t)/1000*(1+p_xfthShift(r,h,f,t)) * p_gfh(r,h,f,fNest,t)
* (pfNest(r,f,fNest,t)/pftH(r,h,f,t)) **p_omegafH(r,f)
$ (p_omegafH(r,f) ne eps)

] s(xfNest,'r,f,fNest,t');

```

The factor income reflects stocks  $xfth$  and related return  $pfth$ , and subtracts household specific shares  $p\_shrDep$  on depreciation:

```

*
* --- Factor income by households, from given (or elasticity/AEZ driven) stock and price determined above
*
e_factYH(rs(r),h,ts(t)) $ (hFlag(r,h) $ (factYH.range(r,h,t) ne 0)) ..

factYH(r,h,t) s(factyh,'r,h,t')

=E= [
$iftheni.hSubr "%hSubr%"=="on"
*
* --- household at sub-national level
*
sum( f $ (sum( a, p_gfh(r,h,f,a,t)) + sum(fNest,p_gfh(r,h,f,fNest,t)) + p_gfh(r,h,f,"natPool",t)),
xftH(r,h,f,t)*pop(r,h,t)/1000 *pfth(r,h,f,t)*(1+p_xfthShift(r,h,f,t)))
$else.hSubr
*
* --- household at national level
*
sum( (r_r(rp,r),f) $ (sum(a, p_gfh(rp,h,f,a,t)) + sum(fNest, p_gfh(rp,h,f,fNest,t)) + p_gfh(rp,h,f,"natPool",t)),
xftH(rp,h,f,t)*pop(rp,h,t)/1000 *pfth(rp,h,f,t)*(1+p_xfthShift(r,h,f,t)))
$endif.hSubr
*
* --- distribution depreciation based on fixed shares shrDep
*
- sum(r_r(r,rNat),p_fdepr(rNat,t)*pi(rNat,t)*kstock(rNat,t))*p_shrDep(r,h,t) ] s(factyh,'r,h,t');

```

(6)

The link to the remaining equations of the standard GTAP model is provided firstly by defining average direct tax rates:

```

*
* --- Weighted national average of household specific direct tax rates
*
e_kappaf(rsNat(rNat),f,ts(t)) $ sum(a, xfflag(rNat,f,a)) ..

(kappaf(rNat,f,t)+kappaShft(rNat,t)) * sum(a $ xfflag(rNat,f,a), pf(rNat,f,a,t)*xf(rNat,f,a,t)) /regy.scale(rNat,t)
=E=
[
$iftheni.hSubr "%hSubr%"=="on"
sum( (r_r(rs,rNat),h) $ (hFlag(rs,h) $ (sum( a, p_gfh(rs,h,f,a,t)) + p_gfh(rs,h,f,"natPool",t))),dirTax(rs,f,h,t))
$else.hSubr
sum(h $ (hFlag(rNat,h) $ sum( r_r(rp,rNat), sum(a, p_gfh(rp,h,f,a,t)) + p_gfh(rp,h,f,"natPool",t))),dirTax(rNat,f,h,t))
$endif.hSubr
] / regy.scale(rNat,t);

```

(7)

These enter the usual tax income equations which hence need not to reflect the different households and related factor ownership shares. Secondly, adding up from savings of the households and governments defines the regional savings used in the standard GTAP model:

```
*
* --- regional savings, sum of savings by households and saving by government
*
* e rsav(rsNat(rNat),ts(t)) ..
  rsav(rNat,t)/rsav.scale(rNat,t) =e= (sum( (r_r(rs,rNat),h) $ hFlag(rs,h),rsavC(rs,h,t))+rsavG(rNat,t))/rsav.scale(rNat,t);
```

(8)

Equally, total factor income is defined:

```
*
* --- aggregated factor income
*
* e facty(rsNat(rNat),ts(t)) ..
  facty(rNat,t)/facty.scale(rNat,t) =e= (sum( (r_r(rs,rNat),h) $ hFlag(rs,h), factyH(rs,h,t)))/facty.scale(rNat,t);
```

(9)

Remittance outflows from the households are defined as a share  $p\_remiitShr$  of labor income; the shares are calculated from GTAP-MIG data base if that is available:

```
*
* --- remittance outflows in local currency (global value*lcu) as a share of labor income
*
* e remoh(rs(r),h,ts(t)) $ (hFlag(r,h) $ ((remohShr.range(r,h,t) ne 0) or (remohShr.up(r,h,t) ne 0)) $ (remoh.range(r,h,t) ne 0)) ..
  remoh(r,h,t)*sum(r_r(r,rNat),lcu(rNat,t))/remoh.scale(r,h,t)
  =E=
*
* ---- given share of the household of all regional remittances sent
*
  remohShr(r,h,t)/remoh.scale(r,h,t)
  $iftheni.gmig "%modulesGMIG%"=="on"
*
* --- the GMIG module determines bo-lateral remittances v_remit
*
  * sum((r_r(r,rNat),rNat1) $ sum(l,p_remit(rnat1,rnat,1)), v_remit(rnat1,rnat,t))
  $else.gmig
*
* --- GIG is not active: we distribute a fixed share of the regional labour income
*
  * sum((r_r(r,rNat),1,a) $ xfFlag(r,1,a), pf(r,1,a,t)*xf(r,1,a,t) * p_remitShr(rNat,"shr",1))
  $endif.gmig
;
```

(10)

Remittance inflows use the same data base to define a bi-lateral link:

```
*
* --- remittance inflows in global currency are a share of the global inflows
*
* e remih(rs(r),h,ts(t)) $ (hFlag(r,h) $ (remih.range(r,h,t) ne 0)) ..
  remih(r,h,t)/remih.scale(r,h,t)
  =E=
*
* ---- given share of the household of all regional remittances received
*
  remihShr(r,h,t)/remih.scale(r,h,t)
  $iftheni.gmig "%modulesGMIG%"=="on"
*
* --- the GMIG module determines bo-lateral remittances v_remit
*
  * sum((r_r(r,rNat),rNat1) $ sum(l,p_remit(rnat,rnat1,1)), v_remit(rnat,rnat1,t))
  $else.gmig
*
* --- GIG is not active: we distribute a fixed share of all nations' rnat labour income to the country rNat1
  in which the household is located
*
  * sum((rNat,1,a) $ xfFlag(rNat,1,a), pf(rNat,1,a,t)*xf(rNat,1,a,t)*sum(r_r(r,rNat1),p_remitShr(rNat1,rNat,1)))
  $endif.gmig
;
```

(11)

Similarly, foreign capital income transfers are defined as share of capital income and a share of total global flows:

```
*
* --- foreign global capital inflows in global currency are a share of the global inflows
*
* e_fyih(rs(r),h,ts(t)) $ (hFlag(r,h) $ (fyih.range(r,h,t) ne 0)) ..
  fyih(r,h,t)/fyih.scale(r,h,t)
  =E= fyihShr(r,h,t) * sum( (rp,h0) $ hFlag(r,h0), fyoH(rp,h0,t))/fyih.scale(r,h,t);
```

(12)

And

```

*
* --- foreign global capital outflows in local currency (fyoh*lcu) as share of capital income
*
e_fyoh(rs(r),h,ts(t)) $ (hFlag(r,h) $ (fyoh.range(r,h,t) ne 0)) ..
    fyoh(r,h,t)*sum(r_r(r,rNat),lcu(rNat,t))/fyoh.scale(r,h,t)
    =E= fyohShr(r,h,t)
        * sum((sameas(f,cap),a ) $ xfFlag(r,f,a), pf(r,f,a,t)*xf(r,f,a,t)*p_shrXf(r,h,f,a,t))/fyoh.scale(r,h,t);

```

(13)

Again, in case that these flows are fixed, shares adjust and a correction factor is introduced to main the global balance:

```

*
* --- ensure global balance, only active of fyih is fixed
*
e_fyihbal(ts(t)) ..
    sum( (r,h) $ hFlag(r,h), fyih(r,h,t))*fyihShrCor(t)/sum(r, regy.scale(r,t))
    =E= sum( (r,h) $ hFlag(r,h), fyoh(r,h,t))/sum(r, regy.scale(r,t));

```

(14)

Foreign aid paid by the government is defined as a share of factor income:

```

*
* --- foreign transfer in local currency as share of factor income
*
e_aido(rsNat,ts(t)) $ (aido.range(rsNat,t) ne 0) ..
    aido(rsNat,t)*lcu(rsNat,t)/regy.scale(rsNat,t) =E= aidoShr(rsNat,t) * facty(rsNat,t)/regy.scale(rsNat,t);

```

(15)

Inflowing aid is defined as a share of global outflows:

```

*
* --- share of incoming foreign aid in total global aid
*
e_aidi(rsNat,ts(t)) $ (aidi.range(rsNat,t) ne 0) ..
    aidi(rsNat,t)/regy.scale(rsNat,t) =E= aidiShr(rsNat,t) * sum(rNat, aido(rNat,t))/regy.scale(rsNat,t);

```

(16)

In case that income aid is fixed, shares are endogenous and a correction factor ensures global balancing:

```

*
* --- ensure that sum of aidi = sum of aido, only active if aidi is fixed
*
e_aidiBal(ts(t)) ..
    sum(rNat, aidi(rNat,t))*aidiShrCor(t)/sum(rNat, regy.scale(rNat,t))
    =E= sum(rNat, aido(rNat,t))/sum(rNat, regy.scale(rNat,t));

```

(17)

Government transfers to household *trng* can be endogenized using a distribution keys *p\_trng* and various drivers as shown below. The first driver for government expenditures, the second one population size, the third one changes in the consumer price indices, the fourth on changes in the wage index and finally, exogenous allocations to each household:



```

*
* --- definition transfers from government to household
*
e_trng(rs(r),h,ts(t)) $ (hFlag(r,h) $ (trng.range(r,h,t) ne 0)) ..
    trng(r,h,t)s(trng,'r,h,t') =E= trng.l(r,h,"%t0%")s(trng,'r,h,t')
*
*
* --- government expend.
*
* (1 + p_trng(r,"pg",t) * [sum(r_r(r,rNat), pg(rNat,t))-1])
*
* --- pop
*
* (1 + p_trng(r,"pop",t))
*
* --- update based on change in CPI
*
* (1 + p_trng(r,"CPI",t) * [pCons(r,h,t)-1])
*
* --- update based on change in real wages
*
* (1 + p_trng(r,"wageIndex",t) * [sum(r_r(r,rNat), wageIndex(rNat,t))-1])
*
* --- additional fixed component
*
+ sum(sameas(h,trngDriver),p_trng(r,trngDriver,t))s(trng,'r,h,t');
-

```

(18)

The wage index is defined as the stock weighted average of all labor categories:

```

--- wage index in the economy
e_wageIndex(rsNat,ts(t)) ..
    wageIndex(rsNat,t) * sum(l, xft(rsNat,l,t))/sum(l, xft.l(rsNat,l,"%t0%")*pft.l(rsNat,l,"%t0%"))
    =E= sum(l, xft(rsNat,l,t)*pft(rsNat,l,t))/sum(l, xft.l(rsNat,l,"%t0%")*pft.l(rsNat,l,"%t0%"));

```

(19)

The transfer paid out or received are in sum equal to the tax flow “tg” as defined in the

That leads to the following two equations that distribute government income to consumption  $yg$  and savings  $rsavg$ :

```

--- government account: gov consumption
e_yg(rsNat(rNat),ts(t)) $ (yg.range(rNat,t) ne 0) ..
    yg(rNat,t)s(yg,'rnat,t') =E= [ ytaxTot(rNat,t)
    + aidi(rNat,t)*aidiShrCor(t)*lcu(rNat,t) - aido(rNat,t)*lcu(rNat,t)]*(1-betaSG(rnat,t))s(yg,'rnat,t');
--- government account: gov savings
e_rsavg(rsNat(rNat),ts(t)) $ (rsavg.range(rNat,t) ne 0) ..
    rsavg(rNat,t)s(rsavg,'rnat,t') =E= [ ytaxTot(rNat,t)
    + aidi(rNat,t)*aidiShrCor(t)*lcu(rNat,t) - aido(rNat,t)*lcu(rNat,t)]*betaSG(rnat,t)s(rsavg,'rnat,t');

```

(20)

As both the government and the private households have their own saving variables, we have to add up to link up to the equation structure of GTAP standard to arrive a physical savings  $xsavG$  at national level:

```

*
* --- Savings by government, quantity index
*
e_xsavG(rsNat(rNat),ts(t)) ..
    xsavG(rNat,t)s(yg,'rNat,t') * psave(rNat,t) =E= rsavG(rNat,t)s(yg,'rNat,t');

```

(21)

The value of regional savings  $rsav$  is defined as:

```

--- regional savings, sum of savings by households and saving by government
e_rsav(rsNat(rNat),ts(t)) ..
    rsav(rNat,t)s(rsav,'rnat,t') =E= (sum( (r_r(rs,rNat),h) $ hFlag(rs,h),rsavC(rs,h,t))+rsavG(rNat,t))s(rsav,'rnat,t');

```

(22)



Equally, we need to introduce the net income transfer with the rest-of-the-world finc which enters the regional income equation, defined as:

```
*
* --- Net income transfers with ROW in local currency, enters regional income equation
* (the equation does not comprise the debt servings of G-RDEM; that is a separate entry
* in the household's income definition)
*
e_finc(rNat,ts(t)) $ rs(rNat) ..
finc(rNat,t) =E= + [aidi(rNat,t) + sum( (r_r(rs,rNat),h) $ hFlag(rs,h), remiH(rs,h,t)) + sum(h $ hFlag(rNat,h), fyiH(rNat,h,t))
- aido(rNat,t) - sum( (r_r(rs,rNat),h) $ hFlag(rs,h), remoH(rs,h,t)) - sum(h $ hFlag(rNat,h), fyoh(rNat,h,t))] *lcu(rNat,t);
```

(23)

Again, shares are defined as variables to allow for closure swaps. All other equations are untouched.

The following three helper equations define for information only the expenditure shares of price, government and investments on regional income:

```
e_betag(rsNat(rnat),ts(t)) ..
betag(rNat,t) =E= yg(rNat,t)/regy(rNat,t);
e_betas(rsNat(rnat),ts(t)) ..
betas(rNat,t) =E= rsav(rNat,t)/regy(rNat,t);
e_betap(rsNat(rnat),ts(t)) ..
betap(rNat,t) =E= sum((r_r(rs,rNat),h) $ hFlag(rs,h), yc(rs,h,t))/regy(rNat,t);
```

### 3.2.4 Government debt dynamics

In the standard layout discussed above, the government's consumptions and saving are directly driven by its tax income as shown above. An alternative mechanism drives these variables by GDP and introduces government bonds to close the account. This approach is described in this section.

With this layout, it is assumed that the saving rates of household is sensitive to changes in the relation between the consumer price index  $pCons$  and (expected) returns to savings  $pFinh$ . Higher consumer prices decrease and higher returns to savings increase the saving rates based on a CET utility function with transformation elasticity  $p_{\omega}Exp$ . The average price for the household's utility is accordingly a weighted average of these two prices. Let  $p_{aExp}$  depict the expenditure share parameter attached to current consumption. The composite price index for region  $r$  and household  $h$  at time  $t$  for savings and consumption is defined as:

```
--- average expenditure price of households, CET weithed average of private consumption and savings save
e_pExp(rs(r),h,ts(t)) $ (hFlag(r,h) $ (betasH.1(r,h,"%t0%") gt 0)) ..
pExpH(r,h,t) =E= [ p_aExp(r,h,t) * (pCons(r,h,t)/sum(r_r(rs,rNat),pFact(rsNat,t)))**(1+p_omegaExp(r,h))
--- if expected returns to savings go up, invest more
+ (1-p_aExp(r,h,t)) * pFinh(r,h,t)**(1+p_omegaExp(r,h)) ]*(1/(1+p_omegaExp(r,h)));
```

The projected (or benchmark) saving rates  $betash$  of the household as stored on  $p_{betashShift}$  are accordingly sensitive to changes in returns in financial markets (returns to investments and government bonds) relative to the price index defined above.

```
--- the exogenously projected saving share reacts to changes in the returns in financial markets
e_betash(rs(r),h,ts(t)) $ (hFlag(r,h) $ (betasH.1(r,h,"%t0%") gt 0) $ (betash.range(r,h,ts) ne 0)) ..
betash(r,h,t) =E= p_betashShift(r,h,t)
* (1-p_aExp(r,h,t)) * ( pFinh(r,h,t)/pExpH(r,h,t))**p_omegaExp(r,h);
```

Households are assumed to maximize expected revenues from their savings based on a CET function. Total savings are distributed between investments which increase the future capital stock and thus generate capital returns and government bonds which draw a fixed yield. Accordingly, the expected returns to household savings are a weighted average of the expected rate of return to capital,  $r_{ore}^{16}$ , and yields of newly issued government bonds,  $govBondYld$ . Both are normalized with their benchmark values to arrive at a unity price  $pFinH$  at the benchmark:

```
*
* --- price index of expected return to capital and gov bond yields
* (CET revenue max)
*
e_pFinH(rs(r),h,ts(t)) $ (hFlag(r,h) $ (betasH.1(r,h,"%t0%") gt 0)) ..

pFinH(r,h,t) =E=
[
  p_aFinRore(r,h,t) * sum(r_r(r,rnat),rore(rnat,t)/rore.1(rnat,"%t0%"))**(1+p_omegaFinH(rs,h))
  + (1-p_aFinRore(r,h,t)) * sum(r_r(r,rnat),govBondYld(rnat,t)/govBondYld.1(rnat,"%t0%"))**(1+p_omegaFinH(rs,h))
]**(1/(1+p_omegaFinH(rs,h)));
```

The amount of government bonds bought  $newGovBond$  is hence determined by the savings of the households  $rSavc$ , benchmark shares into investment  $p_aFinRore$  and the interest paid by the government  $govBondYld$  relative to the price index above:

```
--- new government bonds are a share of the households savings, multiplied with the price relations
between yield and the average returns to savings

e_newGovBond(rsNat(rnat),ts(t)) $ (newGovBond.range(rnat,t) ne 0) ..

newGovBond(rnat,t)s(yTaxTot,'rnat,t')
=E= sum((r_r(rs,rnat),h) $ (hFlag(rs,h) $ p_aFinRore(rs,h,t) $ (betasH.1(rs,h,"%t0%") gt 0)),
  rsavC(rs,h,t) * (1-p_aFinRore(rs,h,t))
  * (govBondYld(rnat,t)/[pFinH(rs,h,t)*govBondYld.1(rnat,"%t0%")])**p_omegaFinH(rs,h))s(yTaxTot,'rnat,t');
```

Government consumption  $yg$  is driven partly by regional income  $regy$  and partly by total tax income  $yTaxTot$ , minus the costs of MAC  $valMaccs$  and minus debt servings  $p_payPastGovBonds$ . The share driven by the corrected tax income is the sum of variables  $spendShrByGovInc$  and  $govBondDamp$ :

```
--- government account: gov consumption as a share of GDP (benchmark relation)

e_yg(rsNat(rNat),ts(t)) $ (yg.range(rNat,t) ne 0) ..

yg(rNat,t)s(yg,'rnat,t')
=E= (1+betaSG(rnat,t))*yg.1(rnat,"%t0%")s(yg,'rnat,t')
  * [ regy(rnat,t)/regy.1(rnat,"%t0%")
    * (1 - spendShrByGovInc(rnat,t) + govBondDamp(rNat,t) $ (newGovBond.range(rnat,t) ne 0))
  + (yTaxTot(rnat,t)+aidi(rnat,t)
  $Sifi "%MACCsProcessEmis"=="on" - valMaccs(rnat,t)
  -sum((r_r(rNat,rs),h) $ hFlag(rs,h) $ p_payPastGovBonds(rs,h,t)))/yTaxTot.1(rnat,"%t0%")
  * (spendShrByGovInc(rnat,t) + govBondDamp(rNat,t) $ (newGovBond.range(rnat,t) ne 0))
  ];
```

A similar equation drives government savings  $rsavg$  (regional savings by government):

<sup>16</sup> The expected revenues are defined based on the variable  $arent$  as the returns to capital minus the ad-valorem direct tax rate on capital. This variable is then divided by the price index of investments and corrected for the depreciation rate to yield the variable  $rorc$  which is called the net rate of return to capital. From there, the expected rate of return to capital is defined by a dampening factor which reflect the end of year capital over the beginning of year capital stock.

```

*
* --- government account: gov savings as a share of GDP (benchmark relation)
*
e_rsavg(rsNat(rNat),ts(t)) $ ((rsavg.range(rNat,t) ne 0) $ (newGovBond.range(rnat,t) ne 0)) ..
    rsavg(rNat,t)s(rsavg,'rnat,t')
    =e= rsavg.l(rnat,"%t0%")s(rsavg,'rnat,t')

    * [ gdpmp(rnat,t)/gdpmp.l(rnat,"%t0%") * (1 - spendShrByGovInc(rnat,t))

*
* --- let government savings increase if it would otherwise run a surplus ...
*
    +((p_govBondFudge(rnat)-ncpVUSin(newGovBond(rnat,t)/yTaxTot(rNat,t),p_govBondFudge(rnat),1.E-5)] * 5000) $ (newGovBond.range(rnat,t) ne 0)

    + (yTaxTot(rnat,t)+aidi(rnat,t)
    $$ifi "%MACCsProcessEmis%"="on" - valMaccs(rnat,t)
    -sum((r_r(rNat,r),h) $ hFlag(r,h), p_payPastGovBonds(r,h,t))/yTaxTot.l(rnat,"%t0%")
    * spendShrByGovInc(rnat,t)]

*
* --- demand for MACC
*
    $$ifi "%MACCsProcessEmis%"="on" + valMaccs(rnat,t)s(rsavg,'rnat,t')
;

```

The share *spendShrByGovInc* which is driven by the government income corrected for payments to past debt depends on the debt servings relative to total government income minus debt servings:

```

*
* --- Share of government consumptions and saving driven by government income and not by GDP
*
e_spendShrByGovInc(rsNat(rNat),ts(t)) $ ((yg.range(rNat,t) ne 0) $ (newGovBond.range(rnat,t) ne 0)) ..
    spendShrByGovInc(rnat,t) =E= [sum((r_r(rNat,r),h) $ hFlag(r,h), p_payPastGovBonds(r,h,t))
    /(-ncpVUSin(-sum((r_r(rNat,r),h) $ hFlag(r,h), p_payPastGovBonds(r,h,t)),
    -[yTaxTot(rnat,t)*m_var2Scale(yTaxTot,rnat,t)
    $$ifi "%MACCsProcessEmis%"="on" - valMaccs(rnat,t)
    - sum((r_r(rNat,r),h) $ hFlag(r,h), p_payPastGovBonds(r,h,t))],1.E-3))]**(1/4);

```

The *ncpVusin* function is a smooth min approximation and ensure that the shares stay below unity even if the debt servings would exceed the total tax income minus the debt serving. Due to the functional form, if 50% of the governments tax income is spend on serving debt, consumption and savings are solely driven by the tax income minus the debt servings and decoupled from GDP growth. This would imply that no new bonds are issued and the debt servings would start to fall. The exponent (1/4) will imply a rather sensitive reaction of this fiscal stability mechanism if debt serving relative to tax income increases.

The share that is driven by tax income is updated endogenously depending on the sum of debt payments relative to total tax income minus the sum of debt payments and spendings on abatement cost:

```

--- Share of government consumptions and saving driven by government income and not by GDP
e_spendShrByGovInc(rsNat(rNat),ts(t)) $ ((yg.range(rNat,t) ne 0) $ (newGovBond.range(rnat,t) ne 0)) ..
    spendShrByGovInc(rnat,t) =E= -ncpVUSin(0,-[sum((r_r(rNat,rs),h) $ hFlag(rs,h), p_payPastGovBonds(rs,h,t))
    /(-ncpVUSin(-sum((r_r(rNat,rs),h) $ hFlag(rs,h), p_payPastGovBonds(rs,h,t))-1.E-6*yg.l(rNat,"%t0%"),
    -[yTaxTot(rnat,t)+aidi(rnat,t)
    $$ifi "%MACCsProcessEmis%"="on" - valMaccs(rnat,t)
    - sum((r_r(rNat,rs),h) $ hFlag(rs,h), p_payPastGovBonds(rs,h,t))],1.E-3))]**(1/4)
    + 1.E-3** (1/4));

```

The second variable *govBondDamp* defining government consumption and savings is driven by the bond yield and introduced to prevent that the bond yield can drop to zero. If this would happen, the average price of returns to saving *p\_finH* could behave unreasonable. The function *ncpVUSin* is a smoothed min expression. The parameters are chosen that such that a bond yield of 5%, *govBondDamp* is equal to zero, at 1% is increases to 0.018 and at 0.5% is increases to 0.5.

```

--- let government savings and consumption increase if it would otherwise run a surplus ...
e_govBondDamp(rsNat(rNat),ts(t)) $ (newGovBond.range(rnat,t) ne 0) ..
    govBondDamp(rNat,t) =E= 1-ncpVUSin(1,govBondYld(rNat,t)*100,1.E-1);

```

The yield of government bonds is implicitly defined by closing the government account with the newly issued bonds:

```

*
* --- government account: demand for new bounds, implicitly defines their price
*
e_govAccountYld(rNat(rNat),ts(t)) $ (newGovBond.range(rNat,t) ne 0) ..

0 =e= [ytaxTot(rNat,t)
      + aidi(rNat,t)*aidiShrCor(t)*lcu(rNat,t) - aido(rNat,t)*lcu(rNat,t)
      - yg(rNat,t)
      - rsavg(rNat,t)
      - p_govAdaptInv(rNat,t)
      - sum((r_r(rNat,r),h) $ hFlag(r,h), p_payPastGovBonds(r,h,t))
      + newGovBond(rNat,t)]s(yTaxTot,'rNat,t');

```

The same equation defines total government savings *rsavg* in case the amount of new government bonds is fixed:

```

e_govAccountSav(rNat(rNat),ts(t)) $ (newGovBond.range(rNat,t) eq 0) ..

0 =e= [ytaxTot(rNat,t)
      + aidi(rNat,t)*aidiShrCor(t)*lcu(rNat,t) - aido(rNat,t)*lcu(rNat,t)
      - yg(rNat,t)
      - rsavg(rNat,t)
      - p_govAdaptInv(rNat,t)
      - sum((r_r(rNat,r),h) $ hFlag(r,h), p_payPastGovBonds(r,h,t))
      + newGovBond(rNat,t)]s(yTaxTot,'rNat,t');

```

Other elements in the government account are net transfers to households *trng* and aid received from other governments *aidi* and paid out to other governments *aido*. These elements can also be zero depending on the underlying data (or assumptions).

Regional savings *rsav* as one component of savings besides depreciation and foreign savings in (I=S) are defined as the sum of household and government savings minus the value of newly issued government bonds:

```

*
* --- regional savings, sum of savings by households and saving by government
*
e_rsav(rsNat(rNat),ts(t)) ..

rsav(rNat,t)/rsav.scale(rNat,t) =e= (sum((r_r(rs,rNat),h) $ hFlag(rs,h),rsavC(rs,h,t))-newGovBond(rNat,t)+rsavG(rNat,t))/rsav.scale(rNat,t);

```

The GTAP Data Base does not comprise data on government savings which would allow to determine the deficit or surplus in its account. Available are data on total tax income and government final consumption. The rest are a contribution to the savings accounts and/or transfers. Re-bookings between different types of final consumption are a complex exercise, as expenditure shares of products, indirect tax rates and shares of the domestic and imported agent differ across the final consumption categories (private, government, investments).

Without touching the final demands and keeping I=S implies that a government deficit increases the government contribution of savings for investments and reduction in the contribution of final household. The newly issued bonds close the resulting gap in the budgets.

Benchmark of the equation structure above requires data on (1) the government deficit in the current year and on (2) yields of government bonds. For simplicity, we assume no repayment of past bonds at the benchmark.

In a first step, we increase the government savings by the assumed deficit, defined as a share of government consumption and savings (see *myGtap/myGtap\_cal.gms*):

```

*
*   --- newly issued government bonds at benchmark, assume 5% gov deficit
*
newGovBond(rNat,t0) = p_budDeficit(rNat) * (yg.l(rnat,t0) + rsavg(rnat,t0));
*
*   --- increase gov savings accordingly
*
rSavg(rnat,t0) = rsavg(rnat,t0) + newGovBond(rNat,t0);
xSavg(rnat,t0) = rsavg(rnat,t0);

```

Next, as we don't have information on the portfolio composition of individual households, we calculate total private savings as the sum over the households and derived from there the average expected returns:

```

:
: *
: *   --- assume yield equal to half of expected return to capital
: *
govBondYld(rnat,t0) = p_bondYield(rNat);
:
: *
: *   --- calculate average return to private savings
: *
parameter p_privSav(rNat,t0);
p_privSav(rNat,t0) = sum( (r_r(r,rNat),h) $ hFlag(r,h), rsavc(r,h,t0));
:
:
: pFinh(r,h,t0) $ hFlag(r,h) = sum(r_r(r,rNat), [ newGovBond(rNat,t0) * govBondYld(rnat,t0)/govBondYld(rnat,t0)
:                                     + (p_privSav(rNat,t0)-newGovBond(rNat,t0) ) * rore(rnat,t0)/rore(rnat,t0) ]
:                                     /p_privSav(rNat,t0));
:

```

Note that this calculation serves as a test only. Both prices are normalized with their benchmark levels and the spending on bonds and investment exhaust the total private savings. Accordingly, the resulting average revenue index is unity by definition.

Next, we set the transformation elasticity  $p\_omegaFinH$  for the households and calculate the share parameter for the investments for each private household  $p\_aFinRore$ :

```

*
*   --- transformation elasticity in private savings between investments and gov bonds
*
p_omegaFinH(r,h) = 2;
*
*   --- assume same share of gov bonds in total household savings for all households
*
p_aFinRore(r,h,t0) $ hFlag(r,h) = sum(r_r(r,rNat), (p_privSav(rNat,t0)-newGovBond(rNat,t0))/p_privSav(rnat,t0)
*                                     * (rore(rnat,t0)/[pFinh(r,h,t0)*rore(rnat,t0)])**p_omegaFinh(r,h));

```

Finally, we set the transformation elasticity  $p\_omegaExph$  for the distribution of household expenditures between final consumptions and savings and calculate its price index  $pExph$  and the share parameter for the expenditure on final demand on total household income  $p\_aExph$ . The latter must be equal to one minus the given saving rates of the household.

```

*
*   --- transformation elasticity between savings and final consumption
*
p_omegaExph(r,h) = 0.2;
*
*   --- average price of household expenditure (weighted average of consumer price index and returns to savings)
*
pExph(r,h,t0) $ hFlag(r,h) = [ (1-betash(r,h,t0))*pcons(r,h,t0) + betash(r,h,t0)*pFinh(r,h,t0) ];
*
*   --- share parameter of private expenditures
*
p_aExph(r,h,t0) $ hFlag(r,h) = (1-betash(r,h,t0));
*

```

In order to determine the revenue stream from past bonds bought by each household during simulation, the model needs to keep track of yield levels and bought bonds by each household. This

achieved by an additional parameter  $p\_bondReport$ <sup>17</sup> that stores how much bonds are bought at which yields in each simulated period, and calculates from there the revenues and add them up over past simulation points before the model is solved for the current period:

```
*
* --- calculate payments from yields of old gov bonds
*

p_bondReport(r,h,"newBonds",tSim-1) $ (hFlag(r,h) $ (p_aFinRore(r,h,tSim-1) ne 1))
= sum( r_r(r,rnat) $ p_aFinRore(r,h,tSim-1),
      rsavC(r,h,tSim-1) * (1-p_aFinRore(r,h,tSim-1))
      * (govBondYld(rnat,tSim-1)/[pFinH(r,h,tSim-1)*govBondYld.1(rnat,"%t0%")]])**p_omegaFinH(r,h));

p_bondReport(r,h,"newBondsToSavings",tSim-1) $ hFlag(r,h)
= p_bondReport(r,h,"newBonds",tSim-1)/rsavC(r,h,tSim-1);

p_bondReport(rNat,"gov","spendShrByGovInc",tSim-1) = spendShrByGovInc(rnat,tSim-1);
p_bondReport(rNat,"gov","newBonds",tSim-1) = newGovBond(rnat,tSim-1);
p_bondReport(rNat,"gov","NewBondsToTax",tSim-1) = newGovBond(rnat,tSim-1)/yTaxTot(rnat,tSim-1);

p_bondReport(r,h,"bondYld",tSim-1) $ hFlag(r,h) = sum(r_r(r,rnat), govBondYld(rnat,tSim-1));

p_bondReport(r,h,"rore",tSim-1) $ hFlag(r,h) = sum(r_r(r,rnat), rore(rnat,tSim-1));

p_bondReport(r,h,"bondRev",tSim-1) $ hFlag(r,h) = p_bondReport(r,h,"newBonds",tSim-1)*p_bondReport(r,h,"bondYld",tSim-1);

p_payPastGovBonds(r,h,tSim)
= sum(tRun $ (tRun.pos lt tSim.pos), p_bondReport(r,h,"bondRev",tRun)
      $if defined p_timeStep * p_timeStep(tRun)
      );
```

The report comprises some other positions to track the behavior of the mechanism

```
p_bondReport(rNat,"gov","payPast",tSim-1) = sum( (r_r(r,rNat),h),p_payPastGovBonds(r,h,tSim-1));
p_bondReport(rNat,"gov","payPastRel",tSim-1) = p_bondReport(rNat,"gov","payPast",tSim-1)/yTaxTot(rnat,tSim-1);
p_bondReport(rNat,"gov","GDPmP",tSim-1) = gdpmp(rnat,tSim-1);
p_bondReport(rNat,"gov","yTaxTot",tSim-1) = yTaxTot(rnat,tSim-1);
p_bondReport(rNat,"gov","yTaxTotAfterPayBond",tSim-1) = yTaxTot(rnat,tSim-1)-p_bondReport(rNat,"gov","payPast",tSim-1);
p_bondReport(rNat,"gov","yg",tSim-1) = yg(rnat,tSim-1);
p_bondReport(rNat,"gov","rsavg",tSim-1) = rsavg(rnat,tSim-1);
p_bondReport(rNat,"gov","grest",tSim-1) = -aidi(rNat,tSim-1)*aidiShrCor(tSim-1)*lcu(rNat,tSim-1) + aido(rNat,tSim-1)*lcu(rNat,tSim-1)
+ sum((r_r(rs,rNat),h) $ hFlag(rs,h), trng(rs,h,tSim-1)*pg(rNat,tSim-1));

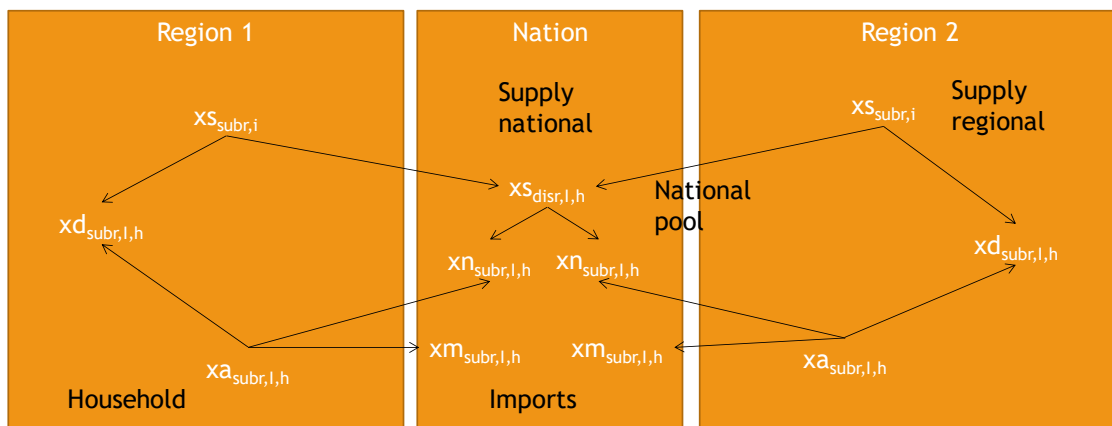
p_bondReport(rNat,"gov","SpendToTax",tSim-1) = ( p_bondReport(rNat,"gov","yg",tSim-1)
+ p_bondReport(rNat,"gov","rsavg",tSim-1)
+ p_bondReport(rNat,"gov","grest",tSim-1))/p_bondReport(rNat,"gov","yTaxTot",tSim-1);

p_bondReport(rNat,"gov","SpendToTaxNet",tSim-1) = ( p_bondReport(rNat,"gov","yg",tSim-1)
+ p_bondReport(rNat,"gov","rsavg",tSim-1)
+ p_bondReport(rNat,"gov","grest",tSim-1))/p_bondReport(rNat,"gov","yTaxTotAfterPayBond",tSim-1);
```

### 3.2.5 Methodology: Household demand

In opposite to GTAP standard, the myGtap module considers potentially three different origins in the top Armington nest: (1) sub-national, (2) national pool and (3) import demand. For all Armington agents which are not private households, the national pool demand layer is not present and defaults to what it called “regional” demand in the model, i.e. domestic demand. That is somewhat confusing ( $xd$  and  $pd$  might refer to the nation or the sub-region) but keeps the changes in the code structure minimal.

<sup>17</sup> The report can be added to the solve report output which is updated in between simulation points in recursive-dynamic runs to track the model behavior.



**Figure 5: Structure of demand system considering subsistence production**

The macro *m\_padef* defining the Armington composite prices shown in equation (24) below hence comprises potentially three share parameters *p\_alphan* (= national pool), *p\_alphad* (domestic = sub-national region for private household, domestic for all other agents) and *alpham* (imports). The macro *p\_pdp* defines tax inclusive prices faced by the agents for the regional and the national origin. In case of the national pool, national prices are introduced as captured by the cross-set *r\_r*. The domestic (or sub-national) demand uses the price of that region directly. The macro *p\_pmp* defines the import price plus agent specific taxes.

$$\begin{aligned}
 \text{\$macro } m\_padef(r,i,aa,t) \{ & ( \quad \backslash \\
 & + [alphan(r,i,aa,t)*sum(rr\_rr(r,rNat),m\_pdp(rNat,i,aa,t)**(1-sigmam(r,i)))] \$ (alphan(r,i,aa,t) \text{ gt eps}) \quad \backslash \\
 & + [alphad(r,i,aa,t)*m\_pdp(r,i,aa,t)**(1-sigmam(r,i))] \$ (alphad(r,i,aa,t) \text{ gt eps}) \quad \backslash \\
 & + [alpham(r,i,aa,t)*m\_pmp(r,i,aa,t)**(1-sigmam(r,i))] \$ (alpham(r,i,aa,t) \text{ gt eps}) \quad \backslash \\
 & ) ** (1/(1-sigmam(r,i))) \} \quad \quad \quad (24)
 \end{aligned}$$

The version of the macro shown above is only used if the global *hSubr* (= Household at sub-national level) is set to “on” by a switch on the interface. Otherwise, the simpler default version in equation (25) is used which only considers the domestic and imported origin:

$$\begin{aligned}
 \text{\$macro } m\_padef(r,i,aa,t) \{ & ( \quad [p\_alphad(r,i,aa,t) * pdp(r,i,aa,t) \quad ** (1-p\_sigmam(r,i))] \$ p\_alphad(r,i,aa,t) \quad \backslash \\
 & + [p\_alphan(r,i,aa,t) * (pmp(r,i,aa,t)/lambdaxm(r,i,aa,t)) ** (1-p\_sigmam(r,i))] \$ p\_alphan(r,i,aa,t) \quad \backslash \\
 & ) ** (1/(1-p\_sigmam(r,i))) \} \quad \quad \quad (25)
 \end{aligned}$$

The resulting Marshallian demands are standard ones for a CES demand system and the usual way to implement the Armington assumption; the national pool demand *xn* for households in sub-national regions is defined as shown in equation (26) below. The left hand *xn* depends on the total Armington demand defined in the macro *mm\_xa*, pre-multiplied with the share parameter for the national pool *p\_alphan*. The resulting quantity is updated with the price relation between the Armington price *m\_pa* and the national price *m\_pdp* faced by agent, considering the substitution elasticity between the imported and domestic origin *p\_sigmam* as supplied by the GTAP data base:

$$\begin{aligned}
 \text{\$xn\_subr,iIn(i,aa,ts) } & \$ (p\_alphan(subr,i,aa,ts) \text{ gt eps}) \quad \quad \quad (26) \\
 \text{xn(subr,i,aa,ts)s(xn,'subr,i,aa,ts')} & = e = p\_alphan(subr,i,aa,ts)*mm\_xa(subr,i,aa,ts)s(xn,'subr,i,aa,ts') \\
 & + ((m\_pa(subr,i,aa,ts) \$ (not addCes(subr,i,'imp')) + mm\_padef(subr,i,aa,ts) \$ addCes(subr,i, \\
 & \quad /sum(r\_r(subr,disr),m\_pdp(disr,i,aa,ts)))*p\_sigmam(subr,i);
 \end{aligned}$$

Note again the use of the *r\_r* cross-set which puts the national price into the price relation.

The equation for the sub-national (respectively domestic) demand *xd* in equation (27) is unchanged from the existing CGEBox code, it comprises the same structural elements as above:



```
e_xd(rs,iIn(i),aaDM,ts) $ ( (p_alphad(rs,i,aaDM,ts) gt eps) $ (xd.range(rs,i,aaDM,ts) ne 0) ) ..
xd(rs,i,aaDM,ts)s(xd,'rs,i,aaDM,ts') =e= p_alphad(rs,i,aaDM,ts)*mm_xa(rs,i,aaDM,ts)s(xd,'rs,i,aaDM,ts')
+ ((m_pa(rs,i,aaDM,ts) $ (not addCes(rs,i,"imp"))) + m_xaidd(rs,i,aaDM,ts) $ addCes(rs,i,"imp"))
/m_pdp(rs,i,aaDM,ts))*p_sigmam(rs,i);
```

(27)

The total import demand equation  $e_{xmt}$  (28) now comprises an additional line which adds the import demand of sub-national households to the import demand of agents defined at national level (first line, the default from the GTAP Standard):

```
--- Total import demand, equation (50) in VDM 2018
e_xmt(rsNat,iIn(i),ts) $ xmtFlag(rsNat,i) ..
xmt(rsNat,i,ts)s(xmt,'rsNat,i,ts') =e= sum(aaDM $ (p_alpham(rsNat,i,aaDM,ts) gt eps), xm(rsNat,i,aaDM,ts))s(xmt,'rsNat,i,ts')
$$if "ModulesGTAP_NUTS2%"=="ON" +sum( (subr,h) $ (r_r(subr,rsNat) $ (p_alpham(subr,i,h,ts) gt eps)),xm(subr,i,h,ts))s(xmt,'rsNat,i,ts')
;
```

(28)

The equilibrium for the national market depicted in the equation  $e_{pd}$  (29) is equally expanded by a second expression which considers demand of the private households from the national pool. The preprocessor `$$ifthen-$$endif` commands allow to introduce the expression only in case the sub-national household layer is active:

```
--- Domestic goods market equilibrium, equation (67) in VDM 2018
e_pd(rsNat,iIn(i),ts) $ xdFlag(rsNat,i) ..
xds(rsNat,i,ts)s(xds,'rsNat,i,ts')
=e= sum(aa $ (m_alphad(rsNat,i,aa,ts) gt eps), xd(rsNat,i,aa,ts))s(xds,'rsNat,i,ts')
$$iftheni.hsubr %hSubr%==on
+ sum( (rs,h) $ (r_r(rs,rsNat) $ (p_alphan(rs,i,h,ts) gt eps)),xn(rs,i,h,ts))s(xds,'rsNat,i,ts')
$$endif.hsubr
;
```

(29)

The equation  $e_{psSubreg}$  (30) defines the share of domestic supply  $xs$  for dis-aggregated nations  $disr$  demanded from their sub-regions  $subr$  now considers the possibility that part of the demand stems directly from households in the same regions. That demand is subtracted in the first line from regional supply:

```
e_psSubreg(subr,iIn(i),ts) $ ( sum(a,xFlag(subr,a,i)) $ (xs.range(subr,i,ts) ne 0) $ diag(i)) ..
(xs(subr,i,ts)
$$iftheni.hsubr "%hSubr%"=="on"
- sum(h $ (p_alphad(subr,i,h,ts) gt eps), xd(subr,i,h,ts))
$$endif.hsubr
- sum(aaDM,p_subrDem(subr,i,aaDM,ts))
)s(xs,'subr,i,ts')
=E= sum(r_r(subr,disr) $ p_gfi(subr,i,disr,ts),
xs(disr,i,ts) * p_gfi(subr,i,disr,ts) * (ps(disr,i,ts)/ps(subr,i,ts))*p_sigmaSubr(disr,i))s(xs,'subr,i,ts');
```

(30)

Thus, the output quantity of the sub-national region  $xs$  is reduced by the demand of households for produce from the same sub-nation. Only the remaining part feeds into the national output quantity  $xs$  of the dis-aggregated nation  $disr$ .

## Closures

The code currently supports two types of closures: either the driving equations which use a fixed share on a specific total as depicted above or fixing the transactions and rendering the shares endogenous:



```
$$iftheni "%closure_foreign_capital_income_sent%"=="Driven by capital income"
    fyoheq.fyoh
$$else
    fyoheq.fyohShr
$$endif
$$iftheni "%closure_foreign_capital_income_received%"=="Share on global total"
    fyiheq.fyih
$$else
    fyiheq.fyihShr
$$endif
$$iftheni "%closure_remittance_sent%"=="Driven by labor income"
    remoheq.remoh
$$else
    remoheq.remohShr
$$endif
$$iftheni "%closure_remittance_received%"=="Share on global total"
    remiheq.remih
$$else
    remiheq.remihShr
$$endif
$$iftheni "%closure_foreign_aid_sent%"=="Driven by factor income"
    aidoeq.aido
$$else
    aidoeq.aidoShr
$$endif
$$iftheni "%closure_foreign_aid_received%"=="Share on global total"
    aidieq.aidi
$$else
    aidieq.aidiShr
$$endif
```

The closures can be defined via the GUI as discussed below. As seen, that solution works currently globally. It is possible to fix also transactions for single countries and render shares endogenous, however, in that case, only the CNS and not the MCP solution algorithm can be used.

### Integrating data from household surveys

The myGTAP module requires only very basic information for benchmarking, assumed to be available from any household survey, namely factor income and population shares. Data on differences in direct tax rates or shares on total private savings can be provided additionally; otherwise, identical savings and direct tax rates for all households are used as derived from the GTAP SAMs. International transfers, between households and from government to household are left out from the model if no data are entered.

Data are inputted as follows:

1. Factor income, savings and population are inputted as shares of each private household type on economy-wide totals as found in the GTAP data base
2. Direct tax rates are relative to the direct tax rates for the aggregate average household as found in the GTAP data base
3. Transfers (foreign, between households, from government) are relative to factor income minus direct taxes
4. Foreign aid (in/out) is defined relative to total factor income

```

table shrHhsls(r,hhsls,*,t)
*
*   --- note: factor shares, savings and populationn as shares on aggregate total
*
*           land.(set.t)  unSkLab.(set.t)  skLab.(set.t)  capital.(set.t)  natRes.(set.t)
(set.r).poor    0.2        0.8          0.4          0.2          0.1
(set.r).rich    0.8        0.1          0.3          0.8          0.9
*
*   --- note: for direct taxes, these are relative differences in tax rates!
*
*           dirTax.(set.t)  savings.(set.t)  pop.(set.t)
(set.r).poor    0.5        0.3          0.8
(set.r).rich    1.5        0.7          0.2
*
*   --- note: for remiH, remoH,fyiH,fyoH,trng
*           these are shares on each household factor income minus direct taxes
*
*           remiH.(set.t)  remoH.(set.t)  fyiH.(set.t)  fyoH.(set.t)  trng.(set.t)
(set.r).poor    0.01      0.00        0.01        0.00        0.25
(set.r).rich    0.00      0.02        0.00        0.02        0.05
*
*   --- note: for household transfers (hsls (rows) - > hhslds (columns))
*           these are shares on each household factor income minus direct taxes
*
*           poor.(set.t)  rich.(set.t)
(set.r).poor    0.00      0.01
(set.r).rich    0.05      0.00
;

*
*   --- shares of foreign income to governments (in/ou), relative to factor income
*
table shrAid(r,*,t)
      aidl.(set.t)  aido.(set.t)
(set.r)    0.01      0.02
;

```

The shares on factor income, savings and population are scaled to unity by the benchmarking code. This allows to either enter the data as shares (as shown below) or as totals. The scaling will in both cases take care of potential rounding or other editing errors. Note that for the other entries, absolute numbers must be translated to the respective shares. That can be done by data transformations in the very same file using the SAM entries already available at the point in the code where the data information is included in the general calibration routine “*cal/cal.gms*”.

The number of household types is comprised in the same file. Note the necessary “*\$ftheni*” clause which ensures that the information is made available to the code when the general set definitions are set-up in “*build/loadGTAPagg.gms*”:

```

$iftheni %l==decl
  set set_hhsls(*) / poor,rich/;
$else

```

The code is set-up such that the household split up is introduced for one, some or all regions, solely driven by the data provided. Currently, the parameters of the CDE demand system are identical. The code allows for two options with regard to final demand by the disaggregated households:

1. The consumption shares between all households are identical.
2. The expansion parameters of the CDE system are used to update the consumption shares, and afterwards, the Armington demands for the different households are adjusted to simultaneously exhaust the given SAM entries and the final consumption total of the disaggregated households.

The user could provide household specific substitution elasticities between domestic and imports based on an extended parameter file, an option generally available for CGEBox.

An alternative is to assign factor income based on activities:

```
*
*      . . . . .
* --- assign income from agricultural activities to agricultural household
*
shrHhsls(r,"agr",agr,t) = 1;
```

Which can be combined with other shares. If data for activities are given, these have priorities.

### SAM rebalancing

The integration of new transactions between model regions such as remittances requires that such transaction add up to zero globally, for instance, the sum of remittances received by all households globally must be equal to the sum of remittances sent. The balancing condition is achieved by applying uniform relative correction factors to the flows. The correction factor distributes any original imbalances, here shown for the case of remittance flows:

```
*
* --- balance global flow of remittances
*
parameter p_surPlus(*,*,t);
p_surPlus("wor","rem",t) = sum( (r,hhsls), remoH.l(r,hhsls,t) - remiH.l(r,hhsls,t));
p_surPlus("wor","wgt",t) = sum( (r,hhsls), remoH.l(r,hhsls,t) + remiH.l(r,hhsls,t));

remoH.l(r,hhsls,t) $ p_surPlus("wor","wgt",t) = remoH.l(r,hhsls,t) * ( 1 - p_surPlus("wor","rem",t)/p_surPlus("wor","wgt",t));
remiH.l(r,hhsls,t) $ p_surPlus("wor","wgt",t) = remiH.l(r,hhsls,t) * ( 1 + p_surPlus("wor","rem",t)/p_surPlus("wor","wgt",t));
```

The first line calculates the imbalance in the original data (“rem”), the second the weights (“wgt”) to derive a uniform correction factor. The two remaining statements correct all remittance flows such that a zero global balance is resulting. The same logic is applied to balance foreign capital income inflows and outflows and foreign aid flows to governments, and also transfers between households inside of one country.

The introduction of new global transfer exchanges will generally imply that the BOP is no longer equal to zero as it is the case in the original SAMS. In order to close the balance again, we first define the new additional net inflows as variable *finc*. A positive sign implies that inflows exceed outflows:

```
*
* --- foreign contribution to income (can be negative)
*
finc.l(r,t) = + aidi.l(r,t) + sum(h,remiH.l(r,h,t)) + sum(h, fyih.l(r,h,t))
              - aido.l(r,t) - sum(h,remoH.l(r,h,t)) - sum(h, fyoh.l(r,h,t));
```

In order to maintain that BOP=0, foreign savings *savf* are downward corrected accordingly:

```
*
* --- correct foreign savings by contribution to income to maintain BOT
*
savf.l(r,t) = savf.l(r,t) - finc.l(r,t);
```

Which in turn requires, according to I=S, that regional savings are re-balanced as well, using the same absolute correction:

```
*
* --- and correct regional savings in opposite direction to main I=S
*
rsav.l(r,t) = rsav.l(r,t) + finc.l(r,t);
```

That ensures that aggregate private consumption, government consumption and savings for each commodity need not to be rebalanced.

### Estimating final demand shares

In case the CDE expansion parameters are used to define final demand, the following procedure is used. As a first step, the Armington demands are derived from the given per capita income differences:

```
xa.1(r,i,hhslds,t) $ xa.1(r,i,"hhsld",t) = xa.1(r,i,"hhsld",t)
* yc.1(r,hhslds,t)/yc.1(r,"hhsld",t)
* [(yc(r,hhslds,t)/pop(r,hhslds,t))/(yc(r,"hhsld",t)/pop(r,"hhsld",t))] ** (eh(r,i,t));
```

The first line reflects differences the distribution of total final household consumption *yc* of the aggregate household as found in the GTAP data base to the different households *hhslds*. The second line reflects the relation between in per capita income of each household to aggregate one, exponent the expansion effect of the CDE demand system.

In order to render the updated Armington demands for the different households consistent both with the given total spent for private consumption *yc* for each household and with the aggregate SAM entries for private consumption for each commodity, the following two consistency equations are used:

```
*
* --- sum of Armington demands by households must exhaust SAM entry
*
e_xaCons(rs,i,t) $ xa.1(rs,i,"hhsld",t) ..
    sum(hhslds, xa(rs,i,hhslds,t)) =E= xa.1(rs,i,"hhsld",t);

e_ycCons(rs,hhslds,t) ..
*
* --- sum of Armington demands by households must exhaust given
* private consumption
*
sum(i, xa(rs,i,hhslds,t)) =E= yc.1(rs,hhslds,t);
```

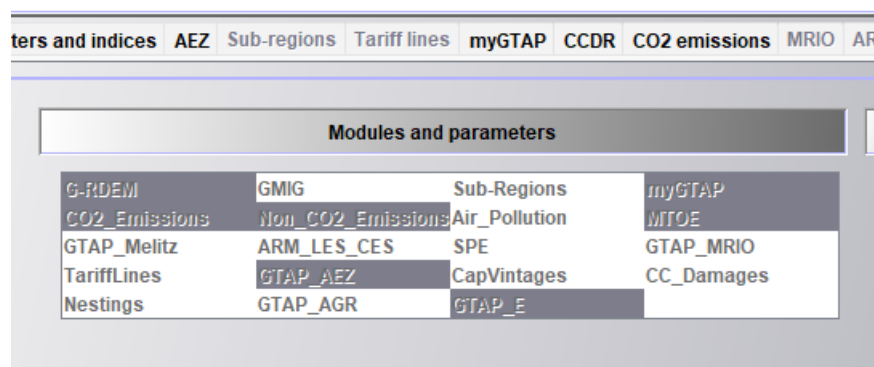
These are combined into a simple model which minimizes squared relative differences from the estimates resulting from applying the expansion elasticities:

```
e_hdpXA ..
v_hdp * sum( (rs,i,hhslds,t) $ xa.1(rs,i,hhslds,t),1)
=E= sum( (rs,i,hhslds,t) $ xa.1(rs,i,hhslds,t),
sqr( (xa(rs,i,hhslds,t)-xa.1(rs,i,hhslds,t))/xa.1(rs,i,hhslds,t)));
```

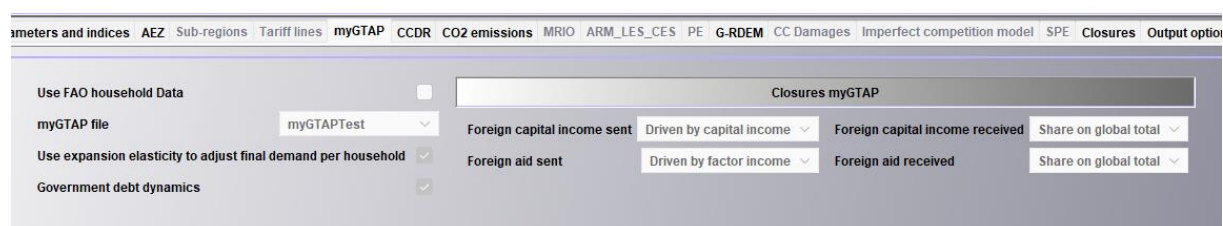
### 3.2.6 Graphical User Interface

#### Options

In order to use the extension, the “myGTAP” module must be switched on the interface:



In that case a tab becomes visible which allows detailing options for the myGTAP module as shown below. The closures have been discussed above.



The selection box of files under “myGTAP file” allows selecting a file with data on the household disaggregation and international transactions as discussed above. These files are stored in “gams\scen\myGTAP”, some of the files are shown below:

The “*regHousehold.gms*” file removes only the regional household from the GTAP Model and introduces separate accounts for the agents without any household differentiation:

```
$iftheni %1==decl
  $$onempty
    set set_hhsls(*) ;;
  $$offempty
$endif
```

The “*myGTAP\_oneHHS�D.gms*” shows how the extension can be used to introduce international transactions without dis-aggregating to different household types, i.e. all shares and relations for that single aggregate private household are set to unity, but remittance, foreign capital income and foreign information is provided:

```

$iftheni %1==decl
  set set_hhslds(*) / agg/;

  set set_myGTAP / dirTax,savings,pop,remiH,remoH,fyiH,fyoH,trng,aidi,aido /;
$else

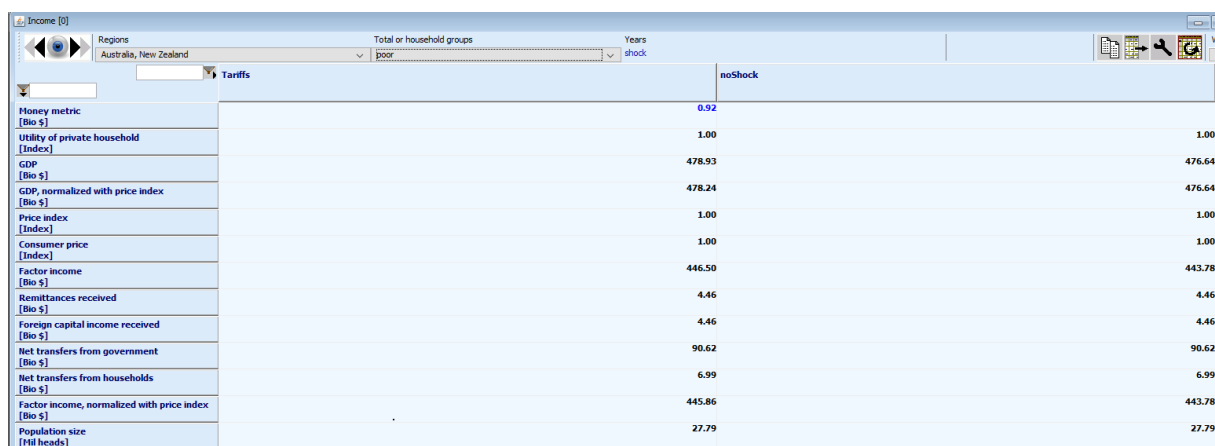
  table shrHhslds(r,hhslds,*,t)
*
*   --- note: for remiH, remoH,fyiH,fyoH
*   these are shares on each household factor income minus direct taxes
*
  (set.r).agg      land.(set.t)  unSkLab.(set.t)  skLab.(set.t)  capital.(set.t)  natRes.(set.t)
*
*   1.0            1.0          1.0          1.0          1.0          1.0
*
*   --- note: for direct taxes, these are relative differences in tax rates!
*
  (set.r).agg      dirTax.(set.t)  savings.(set.t)  pop.(set.t)
*
*   1.0            1.0          1.0
*
*   --- note: for remiH, remoH,fyiH,fyoH,trng
*   these are shares on each household factor income minus direct taxes
*
  (set.r).agg      remiH.(set.t)  remoH.(set.t)  fyiH.(set.t)  fyoH.(set.t)  trng.(set.t)
*
*   0.01           0.00         0.01         0.00         0.10
*
*
*   --- shares of foreign income to goverments (in/ou), relative to factor income
*
  table shrAid(r,*,t)
  (set.r)          aidI.(set.t)  aido.(set.t)
*
*   0.01           0.02
*
$endif

```

The file “myGTAP\_GTAP\_AGR” introduces an agricultural and non-agricultural. All factor income from agriculture accrues to the agricultural household. It uses data from the GTAP-AGR module to define the non-farm income of the agricultural household.

### Exploitation of myGTAP results

The post-processing code aggregates the information from single household to a regional private household and allows analyzing results at the level of the dis-aggregated household types:



|                                                     | noShock | shock  |
|-----------------------------------------------------|---------|--------|
| Money metric [Bio \$]                               | 0.92    |        |
| Utility of private household [Index]                | 1.00    | 1.00   |
| GDP [Bio \$]                                        | 478.93  | 476.64 |
| GDP, normalized with price index [Bio \$]           | 478.24  | 476.64 |
| Price index [Index]                                 | 1.00    | 1.00   |
| Consumer price [Index]                              | 1.00    | 1.00   |
| Factor income [Bio \$]                              | 446.50  | 443.78 |
| Remittances received [Bio \$]                       | 4.46    | 4.46   |
| Foreign capital income received [Bio \$]            | 4.46    | 4.46   |
| Net transfers from government [Bio \$]              | 90.62   | 90.62  |
| Net transfers from households [Bio \$]              | 6.99    | 6.99   |
| Factor income, normalized with price index [Bio \$] | 445.86  | 443.78 |
| Population size [Mil heads]                         | 27.79   | 27.79  |

On top, tables with demand and factor income information at that level are provided:

Household demand, by product [0]

Regions: Australia, New Zealand Oceania | Items: Quantity Q | Origins: Total tot | Years: shock shock | Percentage diff. to Scenarios: noShock noShock

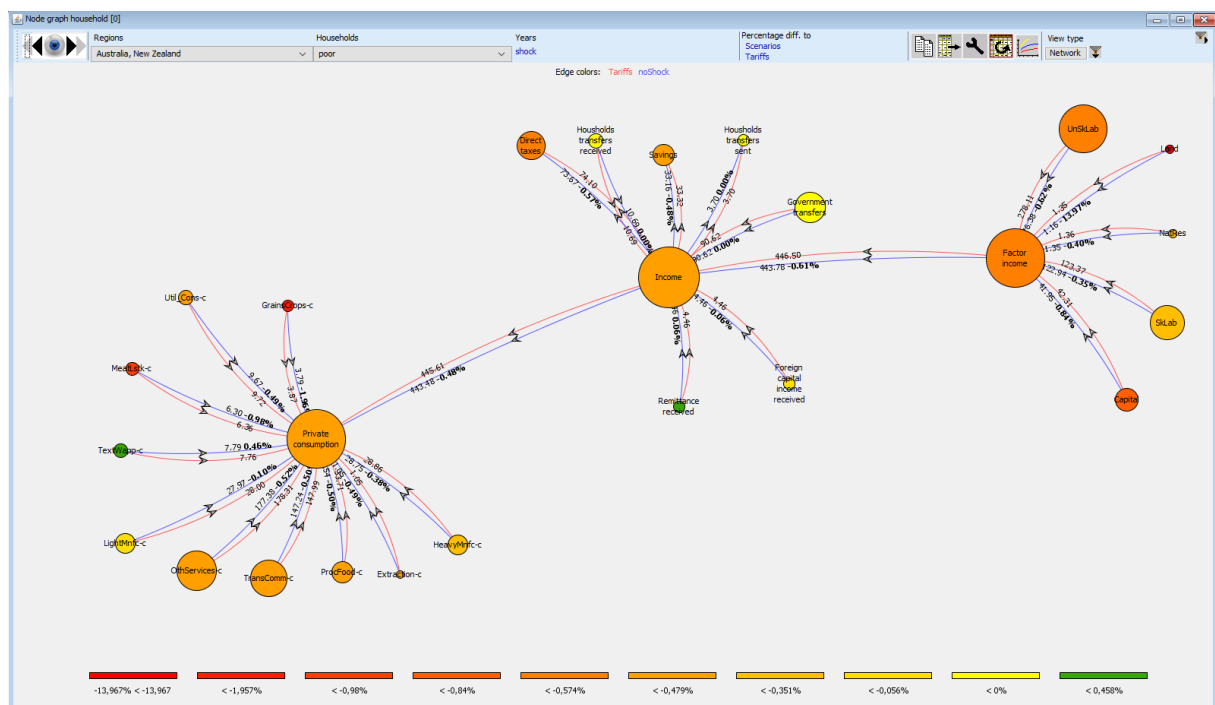
|                                         | Private consumption hhsld |                 |                 | noShock noShock |                 |                 |
|-----------------------------------------|---------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                         | poor poor                 | rich rich       |                 | poor poor       | rich rich       |                 |
| Total totP                              | 529.07<br>0.63%           | 337.15<br>0.50% | 191.92<br>0.86% | 525.76<br>0.86% | 335.49<br>0.86% | 190.28<br>0.86% |
| Grains and Crops GrainsCrops-c          | 4.72<br>-0.02%            | 3.01<br>-0.03%  | 1.71<br>-0.01%  | 4.72<br>0.62%   | 3.01<br>0.62%   | 1.71<br>0.62%   |
| Livestock and Meat Products MeatLstk-c  | 8.06<br>-0.40%            | 5.14<br>-0.51%  | 2.92<br>-0.20%  | 8.10<br>0.73%   | 5.17<br>0.73%   | 2.93<br>0.73%   |
| Mining and Extraction Extraction-c      | 1.38<br>0.53%             | 0.88<br>0.41%   | 0.50<br>0.73%   | 1.37<br>0.62%   | 0.88<br>0.62%   | 0.50<br>0.62%   |
| Processed Food ProcFood-c               | 33.61<br>0.42%            | 21.43<br>0.31%  | 12.19<br>0.62%  | 33.47<br>0.62%  | 21.36<br>0.62%  | 12.11<br>0.62%  |
| Textiles and Clothing TextWapp-c        | 8.55<br>2.73%             | 5.45<br>2.60%   | 3.10<br>2.96%   | 8.32<br>2.96%   | 5.31<br>2.96%   | 3.01<br>2.96%   |
| Light Manufacturing LightMnfc-c         | 33.59<br>1.89%            | 21.41<br>1.75%  | 12.19<br>2.13%  | 32.97<br>2.13%  | 21.04<br>2.13%  | 11.93<br>2.13%  |
| Heavy Manufacturing HeavyMnfc-c         | 29.38<br>1.14%            | 18.72<br>1.01%  | 10.66<br>1.38%  | 29.05<br>1.38%  | 18.54<br>1.38%  | 10.51<br>1.38%  |
| Utilities and Construction Util_Conc-c  | 12.04<br>0.57%            | 7.67<br>0.44%   | 4.37<br>0.81%   | 11.97<br>0.81%  | 7.64<br>0.81%   | 4.33<br>0.81%   |
| Transport and Communication TransComm-c | 180.82<br>0.54%           | 115.22<br>0.40% | 65.60<br>0.78%  | 179.85<br>0.78% | 114.76<br>0.78% | 65.09<br>0.78%  |
| Other Services OthServices-c            | 216.92<br>0.45%           | 138.23<br>0.32% | 78.69<br>0.69%  | 215.94<br>0.69% | 137.79<br>0.69% | 78.15<br>0.69%  |

Factor income [0]

Regions: Australia, New Zealand | Items: Value (net of taxes) | Origins: dom | Years: shock

|                     | Total           | Land           | UnSkLab         | SkLab           | Capital         | NatRes |
|---------------------|-----------------|----------------|-----------------|-----------------|-----------------|--------|
| Aggregate Household | 762.02<br>0.91% | 6.73<br>15.85% | 313.46<br>0.82% | 216.38<br>0.58% | 212.01<br>1.08% |        |
| poor                | 447.37<br>0.81% | 1.35<br>15.85% | 278.63<br>0.82% | 123.65<br>0.58% | 42.40<br>1.08%  |        |
| rich                | 314.65<br>1.06% | 5.39<br>15.85% | 34.83<br>0.82%  | 92.74<br>0.58%  | 169.61<br>1.08% |        |

Node graphs (in value, in value normalized with price index, and per capita) depict graphical income generation and use for each household type, including transfer transactions:



### 3.2.7 Code implementation

#### Data entry and calibration

The information on the list of dis-aggregated household types is provided by the include file chosen on the user interface. It defines the household types which technically co-exist with the aggregate private household found in the standard option termed “*hhsld*”:

```
set h(fd) Households /
  hhsld      "Aggregate Household"
  set.set_hhslds
/ ;
```

That aggregate “*hhsld*” is only used when the data are loaded, in the calibration code “*cal.gms*” and are again introduced post-model for reporting. The household split-up is introduced in “*cal.gms*”. In case that only the regional household mechanism is replaced, but one aggregate private household agent is used, the split-up parameters are simply set to all zeros:

```
$$ifthen not defined ShrHhslds
*
*   --- no data on share of households, assume
*       top level regional household, i.e. one aggregate household own all factors
*
  parameter shrHhslds(r,hhslds,*,t);
  option kill=shrHhslds;
  option kill=shrXF;
  set dummyHslds/dirTax,Savings/;
$$else
  shrXF(r,hhslds,f,t) = shrHhslds(r,hhslds,f,t);
$$endif
```

In that case or if for some country, no distributional data is provided., the ownership share to factors are all unity:

```
*
*   --- if factor shares are missing for all households, use the standard aggregate
*       (allows to use dis-aggregated data for one country in full global model)
*
loop(r $ (not sum( (hhslds,f,t), shrXF(r,hhslds,f,t))),
  shrXF(r,"hhsld",f,t) = 1;
);
```

Next, the code checks if the factor share information is consistent, i.e. if for one region, information for each household type and factor is edited:

```
loop( (r,f) $ (card(hhslds) $ (not sum( (hhslds,t), shrXF(r,hhslds,f,t)))),
  shrXF(r,hhslds,f,t) = eps;
  abort "Factor shares for some households reported, but for other missings ",shrXF;
);
```

If the user wants to introduce zero shares, it should edit an “*eps*” value in the table shown above. The code will automatically scale the data to unity which allows to provide absolute data in the table instead of shares:



```

: *
: *   --- scale factor shares to unity
: *
:   shrXF(r,h,f,t) = shrXF(r,h,f,t)/sum(h0,shrXF(r,h0,f,t));
:

```

Next, differences in direct tax rates are introduced. A similar test as above is implemented. If no information is given, uniform rate are used:

```

*
*   --- shift direct taxes by given difference in direct taxation
*
loop( r $ ( sum( (hhslds,t), shrHhslds(r,hhslds,"dirTax",t))
              $ (sum( (hhslds,t) $ shrHhslds(r,hhslds,"dirTax",t),1) ne card(t)+card(hhslds))),
      shrHhslds(r,hhslds,"dirTax",t) $ (not shrHhslds(r,hhslds,"dirTax",t)) = eps;
      abort "Direct tax shares for some households reported, but for other missings, look out for eps ",shrHhslds;
);

kappafH.1(r,f,"hhsld",t) = kappaf.1(r,f,t);
kappafH.1(r,f,hhslds,t) $ (not shrHhslds(r,hhslds,"dirTax",t)) = kappaf.1(r,f,t);
kappafH.1(r,f,hhslds,t) $ shrHhslds(r,hhslds,"dirTax",t) = kappaf.1(r,f,t) * shrHhslds(r,hhslds,"dirTax",t);

```

From that information, household specific factor income and direct taxes are derived:

```

*
*   -- factor income for each household, based on share on factor ownership
*
factYH.1(r,h,t) = sum(f, pft.1(r,f,t)*xft.1(r,f,t)*shrXf(r,h,f,t))
                  - fdepr(r,t)*pi.1(r,t)*kstock.1(r,t)*sum(sameas(f,cap), shrXf(r,h,cap,t))
                  + sum((fnm,a), pf.1(r,fnm,a,t)*xf.1(r,fnm,a,t)*shrXf(r,h,fnm,t));

factY.1(r,t) = sum(f, pft.1(r,f,t)*xft.1(r,f,t)) - fdepr(r,t)*pi.1(r,t)*kstock.1(r,t)
              + sum((fnm,a), pf.1(r,fnm,a,t)*xf.1(r,fnm,a,t));
factYH.1(r,"hhsld",t) = factY.1(r,t);

*
*   --- direct taxes by household
*
dirTax.1(r,f,h,t) =
+ ( sum(sameas(fm,f), kappafh(r,fm,h,t) * pft(r,fm,t)*xft(r,fm,t)*shrXf(r,h,fm,t))
+ sum((sameas(f,fnm),a), kappafh(r,fnm,h,t) * pf(r,fnm,a,t)*xf(r,fnm,a,t)*shrXf(r,h,fnm,t)) );

dirTax.1(r,f,"hhsld",t) =
+ ( sum(sameas(fm,f), kappaf(r,fm,t) * pft(r,fm,t)*xft(r,fm,t))
+ sum((sameas(f,fnm),a), kappaf(r,fnm,t) * pf(r,fnm,a,t)*xf(r,fnm,a,t)) );

```

Which allows defining direct tax rates which recovers economic wide totals. Next direct taxes paid by the household types are updated:

```

kappaPh.1(r,f,hhslds,t) $ dirTax.1(r,f,hhslds,t)
= kappaPh.1(r,f,hhslds,t)/[sum(hhslds0,dirTax.1(r,f,hhslds0,t))/dirTax.1(r,f,"hhsld",t)];

dirTax.1(r,f,hhslds,t) =
+ ( sum(sameas(fm,f), kappafh(r,fm,hhslds,t) * pft(r,fm,t)*xft(r,fm,t)*shrXf(r,hhslds,fm,t))
+ sum((sameas(f,fnm),a), kappafh(r,fnm,hhslds,t) * pf(r,fnm,a,t)*xf(r,fnm,a,t)*shrXf(r,hhslds,fnm,t)) );

```

A similar approach is used for saving rates, i.e. if no information is given, the same shares as for the aggregate private household are used:

```

*
*   --- determine savings
*
loop( r $ ( sum( (hhslds,t), shrHhslds(r,hhslds,"savings",t))
              $ (sum( (hhslds,t) $ shrHhslds(r,hhslds,"savings",t),1) ne card(t)+card(hhslds))),
      shrHhslds(r,hhslds,"savings",t) $ (not shrHhslds(r,hhslds,"savings",t)) = eps;
      abort "Savings shares for some households reported, but for other missings, look out for eps ",shrHhslds;
);

*
*   --- scale saving shares to unity
*
shrHhslds(r,hhslds,"savings",t) = shrHhslds(r,hhslds,"savings",t)/sum(hhslds0,shrHhslds(r,hhslds0,"savings",t));

```

In order to distribute regional savings to the different accounts, first the government savings are calculated residually:

```

*
* --- savings of government: by definition difference between total tax income and government consumption
*
*   rsavg.l(r,t)      = ytaxTot(r,t) - yg(r,t);

```

In absence, of any information of foreign contributions to the regional government, it is assumed that the government saving comes from regional savings:

```

*
* --- regional savings of the households are defined as the difference between total regional savings and
*   government savings
*
*   rsavc.l(r,"hhsld",t) = rsav(r,t) - rsavg(r,t);
*

```

These savings are distributed to the household types either by the information provided or using uniform rates:

```

*
* --- if information on households's share on total saving are given, apply them
*
*   rsavc.l(r,hhslds,t) $ shrHhslds(r,hhslds,"savings",t)      = rsavc.l(r,"hhsld",t) * shrHhslds(r,hhslds,"savings",t);
*
* --- otherwise, use average regional saving rate
*
*   rsavc.l(r,hhslds,t) $ (not shrHhslds(r,hhslds,"savings",t))
*   = rsavc.l(r,"hhsld",t)*(factYh(r,hhslds,t) - sum(f,dirTax(r,f,hhslds,t)))/(factY(r,t)-yTax(r,"dt",t));

```

From there, saving rates per household type are calculated, and the private consumption spending is calculated as well as the related value share:

```

*
* --- calculate savings rates for the households
*
*   betaSh.l(r,h,t) $ factYh(r,h,t) = rsavc(r,h,t)/(factYH(r,h,t) - sum(f,dirTax(r,f,h,t)));
*
* --- calculate final consumption and final consumption share residually
*
*   yc.l(r,h,t) = factYh(r,h,t) - sum(f,dirTax(r,f,h,t)) - rsavc.l(r,h,t);
*   betaPh.l(r,h,t) $ factYh(r,h,t) = yc(r,h,t)/(factYh(r,h,t) - sum(f,dirTax(r,f,h,t)));

```

The household type use the same consumption tax rates and split-up of domestic and import demand. Note that the information is only introduced if different household types are introduced:

```

:
:   loop(r $ sum( (hhslds,t) , yc.l(r,hhslds,t)) ,
: *
:   --- same prices and tax rates as for aggregate, same consumption composition
: *
:
:   pa.l(r,i,h,t) = pa.l(r,i,"hhsld",t);
:   pdp.l(r,i,h,t) = pdp.l(r,i,"hhsld",t);
:   pmp.l(r,i,h,t) = pmp.l(r,i,"hhsld",t);
:   dintx.l(r,i,h,t) = dintx.l(r,i,"hhsld",t);
:   mintx.l(r,i,h,t) = mintx.l(r,i,"hhsld",t);
:   xa.l(r,i,h,t) = xa.l(r,i,"hhsld",t) * yc.l(r,h,t)/yc.l(r,"hhsld",t);
:   xm.l(r,i,h,t) = xm.l(r,i,"hhsld",t) * yc.l(r,h,t)/yc.l(r,"hhsld",t);
:   xd.l(r,i,h,t) = xd.l(r,i,"hhsld",t) * yc.l(r,h,t)/yc.l(r,"hhsld",t);
:

```

In that case, information on the aggregate household will be deleted:

```

: *
: *      --- delete data for aggregate
: *
:      yc.l(r,"hhsld",t) = 0;
:      xa.l(r,i,"hhsld",t) = 0;
:
:      betaSh.l(r,"hhsld",t) = 0;
:      betaPh.l(r,"hhsld",t) = 0;
:      factYh.l(r,"hhsld",t) = 0;
:      rsavc.l(r,"hhsld",t) = 0;
:      dirTax.l(r,f,"hhsld",t) = 0;
:
:      xaFlag(r,i,h) = xaFlag(r,i,"hhsld");
:      xaFlag(r,i,"hhsld") = 0;

```

## Post processing code

The post processing code basically reverts the code the code from the calibration, i.e. it aggregates from the individual household types to the aggregate private household “hhsld”:

```

if ( sum( (rs,hhslds), pcons.l(rs,hhslds,tSim)),

pdNest.l(rs,"top",h,tsim) = pcons.l(rs,h,tsim);
xdNest.l(rs,"top",h,tsim) $ pcons.l(rs,h,tsim) = yc.l(rs,h,tsim)/pcons.l(rs,h,tsim);

yc.l(rs,"hhsld",t) = sum(hhslds $ pcons.l(rs,hhslds,t), yc(rs,hhslds,t));
xcshr.l(rs,i,"hhsld",t) = sum(hhslds $ xaFlag(rs,i,hhslds), xa(rs,i,hhslds,t)*pa(rs,i,hhslds,t))/yc(rs,"hhsld",t);
xa.l(rs,i,"hhsld",t) = sum(hhslds $ xaFlag(rs,i,hhslds), xa(rs,i,hhslds,t));
xm.l(rs,i,"hhsld",t) = sum(hhslds $ xaFlag(rs,i,hhslds), xm(rs,i,hhslds,t));
xd.l(rs,i,"hhsld",t) = sum(hhslds $ xaFlag(rs,i,hhslds), xd(rs,i,hhslds,t));
factYh.l(rs,"hhsld",t) = sum(hhslds $ pcons.l(rs,hhslds,t), factYh.l(rs,hhslds,t));

pa.l(rs,i,"hhsld",t) $ xa.l(rs,i,"hhsld",t)
= sum(hhslds $ xaFlag(rs,i,hhslds), xa(rs,i,hhslds,t)*pa(rs,i,hhslds,t))/xa(rs,i,"hhsld",t);

pdp.l(rs,i,"hhsld",t) $ xd.l(rs,i,"hhsld",t)
= sum(hhslds $ xaFlag(rs,i,hhslds), xd(rs,i,hhslds,t)*pdp(rs,i,hhslds,t))/xd(rs,i,"hhsld",t);

pmp.l(rs,i,"hhsld",t) $ xd.l(rs,i,"hhsld",t)
= sum(hhslds $ xaFlag(rs,i,hhslds), xm(rs,i,hhslds,t)*pmp(rs,i,hhslds,t))/xm(rs,i,"hhsld",t);

uh.l(rs,"hhsld",t) = sum(hhslds, yc(rs,hhslds,"check")*uh(rs,hhslds,t))/sum(hhslds, yc(rs,hhslds,"check"));
pcons(rs,"hhsld",t) = sum( i,xcshr(rs,i,"hhsld",t)*pa(rs,i,"hhsld",t));
xaFlag(r,i,"hhsld") $ sum(h, xaFlag(r,i,h)) = 1;

betaP(rs,tsim) = yc.l(rs,"hhsld",tsim) / regy.l(rs,tsim);
betaS(rs,tsim) = rsav.l(rs,tsim) / regy.l(rs,tsim);
betaG(rs,tsim) = yg.l(rs,tsim) / regy.l(rs,tsim);

);

```

## 3.2.8 Integration FAO household data into CGEBOX

Wolfgang Britz, October 2019

### Motivation

The following section provides documentation on the integration of FAO household survey data into CGEBox. The use of the data is twofold: first, to define flexibly aggregate household types for use with the myGTAP module and, second, to develop a post-model micro-simulation tool based on the data set. Both are fully operational and encompass the two usual ways household survey data sets are used in CGE work.

Post-model simulation does not require a household differentiation in the CGE model itself. Thus, it has the advantage of keeping the structure of the CGE model separated from the specifics of the household data set. Indeed, the micro-simulation itself can be conducted completely independent from the software code of the CGE model. This argument is not of relevance here as we integrate the post-model micro-simulation transparently into the code of CGEBox itself. Introducing (aggregate) households in the model structure which manage their own factor stocks to maximize revenues can add depth to the analysis which might be missed in post-model simulation. Equally, household

demand behaviour might be more appropriate depicted if aggregate household are introduced in the CGE itself. CGEBox allows now both, also in combination.

### Description of the data set

The FAO household data set<sup>18</sup> covers 19 countries (Kenya, Ethiopia, Malawi, Tanzania, Niger, Uganda, Nigeria, Ghana, Bangladesh, Vietnam, Nepal, Cambodia, Indonesia, Nicaragua, Guatemala, Bolivia, Panama, Albania, Tajikistan), in many cases, also the region inside the country in which the household is located is available. In total, the data set covers more than 193tsd households. The data had been collected between 1992 and 2013 (see Table 4). While there are in total more than 500 items, many of the items are comprised only in a specific sample such as e.g. household possession of certain equipment. The common set covers 115 variables, mostly related to income composition and some classification variables (rural / urban, male or female head, household size etc.).

**Table 4: Data coverage**

| Country    | years          | # of households | # of regions |
|------------|----------------|-----------------|--------------|
| Kenya      | 2005           | 12547           | 8            |
| Ethiopia   | 2012           | 3789            | 10           |
| Malawi     | 2004, 2011     | 23431           | 3            |
| Tanzania   | 2009,2011,2013 | 11923           | 26           |
| Niger      | 2011           | 3910            | 8            |
| Uganda     | 2005,2012      | 8259            | 4            |
| Nigeria    | 2020,2013      | 4936            | 6            |
| Ghana      | 2005,2013      | 24626           | 10           |
| Bangladesh | 2005           | 10058           | 6            |
| Vietnam    | 1992,2002,2008 | 43406           | 8            |
| Nepal      | 2003           | 3908            | 5            |
| Cambodia   | 2004           | 14978           | 5            |
| Indonesia  | 2000           | 10063           | 19           |
| Nicaragua  | 1998,2005      | 10900           | 4            |
| Guatemala  | 2006           | 13640           | 8            |
| Bolivia    | 2005           | 4021            | 9            |
| Panama     | 2003           | 6314            | 6            |
| Albania    | 2005           | 3818            | 4            |
| Tajikistan | 2003,2007      | 7697            | 5            |

<sup>18</sup>

[http://www.fao.org/fileadmin/user\\_upload/smallholders\\_dataportrait/docs/Data\\_portrait\\_variables\\_description\\_new2.pdf](http://www.fao.org/fileadmin/user_upload/smallholders_dataportrait/docs/Data_portrait_variables_description_new2.pdf). Thanks to the FAO for making this highly valuable data set available for research.

### Income generation

We focus here mostly on income generation – how the income of the household is composed – at it is one main link of household survey data to CGE model results. As the data set is tailored to FAO needs and covers developing countries, it takes care to reflect the informal character of the agricultural sector in developing countries with subsistence features. This explains partly the top-level income categories found in the data set:

**Table 5: Top level categories in income generation**

| Category label    | Explanation                                                                       |
|-------------------|-----------------------------------------------------------------------------------|
| Selfemp           | Income (partly imputed) from being self-employed outside the agriculture sector   |
| Livstincome1gross | Income from livestock (mostly imputed, comprises the value of own-consumed crops) |
| Cropinome1gross   | Income from crops (mostly imputed, comprises the value of own-consumed crops)     |
| Transfer gross    | Income from private and government transfers                                      |
| agr_wge           | Wage income from agriculture                                                      |
| Nonagr_wge        | Wage income non-agriculture                                                       |
| Other             | Mainly farm and non-farm rents                                                    |

Alternatively, these items are expressed as shares on total income which might ease data processing. Checks have shown that the shares generally add up to unity while adding the absolute might show some imbalances. For the work with CGEBox, using shares is hence the implemented approach.

**Table 6 Top level share categories**

| Category label | Explanation                                            |
|----------------|--------------------------------------------------------|
| Sh_agin        | Income from agriculture (crops, livestock, farm rents) |
| shagwge        | Wages from agriculture                                 |
| shoffarm       | Wages and self-employed non-agriculture                |
| Shtrans        | Income from private and government transfers           |
| shother        | Mainly non-farm rents                                  |

The shares of the top-level categories in Table 6 are broken down more detailed ones as shown in Table 7. The details allow linking the different wage and self-employed income categories of the households to some aggregate sectors in the economy, and in case of wages, also to two aggregate labor skill categories. That clearly adds depth to the link between the micro-data set and the macro-simulation in the CGE model.

**Table 7: Second level categories shares**

| Category label | Explanation                                                                     | Part of |
|----------------|---------------------------------------------------------------------------------|---------|
| Selfemp        | Income (partly imputed) from being self employed outside the agriculture sector |         |

|                                           |                                                                                                                                                                                                                                    |          |
|-------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| Livstincome1gross                         | Income from livestock including wages                                                                                                                                                                                              | sh_agin  |
| Cropinomce1gross                          | Income from crops including wages                                                                                                                                                                                                  | sh_agin  |
| Sh1privtransfer                           | Income from private transfers                                                                                                                                                                                                      | shtrans  |
| Sh1pubtransfer                            | Income from public transfers                                                                                                                                                                                                       | shtrans  |
| sh1wge2_1..3 ..<br>shewge10_1..3          | Wage income, different sectors (mining, manufacturing, electricity/utilities, constructions, commerce, transport/storage/communication, finance/insurance/real estate, other serices, other); _1: skilled, _2: unskilled, _3:other | Shoffarm |
| Sh1selfemp<br>(she2Self ...<br>self1self) | Non-ag self-employed income, sectoral break as above                                                                                                                                                                               | Shoffarm |

### Data processing

The data in STATA format were converted by Olexandr Nekhay into CSV format. In order to import the data into GAMS, the CSV2GDX conversion routine was used:

```
$call csv2gdx "fao_hhsld.csv" useHeader=y id=p_dataFao index=2,4,7,6 values=8..lastCol acceptBadUels=Y
```

The outcome is a large GDX container with about 18 Mio non-zeros. It comprises a five-dimensional parameter with the household observations (country, region, household identifier, year, item). In order to speed up processing, first three cross sets were generated which show the relation between the country and the year, the country and the household id and the country and region for which data are available:

```
set country_hhid(country,hhid);
country_hhid(country,hhid) $ sum( (region,year), p_dataFao(country,region,hhid,year,"totIncome")) = YES;

set country_year(country,year);
country_year(country,year) $ sum( (region,hhid), p_dataFao(country,region,hhid,year,"totIncome")) = YES;

set country_region(country,region);
country_region(country,region) $ sum( (hhid,year), p_dataFao(country,region,hhid,year,"totIncome")) = YES;
```

Next, we compact the data by removing the year and region dimensions from the data set which are not needed for further applications and store the resulting parameter and the cross-sets to disk for later reuse:

```
p_compact(country_hhid(country,hhid),datVars)
= sum( (country_region(country,region),country_year(country,year)),
       p_dataFao(country,region,hhid,year,datVars));

execute_unload "compact.gdx" p_compact,country_hhid,country,hhid,region,year;
```

It should be noted that the household ids are unique for each country. The same id might however be found in several countries. Equally, one household might be found in multiple years.

### Linking the data set to the GTAP data

The income categories reported in the FAO data set do not match directly the sector resolution of GTAP. Similarly, the income definition comprises elements such as private and government transfers for which no data are available in the standard GTAP data set. The following paragraph briefly

documents how we derive form the income shares data to be used with the myGTAP module in CGEBox.

In order to link wage income to the GTAP data, we link the shares of wages as reported in the FAO household survey to GTAP labour skill categories and GTAP sectors:

```
set datvars_1_a(datVars,1,a) /

    shlwge2_1.unSkLab .(coa-a,oil-a,gas-a,omn-a)
    shlwge2_2.skLab   .(coa-a,oil-a,gas-a,omn-a)
    shlwge2_3.unskLab .(coa-a,oil-a,gas-a,omn-a)

    shlwge3_1.unSkLab .set.manuf
    shlwge3_2.skLab   .set.manuf
    shlwge3_3.unskLab .set.manuf

    shlwge4_1.unSkLab .set.eleDist
    shlwge4_2.skLab   .set.eleDist
    shlwge4_3.unskLab .set.eleDist

    shlwge5_1.unSkLab .cns-a
    shlwge5_2.skLab   .cns-a
    shlwge5_3.unskLab .cns-a

    shlwge6_1.unSkLab .trd-a
    shlwge6_2.skLab   .trd-a
    shlwge6_3.unskLab .trd-a

    shlwge7_1.unSkLab .set.trsStoCom
    shlwge7_2.skLab   .set.trsStoCom
    shlwge7_3.unskLab .set.trsStoCom

    shlwge8_1.unSkLab .set.FinInsReal
    shlwge8_2.skLab   .set.FinInsReal
    shlwge8_3.unskLab .set.FinInsReal

    shlwge9_1.unSkLab .set.services
    shlwge9_2.skLab   .set.services
    shlwge9_3.unskLab .set.services

/;
```

The sets used such as *set.manuf* list simply the GTAP sectors. A similar cross set is used to link sector to self-employed earning shares:

```
set datvars_self(datVars,a) /
    sh1Self2.(coa-a,oil-a,gas-a,omn-a)
    sh1Self3.set.manuf
    sh1Self4.set.eleDist
    sh1Self5.cns-a
    sh1Self6.trd-a
    sh1Self7.set.trsStoCom
    sh1Self8.set.FinInsReal
    sh1Self9.set.services
/;
datvars_self("sh1Self10",nonAgr) = YES;
```

These relations are needed both for the application of the myGTAP module and for the post-model micro-simulation.

### Defining aggregate household types for the myGTAP module of CGEBox

The definition of household types can be managed via the interface where up to three classification

criteria can be defined:

The screenshot shows a software interface titled "Use FAO household Data" with a checked checkbox. It contains three rows for defining classification criteria:

| Classification criterion   | Variable | Value                         |
|----------------------------|----------|-------------------------------|
| Classification criterion 1 | IncPerCa | q 25,50,75,100                |
| Classification criterion 2 | urban    | v rural 0, urban 1            |
| Classification criterion 3 | sh_Agin  | v "0-0.5";0.5-0.75;0.75-1.00" |

The first field captures the name of the up to three items from the FAO household surveys of which the numerical values are used to assign households to classes. As shown in the second example, that can include flags such as rural/urban, male/female household head etc.

The second field defines if the limits entered should be interpreted directly (v=value) or defines quantiles (q). If quantiles are used, the household data are sorted (unweighted) by the item's numerical values and percentiles of these values define country specific class limits. The third field defines the numerical limits for the classes. Here, the following points need to be noted:

1. Classes must be ordered from low to high. (Attention here for flags, an input such as "female 1, male 0" will not produce the correct classification).
2. The last text field defines the elements of a set. Hence, the inputs must be valid GAMS syntax for the text between "SET /" and the closing "/;". That especially means that limits defined as floats with a decimal point must be surrounded by quotation marks as shown in the example for "sh\_agin" (share of agricultural income). Otherwise, GAMS will assume that an element of a two-dimensional cross-set will be defined. Due to the interaction of the Java and GAMS code in the GUI, if quotation marks are used for one of the set elements, the whole text field must be surrounded by a second type of quotation marks as well.
3. If both a label and an explanatory text are given, the second element is interpreted as the numerical threshold. That is useful for cases which indicate flags as in the second field in the example above the label of the household will be comprise texts instead of "0" or "1".
4. If a '-' is found in a label, the limit is automatically derive from the number after the '-', as shown in the "sh\_Agin" example above.
5. If only numbers are given, the name of the item is used as a prefix to ease reading the household names.

The example above will define a set with  $4 \times 2 \times 3 = 24$  household types. In order to reduce manual coding efforts and allow for a GUI driven definition of such classifications, embedded Python code defines the necessary cross-sets and defines sets such as shown in the example below.

The Python code firstly constructs easier to read set-definitions from the user input, for the example above, they look like:

---

```
set incPerCap/  
"incPerCap 25"  
"incPerCap 50"  
"incPerCap 75"  
"incPerCap 100"  
/  
set urban/  
"rural"  
"urban"  
/  
set sh_Agin/  
"incPerCap 0-0.5"  
"incPerCap 0.5-0.75"  
"incPerCap 0.75-1.00"  
/;
```

Secondly, it defines the names of the aggregate household types:



```
SET set_hhslsds /
  "incPerCap 25_rural_sh_Agin 0-0.5"
  "incPerCap 25_rural_sh_Agin 0.5-0.75"
  "incPerCap 25_rural_sh_Agin 0.75-1.00"
  "incPerCap 25_urban_sh_Agin 0-0.5"
  "incPerCap 25_urban_sh_Agin 0.5-0.75"
  "incPerCap 25_urban_sh_Agin 0.75-1.00"
  "incPerCap 50_rural_sh_Agin 0-0.5"
  "incPerCap 50_rural_sh_Agin 0.5-0.75"
  "incPerCap 50_rural_sh_Agin 0.75-1.00"
  "incPerCap 50_urban_sh_Agin 0-0.5"
  "incPerCap 50_urban_sh_Agin 0.5-0.75"
  "incPerCap 50_urban_sh_Agin 0.75-1.00"
  "incPerCap 75_rural_sh_Agin 0-0.5"
  "incPerCap 75_rural_sh_Agin 0.5-0.75"
  "incPerCap 75_rural_sh_Agin 0.75-1.00"
  "incPerCap 75_urban_sh_Agin 0-0.5"
  "incPerCap 75_urban_sh_Agin 0.5-0.75"
  "incPerCap 75_urban_sh_Agin 0.75-1.00"
  "incPerCap 100_rural_sh_Agin 0-0.5"
  "incPerCap 100_rural_sh_Agin 0.5-0.75"
  "incPerCap 100_rural_sh_Agin 0.75-1.00"
  "incPerCap 100_urban_sh_Agin 0-0.5"
  "incPerCap 100_urban_sh_Agin 0.5-0.75"
  "incPerCap 100_urban_sh_Agin 0.75-1.00"
/;
```

Third, it defines the cross-set which links the attribute values of each item to the household type:

```
...
$setglobal hhAggregSet hhlslsds_incPerCap_urban_sh_Agin(hhAgg,incPerCap,urban,sh_Agin)
SET hhlslsds_incPerCap_urban_sh_Agin(*,incPerCap,urban,sh_Agin) /
  "incPerCap 25_rural_sh_Agin 0-0.5"."incPerCap 25"."rural"."sh_Agin 0-0.5"
  "incPerCap 25_rural_sh_Agin 0.5-0.75"."incPerCap 25"."rural"."sh_Agin 0.5-0.75"
  "incPerCap 25_rural_sh_Agin 0.75-1.00"."incPerCap 25"."rural"."sh_Agin 0.75-1.00"
  "incPerCap 25_urban_sh_Agin 0-0.5"."incPerCap 25"."urban"."sh_Agin 0-0.5"
  "incPerCap 25_urban_sh_Agin 0.5-0.75"."incPerCap 25"."urban"."sh_Agin 0.5-0.75"
  "incPerCap 25_urban_sh_Agin 0.75-1.00"."incPerCap 25"."urban"."sh_Agin 0.75-1.00"
  "incPerCap 50_rural_sh_Agin 0-0.5"."incPerCap 50"."rural"."sh_Agin 0-0.5"
  "incPerCap 50_rural_sh_Agin 0.5-0.75"."incPerCap 50"."rural"."sh_Agin 0.5-0.75"
  "incPerCap 50_rural_sh_Agin 0.75-1.00"."incPerCap 50"."rural"."sh_Agin 0.75-1.00"
  "incPerCap 50_urban_sh_Agin 0-0.5"."incPerCap 50"."urban"."sh_Agin 0-0.5"
  "incPerCap 50_urban_sh_Agin 0.5-0.75"."incPerCap 50"."urban"."sh_Agin 0.5-0.75"
  "incPerCap 50_urban_sh_Agin 0.75-1.00"."incPerCap 50"."urban"."sh_Agin 0.75-1.00"
  "incPerCap 75_rural_sh_Agin 0-0.5"."incPerCap 75"."rural"."sh_Agin 0-0.5"
  "incPerCap 75_rural_sh_Agin 0.5-0.75"."incPerCap 75"."rural"."sh_Agin 0.5-0.75"
  "incPerCap 75_rural_sh_Agin 0.75-1.00"."incPerCap 75"."rural"."sh_Agin 0.75-1.00"
  "incPerCap 75_urban_sh_Agin 0-0.5"."incPerCap 75"."urban"."sh_Agin 0-0.5"
  "incPerCap 75_urban_sh_Agin 0.5-0.75"."incPerCap 75"."urban"."sh_Agin 0.5-0.75"
  "incPerCap 75_urban_sh_Agin 0.75-1.00"."incPerCap 75"."urban"."sh_Agin 0.75-1.00"
  "incPerCap 100_rural_sh_Agin 0-0.5"."incPerCap 100"."rural"."sh_Agin 0-0.5"
  "incPerCap 100_rural_sh_Agin 0.5-0.75"."incPerCap 100"."rural"."sh_Agin 0.5-0.75"
  "incPerCap 100_rural_sh_Agin 0.75-1.00"."incPerCap 100"."rural"."sh_Agin 0.75-1.00"
  "incPerCap 100_urban_sh_Agin 0-0.5"."incPerCap 100"."urban"."sh_Agin 0-0.5"
  "incPerCap 100_urban_sh_Agin 0.5-0.75"."incPerCap 100"."urban"."sh_Agin 0.5-0.75"
  "incPerCap 100_urban_sh_Agin 0.75-1.00"."incPerCap 100"."urban"."sh_Agin 0.75-1.00"
/;
```

The Python code also automatically defines the thresholds for the active classification criteria:

```

--F --
PARAMETER p_incPerCap(microR,incPerCap) /
(set.microR)."incPerCap 25" 25
(set.microR)."incPerCap 50" 50
(set.microR)."incPerCap 75" 75
(set.microR)."incPerCap 100" 100
/;
PARAMETER p_urban(microR,urban) /
(set.microR)."rural" 0
(set.microR)."urban" 1
/;
PARAMETER p_sh_Agin(microR,sh_Agin) /
(set.microR)."sh_Agin 0-0.5" 0.5
(set.microR)."sh_Agin 0.5-0.75" 0.75
(set.microR)."sh_Agin 0.75-1.00" 1.00
/;

```

In case of percentiles, such as for the item “incPerCap”, the values are replaced based on the statistical analysis. Again, to speed up processing, we define mapping sets for each classification criteria which indicates to which class a certain country household combination belongs. The code shown below uses globals (not shown) as the names of the fields are defined via the GUI and not hard-coded in the GAMS code itself:

```

set country_hhid_incPerCap(microR,hhid,incPerCap);
country_hhid_incPerCap(microR,hhid,incPerCap)
= sum( country_hhid(microR,hhid) $ ( ((p_dataCompact(microR,hhid,"incPerCap") gt p_incPerCap(incPerCap-1)) or (incPerCap.pos eq 1))
and (p_dataCompact(microR,hhid,"incPerCap") le p_incPerCap(incPerCap) ) ),1);

set country_hhid_urban(microR,hhid,urban);
country_hhid_urban(microR,hhid,urban)
= sum( country_hhid(microR,hhid) $ ( ((p_dataCompact(microR,hhid,"urban") gt p_urban(urban-1)) or (urban.pos eq 1))
and (p_dataCompact(microR,hhid,"urban") le p_urban(urban) ) ),1);

set country_hhid_femHead(microR,hhid,femHead);
country_hhid_femHead(microR,hhid,femHead)
= sum( country_hhid(microR,hhid) $ ( ((p_dataCompact(microR,hhid,"femHead") gt p_femHead(femHead-1)) or (femHead.pos eq 1))
and (p_dataCompact(microR,hhid,"femHead") le p_femHead(femHead) ) ),1);

```

Next, we aggregate the household data using the aggregation weights. That requires linking each single household entry hhid to the aggregate category hhagg.:

```

parameter p_HHDataFao(microR,*,datVars);

p_HHDataFao(microR,hhagg,datVars) $ (not sameas(datVars,"weight"))
= sum( (hhllds_incPerCap_urban_femHead(hhAgg,incPerCap,urban,femHead), country_hhid_incPerCap(microR,hhid,incPerCap)
,country_hhid_urban(microR,hhid,urban)
,country_hhid_femHead(microR,hhid,femHead)
),
p_dataCompact(microR,hhid,datVars)*p_dataCompact(microR,hhid,"weight"));

```

We need a separate statement for the weights:

```

p_HHDataFao(microR,hhagg,"weight")
= sum( (hhllds_incPerCap_urban_femHead(hhAgg,incPerCap,urban,femHead), country_hhid_incPerCap(microR,hhid,incPerCap)
,country_hhid_urban(microR,hhid,urban)
,country_hhid_femHead(microR,hhid,femHead)
),
p_dataCompact(microR,hhid,"weight"));

```

After these statements, the data for the aggregate household comprise totals, i.e. the sum of the single observations multiplied with the weights. In order to express the items on a per household basis, we have to divide the sum of the weights. Equally, in order to ease inspecting the data, we express the weights for each aggregate household as percentages on the total:

```

p_agg(country,"all","weight") = sum(incLevl, p_agg(country,incLevl,"weight"));

p_agg(country,incLevl,datVars) $ (p_agg(country,incLevl,"weight") and (not sameas(datVars,"weight")))
= p_agg(country,incLevl,datVars) / p_agg(country,incLevl,"weight");

p_agg(country,incLevl,"weight")
= p_agg(country,incLevl,"weight") / p_agg(country,"all","weight")*100;

```

The resulting data set was checked for the exhaustion conditions discussed above, i.e. total income adding up and adding up of various share positions. The resulting data look like (only a selection):

|                                       | shincprod | sh_agin | shagwge | shoffarm | sl |
|---------------------------------------|-----------|---------|---------|----------|----|
| incPerCap 75_rural_sh_Agin 0.75-1.00  | 0.84      | 0.93    | 0.03    | 0.01     |    |
| incPerCap 75_urban_sh_Agin 0-0.5      | 0.06      | 0.07    | 0.13    | 0.68     |    |
| incPerCap 75_urban_sh_Agin 0.5-0.75   | 0.58      | 0.59    | 0.07    | 0.14     |    |
| incPerCap 75_urban_sh_Agin 0.75-1.00  | 0.81      | 0.90    | 0.06    | 0.01     |    |
| incPerCap 100_rural_sh_Agin 0-0.5     | 0.12      | 0.13    | 0.11    | 0.72     |    |
| incPerCap 100_rural_sh_Agin 0.5-0.75  | 0.55      | 0.62    | 0.12    | 0.20     |    |
| incPerCap 100_rural_sh_Agin 0.75-1.00 | 0.77      | 0.94    | 0.02    | 0.01     |    |
| incPerCap 100_urban_sh_Agin 0-0.5     | 0.03      | 0.03    | 0.03    | 0.86     |    |
| incPerCap 100_urban_sh_Agin 0.5-0.75  | 0.59      | 0.62    | 0.08    | 0.19     |    |
| incPerCap 100_urban_sh_Agin 0.75-1.00 | 0.63      | 0.94    | 0.04    | 0.00     |    |
| all                                   |           |         |         |          |    |

That approach offers quite some flexibility to define household types from the FAO data set without additional coding efforts. For more specific tasks, the code can be used as the basis to develop own aggregation routines which can be called from a file stored under “\scen\myGtap”.

### Integrating the data into the myGTAP accounting framework

Household surveys are a rich source to introduce household specific transactions into a CGE model, but their accounting logic often poorly matches the representation of transactions in a SAM. We present in the following an approach which tries to carry a larger share of the data reported on the earning side from the FAO’s household surveys into the myGTAP module of CGEBox.

The myGTAP module links primary factor earnings from the different production activities to the household by letting them manage their own primary factor stocks. The management is depicted in a stylized way by a revenue maximization problem represented by Constant Elasticity of Transformation (CET) function which captures the different real-world phenomena which render factor mobility sluggish. For instance, that means that households will not (and cannot) simply switch from being employed in one sector to another if wage rates change. Furthermore, it allows for household specific saving rates, direct tax rates and can depict government transfers as well as transfers between household types.

In order to determine spending behaviour, we use the AIDADS demand system built into CGEBox to estimate income dependent changes in the demand shares given the national averages found in the GTAP data base.

### Factor income and link to activities

The household survey data prepared as discussed above report shares of revenue categories (wages, public and private transfers, income from agriculture etc.) on total household income, total household income, household size and its relative weight.

## 1. Agricultural income

Agricultural income is reported for crops and livestock activities. Using the household total income and the reported share for crop or livestock activities should match conceptually the total primary factor income from these activities in the SAM. We hence assign shares of these factor incomes based on the given weight (total income times shares from crop or livestock activities).

In order to produce a better match to the SAM, we first calculate the absolute income for each aggregate household and scale the result to yield shares. To do so, we multiply the per-household income with its population weight which should give the economic weight on total income. The total for each aggregate household is then divided by the sum over all household to yield shares:

```
*
* --- income, converted to shares
*
p_hhDataFao(faoR,hhslds,"incWeight") = p_hhDataFao(faoR,hhslds,"weight") * p_HHDataFao(faoR,hhslds,"totIncome");
p_hhDataFao(faoR,hhslds,"incWeight")
  = p_hhDataFao(faoR,hhslds,"incWeight")/sum(hhslds0,p_hhDataFao(faoR,hhslds0,"incWeight"));
```

Based on these shares, we allocate distribution shares for the agricultural primary factors for crop and livestock activities to the households:

```
*
* ---
* --- assign income from crop and livestock production
*
shrXf(faoR,hhslds,f,crop_a,t0) $ xfFlag(faoR,f,crop_a)
  = p_HHDataFao(faoR,hhslds,"sh1Crop1") * p_HHDataFao(faoR,hhslds,"incWeight");

shrXf(faoR,hhslds,f,lstk,t0) $ xfFlag(faoR,f,lstk)
  = p_HHDataFao(faoR,hhslds,"sh1Livestock") * p_HHDataFao(faoR,hhslds,"incWeight");
```

## 2. Wages

The reported income from wages by skill category in different sectors (total factor income times reported share from wages) is used as a distribution key for wage income from the different sectors:

```
*
* --- assign wages from skilled and unskilled labour
*
shrXF(faoR,hhslds,l,a,t0) $ xfFlag(faoR,l,a)
  = shrXF(faoR,hhslds,l,a,t0)
    + sum(datvars_l_a(wages,l,a), p_HHDataFao(faoR,hhslds,wages)) * p_HHDataFao(faoR,hhslds,"incWeight");
```

## 3. Income from self-employed:

The same logic as for wages is used:

```
*
* --- assign factor income for self employed
*
shrXF(faoR,hhslds,f,a,t0) $ xfFlag(faoR,f,a)
  = shrXF(faoR,hhslds,f,a,t0)
    + sum(datvars_self(self,a), p_HHDataFao(faoR,hhslds,self)) * p_HHDataFao(faoR,hhslds,"incWeight");
```

If there are still missing shares, we use total non-farm income:

```
*
* --- if shares are still missing, use off farm income
*
shrXF(faoR,hhslds,f,a,t0) $ (xfFlag(faoR,f,a) and (not sum(hhslds0, shrXF(faoR,hhslds0,f,a,t0))))
  = p_HHDataFao(faoR,hhslds,"incWeight") * p_HHDataFao(faoR,hhslds,"shOffFarm");
```

We also opted to introduce some assumptions on income dependent factor specific direct tax rates with the following parameter which depicts differences in tax rates. The first assumption is that households which receive a higher share of public transfer than the average across all households will pay lower direct tax rates:

```

parameter p_meanTransFer(r);

p_meanTransfer(faoR) = sum(hhsls, p_HHDataFao(faoR,hhsls,"shipubTransfer")*p_HHDataFao(faoR,hhsls,"totIncome"));

shrHhsls(faoR,hhsls,"dirTax",t0) = eps;

shrHhsls(faoR,hhsls,"dirTax",t0) $ p_HHDataFao(faoR,hhsls,"popWeight")
= ( 1 - (p_HHDataFao(faoR,hhsls,"shipubTransfer")-p_meanTransfer(faoR)) ) * p_HHDataFao(faoR,hhsls,"popWeight");

```

We estimate differences in saving rates based on the regression for macro-saving rates as part of G-RDEM, making sure that we don't yield negative saving rates and start from a minimum level of 5%. As the maximum rate found in the sample use for the regression where below 40%, we restrict the initialization to 50%:

```

shrHhsls(faoR,hhsls,"savings",t0) $ p_HHDataFao(faoR,hhsls,"popWeight")
= p_HHDataFao(faoR,hhsls,"incWeight") *
min(0.5,max(1.E-3,max(0.05,betas.l(faoR,t0)) + (p_HHDataFao(faoR,hhsls,"incPerCap") - p_avgIncome(faoR))*0.0010213/1000))
/max(0.05,betas.l(faoR,t0));

```

## Transfers

The data set reports the shares of private and public transfers received on income. The data cannot be easily used to assign some numbers. Instead a constrained optimization problem is defined. Its aim is not only to define the transfers, but also to generate income levels which in relative terms come close to those reported in the FAO data set.

It comprises the following balancing equations:

1. The reported income share on total income should be recovered, where rs are the regions, hhsls the household and t0 the benchmark period. Household income from the differences sources is defined as hinc and v\_corr are endogenous correction factors to minimize:

```

*
* --- recover given relative incomes
*

e_totInc(rs,t0) ..

v_totInc(rs,t0)
=E= sum(hhsls0 $ p_hhDataFao(rs,hhsls0,"income") , hinc(rs,hhsls0,t0));
e_hincRel(rs,hhsls,t0) $ p_hhDataFao(rs,hhsls,"income") ..

hinc(rs,hhsls,t0)
=E= v_totInc(rs,t0) * p_hhDataFao(rs,hhsls,"income") * v_corr(rs,hhsls,"income");

```

2. Total household income is defined from the given allocated factor income factyH and direct taxes dirTax on primary factors f, plus the endogenously determined transfer from government trng and transfer between households trnh. The variable trnh carries the delivering household on the second index and the receiving household on the third. We also consider remittances received remiH and a share on wages remoh which is sent abroad<sup>19</sup>:

<sup>19</sup> The data on remittances are available at bi-lateral level from the GTAP-MIG data base and define the total of remittances received and sent for each country resp. aggregate of countries.

```

*
* --- factor income plus government plus private transfers
*
e_hinc(rs,hhslds,t0) $ p_hhDataFao(rs,hhslds,"income") ..

    hinc(rs,hhslds,t0) =E= (factYH.1(rs,hhslds,t0)-sum(f,dirTax.1(rs,f,hhslds,t0)))
      + trng(rs,hhslds,t0)
      + remih(rs,hhslds,t0)
      - remoh.1(rs,hhslds,t0)
      + sum(hhslds0, trnh(rs,hhslds0,hhslds,t0))
      - sum(hhslds0, trnh(rs,hhslds,hhslds0,t0))
      ;
*

```

- Income shares from public transfer should be recovered; a correction factor allows for deviations. Note that we don't have these data from the SAM. Instead, they are indirectly constructed by the balancing program.

```

*
* --- relative government transfers
*
e_trngShr(rs,hhslds,t0) $ p_hhDataFao(rs,hhslds,"income") ..

    trng(rs,hhslds,t0)
    =E= [p_hhDataFao(rs,hhslds,"shipubtransfer")+v_corr(rs,hhslds,"trng"]*(hinc(rs,hhslds,t0)+sum(hhslds0, trnh(rs,hhslds,hhslds0,t0)));
*

```

- Transfer income from other households should be recovered, again allowing for deviations based on a correction factor:

```

*
* --- relative private transfers
*
e_trnhShr(rs,hhslds,t0) $ p_hhDataFao(rs,hhslds,"income") ..

    sum(hhslds0, trnh(rs,hhslds0,hhslds,t0))
    =E= [p_hhDataFao(rs,hhslds,"shiprivtransfer")+v_corr(rs,hhslds,"trnh"] * (hinc(rs,hhslds,t0)+sum(hhslds0, trnh(rs,hhslds,hhslds0,t0)));
*

```

- Two equations ensure first that the given remittances are exhausted, and second, that estimated shares come close to given ones:

```

*
* --- remittances
*
e_remitBal(rs,t0) ..

    sum(hhslds $ p_hhDataFao(rs,hhslds,"income"),remih(rs,hhslds,t0)) =E= remih.1(rs,"hhsld",t0);

e_remitShr(rs,hhslds,t0) $ p_hhDataFao(rs,hhslds,"income") ..

    remih(rs,hhslds,t0)
    =E= [p_hhDataFao(rs,hhslds,"remittences")+v_corr(rs,hhslds,"remit"]*(hinc(rs,hhslds,t0)+sum(hhslds0, trnh(rs,hhslds,hhslds0,t0)));
*

```

The above equations define the constraints for a penalty function which minimizes squared differences for the four correction factors, given a high weight for recovering the relative income levels. We add terms which penalize transfer from poorer to richer households.

```

e_objeHHsld ..

    v_objeHHsld * sum( (rs,hhslds,t0) $ p_hhDataFao(rs,hhslds,"income"),1)

    =E= sum( (rs,hhslds,t0) $ p_hhDataFao(rs,hhslds,"income"),

        sqr( v_corr(rs,hhslds,"trnh"))
      + sqr( v_corr(rs,hhslds,"trng"))
      + sqr( v_corr(rs,hhslds,"remit"))
      + sqr( v_corr(rs,hhslds,"income")-1)

        --- transfers from poorer to richer households

      + sum(hhslds0
        $ (( p_hhDataFao(rs,hhslds,"income")/p_HHDataFao(rs,hhslds,"population"))
          < (p_hhDataFao(rs,hhslds0,"income")/p_HHDataFao(rs,hhslds0,"population")))
        $ p_hhDataFao(rs,hhslds0,"income")) ,
        trnh(rs,hhslds,hhslds0,t0)/hinc(rs,hhslds,t0) )

    );

```

**Saving rates and foreign income contributions**

The household surveys do not comprise data on saving rates. Unfortunately, due to the regional household concept, the GTAP data set reports only the total regional savings as the sum of savings (or additional debts) by private households, firms and public institutions. What is available from the SAM are total tax income  $yTaxTot$ , factor income after depreciation  $yFacty$ , direct taxes  $dirtax$  and the private  $yc$  and public consumption  $yg$  expenditures. Equally missing are data of transfers from the government to households  $trng$ . They are not needed for the regional household approach found in the GTAP Standard model as they net out if we add income sources of public and private agents. (The same holds for transfers between different households in the same country).

That means that we know for the government the usual main income source (tax income) and know what it spent for consumption ( $yg$ ), but miss data on savings or transfers to households:

$$\overline{yg} + trng + savg = \overline{yTaxTot} \quad (31)$$

Similar for the private aggregate household, we miss data on its savings (and again on transfers received from government):

$$\overline{yc} + savc = \overline{yFacty} - \overline{dirtax} + trng \quad (32)$$

What is known are total regional savings  $rsav$ :

$$\overline{rsav} = savc + savg \quad (33)$$

From the given balanced data found in the SAM, reflecting the equivalence of regional household income earned (RHS) and spent (LHS), it also holds that:

$$\overline{yg} + \overline{yc} + \overline{rsav} = \overline{yFacty} + \overline{yTaxTot} - \overline{dirtax} \quad (34)$$

These four accounting identities are linear dependent as seen easily from the fact that the government transfers  $trng$  can be set to zero, and  $savg$  and  $savc$  chosen to close the government (1) and private household budgets (2). Adding next (1) and (2) leads over (3) to the identity found in (4).

That implies that we need either additional data or use assumptions to define government transfers  $trng$  and the two savings positions. As governments seldom achieve a budget surplus, a rather simple rule sets  $savg$  to zero in case that total regional saving  $rsav$  is positive. That defines government transfers to households  $trng$  from (1). i.e. tax income is exactly exhausted by public consumption and transfers to households. Identity (32) then reflects that the households performs all regional savings as  $savg$  is zero by assumption. In case of overall negative savings, we set instead  $savc$  to zero (in opposite to government, it is not likely that the aggregate household has a structural deficit), and calculate from there the necessary transfers  $trng$  to close the household's budget. That implies that the total and under that case negative regional savings become in full additional debt of the government.

The balancing becomes only slightly more complex if we take remittances or other income exchanges with the rest-of-the world  $finc$  (for foreign income) into account:

$$\overline{yc} + savc = \overline{yFacty} - \overline{dirtax} + trng + finc \quad (35)$$

These income flows are part of the balance of payment, if introduced in the existing data base, the foreign savings  $fsav$  needs to be corrected by the same amount from the B.O.T. equals B.O.P. condition:

$$\bar{M} - \bar{X} = \text{savf} + \text{finc} \quad (36)$$

As the additional income  $\text{finc}$  in (35) on the RHS will increase savings  $\text{savc}$  of the household on the LHS by the same amount, we also maintain the identity of investments  $I$  and total saving  $S$ , the latter equal to the sum of foreign and regional savings:

$$\bar{I} \equiv \bar{S} = \text{savf} + \text{savc} + \text{savg} \quad (37)$$

Alternative assumptions on the distribution of total regional savings to government and private households are clearly possible as long as the accounting identities are maintained.

### 3.2.9 Post-model micro-simulation

Post-model micro-simulation became fashionable in the first decade of the 2000 (cf. Cockburn 2006<sup>20</sup>). It reflects the growing insights that household are differently affected by policy reforms such as trade liberalization or structural adjustment programs. There are two major pathways how data of individual households are updated in that type of analysis: from the revenue and from the expenditure side. From the revenues side, policy reforms will lead to differentiated impacts on factor returns depending on the factor and the sector. Especially with sluggish factor markets, households can see their factor income develop quite differently depending in which sectors they are engaged, and, from which factors they draw income. Policy reforms can also directly affect factor returns e.g. by changing direct tax rates or factor taxes respectively subsidies faced by the different sectors. Many households depend also on private and public transfers. In the latter case, changes in tax income or how tax income is distributed can affect them. For private transfers, income changes of their donors typically also impact the size of transfers. Impacts from the expenditure side result from changes in consumer prices, driven by changes in domestic and foreign production costs, trade margins and the various taxes and subsidies which link consumer prices to production costs. These general considerations lead to a rather straightforward approach described next which reflects the state-of-the-art in that field.

The post-model micro-simulation builds technically on three steps which reflect the considerations above. First, changes in factor prices and factor quantities as well as tax and regional income from the CGE model simulation are used to update the income position of each household in the FAO household data set. Second, we use the MAIDADS demand system to estimate budget shares; results from the G-RDEM estimation of the saving rates allow to estimate saving rates specific to each household, and finally, we assign to each household a share of government consumption. That allows calculating the money metric for each household. The final step summarizes the results based on a regression of the money metric on income composition and by calculating percentiles of key results and household features, sorted by the money metric.

#### Updating household income

The FAO data set reports shares on total household income for different wage categories, for self-employed, crop and livestock production, public and private transfers and an additional position. We detail next how these different elements are used to update the reported benchmark income for each household.

---

<sup>20</sup> Cockburn, John (2016): "Trade liberalisation and poverty in Nepal: A computable general equilibrium micro-simulation analysis." *Globalisation and Poverty*. Routledge, 2006. 189-212



For wage income, we use changes in factor returns to the specific skill category and the individual GTAP sectors linked to the more aggregate categories in the FAO data set to estimate the change in wage income:

```
*
* --- income from different wage categories: link to change in labour remuneration
*
p_microRes(country_hhid(microR,hhid),wages,tSim)
= p_compact(microR,hhid,"totIncome") * p_compact(microR,hhid,wages)
  * sum( datvars_l_a(wages,l,a) $ xfflag(microR,l,a), xf.l(microR,l,a,tSim) *pf.l(microR,l,a,tSim) * (1-kappaf.l(microR,l,tSim)))
  / sum( datvars_l_a(wages,l,a) $ xfflag(microR,l,a), xf.l(microR,l,a,"%t0%") *pf.l(microR,l,a,"%t0%") * (1-kappaf.l(microR,l,"%t0%")));
```

For instance, for the second wage category linked to extraction, the following sectors are used:

```
set datvars_l_a(datVars,l,a) /

shlwge2_1.unskLab .(coa-a,oil-a,gas-a,omn-a)
shlwge2_2.skLab .(coa-a,oil-a,gas-a,omn-a)
shlwge2_3.unskLab .(coa-a,oil-a,gas-a,omn-a)
```

For self-employed and agriculture income from crop respectively livestock production, all factor returns are used (wages, interest, returns to natural resources and land) as it is probably rather unclear how especially wage income accruing to the household from a farm it manages is accounted for in the SAM and the household survey:

```
*
* --- income from different self-employed categories: link to change in factor remuneration
*
p_microRes(country_hhid(microR,hhid),self,tSim)
= p_compact(microR,hhid,"totIncome") * p_compact(microR,hhid,self)
  * sum( (datvars_self(self,a),f) $ xfflag(microR,f,a), xf.l(microR,f,a,tSim) *pf.l(microR,f,a,tSim) * (1-kappaf.l(microR,f,tSim)))
  / sum( (datvars_self(self,a),f) $ xfflag(microR,f,a), xf.l(microR,f,a,"%t0%") *pf.l(microR,f,a,"%t0%") * (1-kappaf.l(microR,f,"%t0%")));
```

The different aggregates (2-10) for which income from self-employment is reported in the FAO data (such as manufacturing) are mapped to the GTAP sectors:

```
set datvars_self(datVars,a) /
shlSelf2.(coa-a,oil-a,gas-a,omn-a)
shlSelf3.set.manuf
shlSelf4.set.eleDist
shlSelf5.cns-a
shlSelf6.trd-a
shlSelf7.set.trsStoCom
shlSelf8.set.FinInsReal
shlSelf9.set.services
/;
datVars_self("shlSelf10",nonAgr) = YES;
```

Public transfers are updated with changes in total tax income. This assumes that the value share of tax income spent on transfers to each household type is unchanged:

```
*
* --- income from public transfer: tax income
*
p_microRes(country_hhid(microR,hhid),"pubTransfer",tSim)
= p_compact(microR,hhid,"totIncome") * p_compact(microR,hhid,"shlPubTransfer")
  * yTaxTot.l(microR,tSim)/yTaxTot.l(microR,"%t0%");
```

For private transfers, we assume that payments follow the aggregate private consumption changes:

```
*
* --- income from private transfer: private consumption
*
p_microRes(country_hhid(microR,hhid),"privTransfer",tSim)
= p_compact(microR,hhid,"totIncome") * p_compact(microR,hhid,"shlPrivTransfer")
  * yc.l(microR,"hhsld",tSim)/yc.l(microR,"hhsld","%t0%");
```

For the remaining positions under “other” which refer, for instance, to non-farm rental income and remittances (where the data set often fails to provide a detailed composition), we use changes in regional income:

```
*
* --- other income: regional income
*
p_microRes(country_hhid(microR,hhid),"other",tSim)
= p_compact(microR,hhid,"totIncome") * p_compact(microR,hhid,"shlOther")
  * reqv.1(microR,tSim)/reqv.1(microR,"%t0%");
```

After these changes, all income categories found in the FAO data set have been updated based on the CGE model results such that changes in income in real terms from the income generation side are simulated. The notion of real terms reflects that there is not inflation in the CGE model as only relative prices matter for simulated changes in quantities. The usual case in the GTAP model is to use a world factor price index as the numeraire.

### Estimating budget shares, their changes and the money metric

Income changes at inflation free factor prices alone are not sufficient to assess how policy reforms or other changes impact the economic situation of individual household. We need also to reflect which prices they face for the goods they are consuming.

The next step is therefore to estimate the impact of changes in consumer prices. The code is only active for the MAIDADS demand system which is especially suitable to capture impacts of changes in income levels on consumption pattern due to is rather non-linear Engels curves which have been econometrically estimated (Britz 2021<sup>21</sup>). In a first step, we calculate cardinal utility for each household from an econometrically estimated relation to per capita income:

```
p_microRes(country_hhid(microR,hhid),"yc",calcS) $ p_microRes(microR,hhid,"incPerCap",calcS)
= p_microRes(microR,hhid,"incPerCap",calcS) * (p_microRes(microR,hhid,"betap",calcS)+p_microRes(microR,hhid,"betag",calcS));

p_microRes(country_hhid(microR,hhid),"uh",calcS) $ p_microRes(microR,hhid,"yc",calcS)
= (p_estU("cnst") + p_estU("log") * log(p_microRes(microR,hhid,"yc",calcS)/pgdmp.1(micror,calcS))
  + p_estU("logSqr") * sqrt(log(p_microRes(microR,hhid,"yc",calcS)/pgdmp.1(micror,calcS))))
  + p_estU("logSqr") * sqrt(log(p_microRes(microR,hhid,"yc",calcS)/pgdmp.1(micror,calcS)))
  + sum( deltaDrivers, p_estu(deltaDrivers) * p_deltaShifters(micror,deltaDrivers));
```

From there, we define the driver of the exponential functions for the marginal budget shares and commitment terms:

```
p_microRes(country_hhid(microR,hhid),"uhExpD",calcS) $ p_microRes(microR,hhid,"incPerCap",calcS)
= p_parAIDADS("UD","cnst") + v_kappa(microR,"hhsld","UDC")
  + (v_kappa(microR,"hhsld","UDL") + p_parAIDADS("UD","lin")) * p_microRes(microR,hhid,"uh",calcS);

p_microRes(country_hhid(microR,hhid),"uhExpG",calcS) $ p_microRes(microR,hhid,"incPerCap",calcS)
= p_parAIDADS("UG","cnst") + v_kappa(microR,"hhsld","UGC")
  + (v_kappa(microR,"hhsld","UGL") + p_parAIDADS("UG","lin")) * p_microRes(microR,hhid,"uh",calcS);
```

From there, we define the marginal budget share as a linear combination of the marginal budget shares at very low income (= utility) level alphaAIDADS and the ones at very high income (= utility) level betaAIDADS;

```
p_microAlpha(country_hhid(microR,hhid),i,%2) $ p_microRes(microR,hhid,"yc",%2)
= {alphaAIDADS(microR,i,"hhsld")+betaAIDADS(microR,i,"hhsld")} * sqExp(p_microRes(microR,hhid,"uhExpD",%2),10) ]
  / [ 1 + sqExp(p_microRes(microR,hhid,"uhExpD",%2),10) ];

p_microAlpha(country_hhid(microR,hhid),dNest,%2) $ p_microRes(microR,hhid,"yc",%2)
= {alphaAIDADS(microR,dNest,"hhsld")+betaAIDADS(microR,dNest,"hhsld")} * sqExp(p_microRes(microR,hhid,"uhExpD",%2),10) ]
  / [ 1 + sqExp(p_microRes(microR,hhid,"uhExpD",%2),10) ];
```

And similar for the commitment terms:

<sup>21</sup> Britz, W. (2021): Estimating a global MAIDADS demand system considering demography, climate and norms, *Bio-based and Applied Economics*, in: online available 19 July 2021

```

p_microGamma(country_hhid(microR,hhid),i,%2) $ p_microRes(microR,hhid,"yc",%2)
= [gammaLo(microR,i,"hhsld")+gammaHi(microR,i,"hhsld") * sqExp(p_microRes(microR,hhid,"uhExpG",%2),10) ]
/{ 1 + sqExp(p_microRes(microR,hhid,"uhExpG",%2),10) }];

p_microGamma(country_hhid(microR,hhid),dNest,%2) $ p_microRes(microR,hhid,"yc",%2)
= [gammaLo(microR,dNest,"hhsld")+gammaHi(microR,dNest,"hhsld")*sqExp(p_microRes(microR,hhid,"uhExpG",%2),10)]
/{ 1 + sqExp(p_microRes(microR,hhid,"uhExpG",%2),10) }];

```

Given the absolute term of the AIDADS demand system gammaLES, we can define the non-committed income, the i.e. the income distributed by the marginal budget shares defined above<sup>22</sup>:

```

p_microRes(country_hhid(microR,hhid),"nonCommit",%2) $ p_microRes(microR,hhid,"yc",%2)
= max(p_microRes(microR,hhid,"yc",%2)*0.25,p_microRes(microR,hhid,"yc",%2)
- sum(i, p_microGamma(microR,hhid,i,%2)*pa.1(microR,i,"hhsld",%2) )
- sum(dNest, p_microGamma(microR,hhid,dNest,%2)*pdNest.1(microR,dNest,"hhsld",%2)));

```

The non-committed income in combination with the marginal budget shares allow to define the consumer price index for each household:

```

*
* --- consumer price for individual households
*
p_microRes(country_hhid(microR,hhid),"pCons",%2) $ p_microRes(microR,hhid,"incPerCap",%2)
*
* --- Armington demand times price, divided by income = budget share => summed defines the hhsld specific consumer price
*
= [ sum(i,
pa.1(microR,i,"hhsld",%2) * (p_microGamma(microR,hhid,i,%2)
+ p_microAlpha(microR,hhid,i,%2)/pa.1(microR,i,"hhsld",%2)*p_microRes(microR,hhid,"nonCommit",%2)))
+ sum(dNest,
pdNest.1(microR,dNest,"hhsld",%2) * (p_microGamma(microR,hhid,dNest,%2)
+ p_microAlpha(microR,hhid,dNest,%2)/pdNest.1(microR,dNest,"hhsld",%2)*p_microRes(microR,hhid,"nonCommit",%2)))
]
/(p_microRes(microR,hhid,"incPerCap",%2)*p_microRes(microR,hhid,"betap",%2));

```

The contribution to utility from government consumption is based on the simple assumption that each citizen consumes the same share on the total amount:

```

*
* --- distribute government consumption on a per capita basis
*
p_microRes(country_hhid(microR,hhid),"yg",tSim) $ p_microRes(microR,hhid,"incPerCap",tSim)
= yg.1(microR,tSim)*%scale%*1000. / pop.1(microR,"hhsld",tSim);

```

From this assumption follows that the utility of poorer household depends to a larger share on the development of government consumption compared to richer ones:

```

*
* --- calculate implicate share of government consumption on regional income per capita
*
p_microRes(country_hhid(microR,hhid),"betag",tSim) $ p_microRes(microR,hhid,"incPerCap",tSim)
= p_microRes(microR,hhid,"yg",tSim) * (1-p_microRes(microR,hhid,"betas",tSim))
/ (p_microRes(microR,hhid,"incPerCap",tSim)+p_microRes(microR,hhid,"yg",tSim));

```

For savings, we first estimate saving rates depending on per-capita income, using the regression coefficient from the G-RDEM model (Britz and Roson 2019<sup>23</sup>). We cut-off estimates above 50% as the largest macro-saving rate found in any country in the underlying estimation was around 40% while per capita income levels at household level found in the samples easily exceed even the highest mean per capita income found for any country in the GTAP data set used in the econometric analysis. Note also that the CD-utility function does not allow for negative cost shares. We hence exclude cases where the household spends more than he earns (it is not clear if credits fall under other income):

<sup>22</sup> The actual code also considers demand nests.

<sup>23</sup> Britz, W., Roson, R. (2019): "G-RDEM: A GTAP-Based Recursive Dynamic CGE Model for Long-Term Baseline Generation and Analysis", *Journal of Global Economic Analysis* 4(1): 50-96

```

parameter p_avgInc(microR,t);
p_avgInc(microR,tSim) = sum(country_hhid(microR,hhid), p_compact(microR,hhid,"weight"));
p_avgInc(microR,tSim)
= sum(country_hhid(microR,hhid), p_compact(microR,hhid,"weight")*p_microRes(microR,hhid,"incPerCap",tSim))/p_avgInc(microR,tSim);

p_microRes(country_hhid(microR,hhid),"betas",tSim) $ p_microRes(microR,hhid,"incPerCap",tSim)
= min(0.5,max(1.E-3,betas.1(microR,tSim)) + (p_microRes(microR,hhid,"incPerCap",tSim) - p_avgInc(microR,tSim))*0.0010213);

```

With shares of (implicit) income spent to private (betap) and public (betag) consumption and savings (betas) defined, the utility level for each household can be calculated from the indirect utility function of the assumed top-level CD demand system:

```

*
* --- use indirect utility function to derive utility from income and arguments of CD-utility function
*
p_microRes(country_hhid(microR,hhid),"u",tSim) $ p_microRes(microR,hhid,"incPerCap",tSim)
= (p_microRes(microR,hhid,"incPerCap",tSim)+p_microRes(microR,hhid,"yg",tSim))/(1-p_microRes(microR,hhid,"betas",tSim))
* [
  (1/p_microRes(microR,hhid,"betas",tSim))*p_microRes(microR,hhid,"betas",tSim)
  * (1/p_microRes(microR,hhid,"betag",tSim))*p_microRes(microR,hhid,"betag",tSim)
  * (1/p_microRes(microR,hhid,"betap",tSim))*p_microRes(microR,hhid,"betap",tSim)
]
* ( 1 / p_microRes(microR,hhid,"pCons",tSim) ) ** p_microRes(microR,hhid,"betap",tSim)
* ( 1 / pg.1(microR,tSim) ) ** p_microRes(microR,hhid,"betag",tSim)
* ( 1 / pi.1(microR,tSim) ) ** p_microRes(microR,hhid,"betas",tSim);
*

```

Given the utility levels from the indirect CD utility function, we can calculate the money metric as the ultimate measure of changes in the households' purchasing power. The money metric reflects both changes in income (factor income, transfers, other) and the different price indices, considering differences in marginal budget shares across per capita income levels, saving rates and in the importance of government consumption in total household consumption:

```

*
* --- money metric from betas / saving prices and (1-betas) / pcons
*
$batinclude 'util/title.gms' "Calculate money metric for post-model micro-simulation %dataset%, %shock_withoutPath%"

p_microRes(country_hhid(microR,hhid),"au",tSim)
= p_microRes(microR,hhid,"incPerCap",tSim) / [ p_microRes(microR,hhid,"u",tSim)
  * ( 1/(1-max(1.E-3,betas.1(microR,tSim)))**max(1.E-3,betas.1(microR,tSim)) )
  * ( 1/max(1.E-3,betas.1(microR,tSim)) )**max(1.E-3,betas.1(microR,tSim)) ];

p_microRes(country_hhid(microR,hhid),"mm",tSim)
= ( p_microRes(microR,hhid,"u",tSim) - p_microRes(microR,hhid,"u",tSim) ) * p_microRes(microR,hhid,"au",tSim)
  * ( 1 / p_microRes(microR,hhid,"pCons",tSim) ) ** (1-max(1.E-3,betas.1(microR,tSim)))
  * ( pi.1(microR,tSim) / pi.1(microR,tSim) ) ** max(1.E-3,betas.1(microR,tSim));

```

We consider the money metric as the most suitable indicator to assess the well-being of household which can be calculated from the data set.

## Estimating individual demands and nutrient intakes

Given household specific marginal budget shares and commitment terms, the Marshallian demands can be defined, taking potential CES-subnests into account:

```

*
* (1) Directly driven by top-level demand system
*
p_microDem(country_hhid(microR,hhid),i,%2) $ p_hasXA(microR,i,"hhsld")
= p_microGamma(microR,hhid,i,%2) + p_microAlpha(microR,hhid,i,%2)/pa.1(microR,i,"hhsld",%2)
  * p_microRes(microR,hhid,"nonCommit",%2);

*
* (2) Demand nests driven by top-level demand system
*
p_microDem(country_hhid(microR,hhid),dNest,%2) $ sum(dNest_n_fd("top",dNest,h), p_hasXA(microR,dNest,"hhsld"))
= p_microGamma(microR,hhid,dNest,%2) + p_microAlpha(microR,hhid,dNest,%2)/pdNest.1(microR,dNest,"hhsld",%2)
  * p_microRes(microR,hhid,"nonCommit",%2);

*
* (2) In other cases, the homothecy of the CES sub-nesting implies that the demand (at equal prices across households!)
* depend solely on the demand for the nests in the top level demand function
*
p_microDem(country_hhid(microR,hhid),i,%2) $ (xaFlag(microR,i,"hhsld") and (not p_hasXA(microR,i,"hhsld")))
= xa.1(microR,i,"hhsld",%2)
  * sum(topDNest(dNest,i,"hhsld"), p_microDem(microR,hhid,dNest,%2)/xdNest.1(microR,dNest,"hhsld",%2));

```

Given the estimated nutrient contents for the products, nutrient intake per capita and day for calories, protein and fat can be estimated:

```

*
* --- calculate nutrient uptake
*
p_microRes(country, hhid(microR, hhid), nut, %2) $ p_microRes(microR, hhid, "incPerCap", %2)
= sum(i $ p_nutContFabio(microR, i, nut) , [ p_microDem(microR, hhid, i, %2)*0.001
+ ( sum(gov $ xaFlag(microR, i, gov), xa.1(microR, i, gov, %2))
+ sum(inv $ xaFlag(microR, i, inv), xa.1(microR, i, inv, %2))) / pop(microR, "hhsld", %2) ]
* p_nutContFabio(microR, i, nut));

```

## Analysing the micro-simulation results

The samples for the different countries cover each several thousand observations such that it makes limited sense to analyse outcomes for each household separately. Instead, we first summarize findings in percentiles. To do so, we sort the households by their money metric, i.e. their changes in per capita purchasing power, and report for each quantile average earning shares and other results. We need to use the weights reported for each household. As GAMS is not designed for statistical analysis, especially reflecting the weights leads to somewhat clumsy code. We still consider it advantageous to stick to GAMS instead of moving data between different software packages.

As a first step, we use the GDXRank utility to define a sort index. As GDXRank works with a vector, we have to copy the money metric results for each household from the multi-dimensional parameter `p_microRes` to a temporary one-dimensional parameter `p_temp`. The permutation index returned `p_index` is used to define lags and leads in order to generate a sorted copy of the results of interest, as represented by the set `res`.

```

p_temp(curHhid) = p_microRes(microR, curHhid, "mm", tSim);
execute_unload "rank_in.gdx", p_temp;
execute 'gdxrank rank_in rank_out.gdx > %system.nullfile%';
execute_load "rank_out.gdx", p_index = p_temp;
p_index(curHhid) = p_index(curHhid)-ord(curHhid);
p_Sorted(curHhid + p_index(curHhid), res) = p_microRes(microR, curHhid, res, tSim);

```

Next, we sum up as the weights, and set the threshold for the first percentile as 1% of the total weights:

```

sumOfWeights = sum(curHhid, p_compact(microR, curHhid, "weight"));
p_threshHold = sumOfWeights/100;
count = 1;

```

In a loop over all households in the country, we add up the weights over all households `p_curSumW` as well the weights `p_curSumWQ` and variables of interest multiplied with their weights `p_curSumQ` for the next quantile to calculate. If the sum of all weights exceeds the next threshold, we calculate the averages for the result variables by dividing by the sum of weights in the current quantile. The results are collected in the `p_quantiles` parameter. In order to prepare for the next quantile calculation, we reset the sum for the next quantile and update the threshold.

```

p_curSumW      = 0;
p_curSumWQ     = 0;
p_curSumQ(res) = 0;
loop(curHhid,
  p_curSumW      = p_curSumW      + p_sorted(curHhid, "weight");
  p_curSumWQ     = p_curSumWQ     + p_sorted(curHhid, "weight");
  p_curSumQ(res) = p_curSumQ(res) + p_sorted(curHhid, res)*p_sorted(curHhid, "weight");
  if ( p_curSumW ge p_threshHold,
    p_quantiles(microR, Perc, res) $ (perc.pos eq count) = p_curSumQ(res) / p_curSumWQ;
    p_threshHold = p_threshHold + sumOfWeights/100;
    count = count + 1;
    p_curSumQ(res) = 0;
    p_curSumWQ     = 0;
  );
);

```

One might now guess from the income shares etc. reported for the different quantiles where differences stem from. In order to improve here, we estimate an OLS model which use per capita

income and the different shares as explanatory variables. The regression coefficients with high p-values can help to interpret the results. To do so, we use the LS solver from GAMS (or embedded Python for GAMS version 34 or later, not shown here):

```

set expV(datVars);expV(addInc) = YES; expV("incPerCap") = YES;
variable v_ssq,v_cnst,v_par;
equation e_sumsq,e_fit;
*
* --- simple linear model which estimates the money metric
*      from the income shares and total income
*
e_sumsq.. v_ssq =n= 0;

e_fit(country_hhid(curMicroR,curHhid),ts)..

p_microRes(curMicroR,curhhid,"mm",ts) =e= v_cnst + sum(curExpV, v_par(curExpV) * p_compact(curMicroR,curHhid,curExpV));

```

Processing the results is somewhat cumbersome as the parameter and test statistics are not correctly labelled:

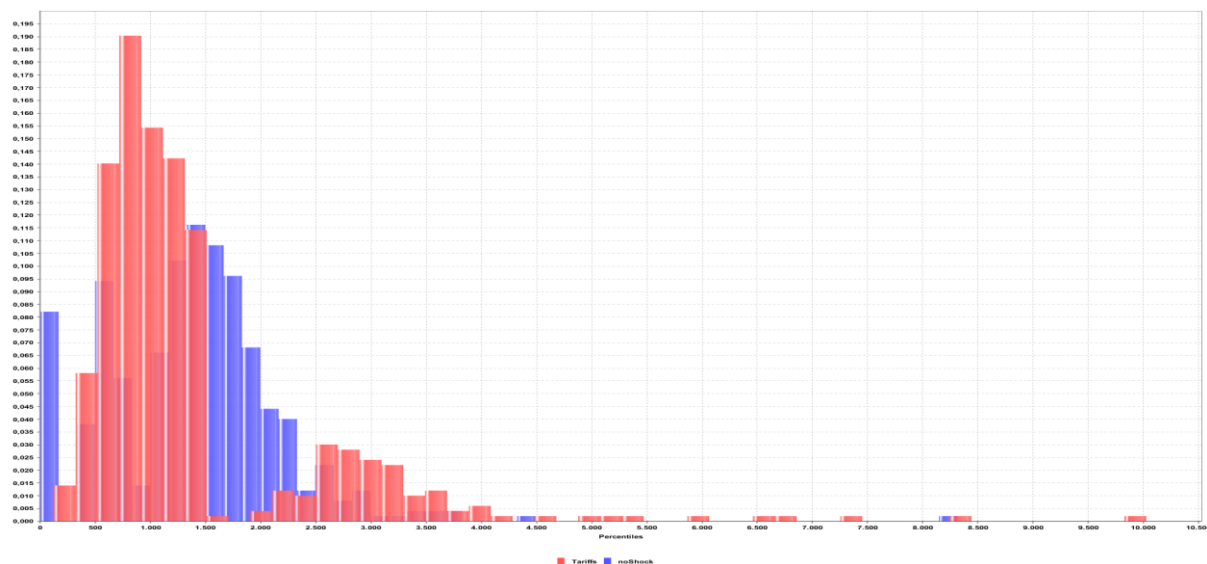
```

solve m_leastsq using lp minimizing v_ssq;
execute_load 'ls', se, sigma, r2, df, tval, pval;
if ( card(tVal) eq card(curExpV)+1,
    p_quantiles(microR,"par",curExpV) = v_par(curExpV);
    p_quantiles(microR,"p-Val",curExpV) = sum(xls $ (xls.pos+1 eq curExpV.pos),pval(xls));
    p_quantiles(microR,"t-Val",curExpV) = sum(xls $ (xls.pos+1 eq curExpV.pos),tval(xls));
    p_quantiles(microR,"se",curExpV) = se(curExpV);
    p_quantiles(microR,"R2","") = r2;
    p_quantiles(microR,"df","") = df;
);

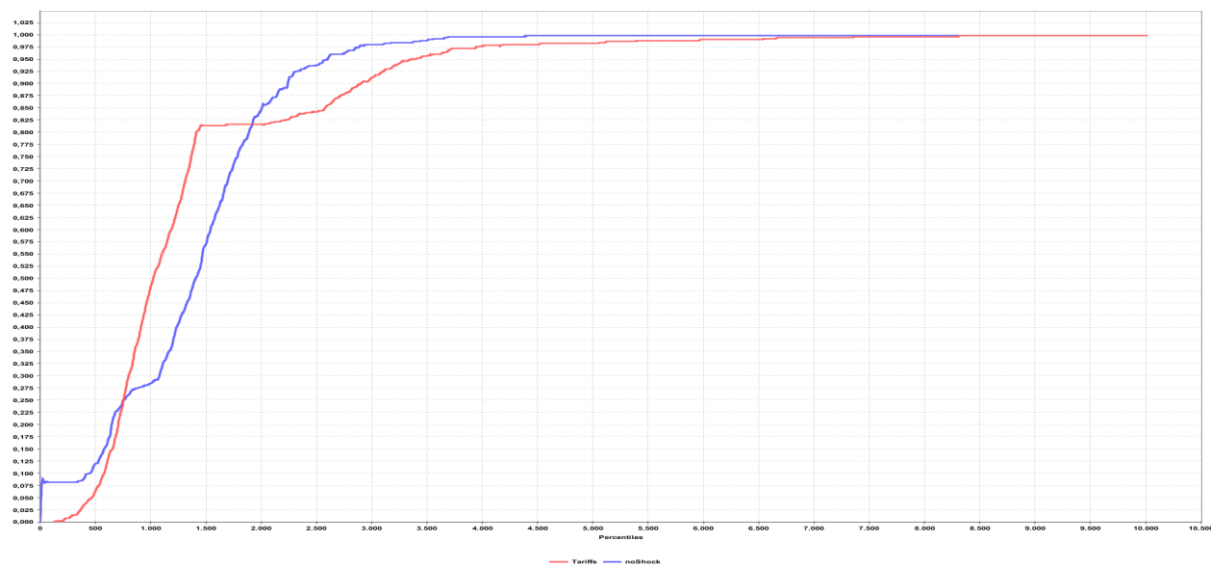
```

The results can be inspected with the interface, we offer:

1. Histograms which show the distribution of the various items (either for aggregated groups or rather details), for instance:



2. The same information, in cumulative format:



Both can also be shown as tables.

3. Finally, we show results for the OLS estimation:

| Microsimulation results, OLS [0]                   |  |       |      |          |
|----------------------------------------------------|--|-------|------|----------|
| Regions                                            |  | Years |      |          |
| Vietnam                                            |  | shock |      |          |
| Tariffs                                            |  | nc    |      |          |
| par                                                |  | R2    | df   | pa       |
| Money metric [cnst \$ per capita]                  |  |       | 0.98 | 43364.00 |
| Crop production [cnst \$ per capita]               |  | 0.12  |      |          |
| Livestock production [cnst \$ per capita]          |  | 0.11  |      |          |
| Public transfers [cnst \$ per capita]              |  | -0.03 |      |          |
| Private transfers [cnst \$ per capita]             |  | 0.10  |      |          |
| Wages, agriculture, unskilled [cnst \$ per capita] |  | 0.14  |      |          |
| Wages, agriculture, skilled [cnst \$ per capita]   |  | 0.15  |      |          |
| Wages, agriculture, other [cnst \$ per capita]     |  | 0.11  |      |          |
| Wages, extraction, unskilled [cnst \$ per capita]  |  | 0.11  |      |          |



### 3.3 A GMIG module for CGEBox

Wolfgang Britz, November 2019

#### 3.3.1 Background and motivation

The GMIG data base (cf. Walsmley and et al. 2011<sup>24</sup>) reports international bi-lateral labour and population migration along with wage differences. Specifically, it extends the GTAP standard data set with data on persons employed of a certain nationality working in other countries and the wages they receive at market, buyers' and sellers' prices. It also adds data on bi-lateral remittances. One of the interesting features of the GMIG data base is that reports persons such that wages per capita can be defined. The data base was developed to source the GMIG model which simulates changes in these variables. The GMIG model otherwise is a rather straightforward extension of the GTAP standard model. Extending CGEBox with that module, especially for long-run analysis, seems therefore inviting.

GMIG considers that the labour stock and population of a country consist of nationals and migrants. The GMIG module in CGEBox assumes that an aggregate agent in each home country manages the labour stocks by skill category of nationals and distributes these stocks to the home country and abroad according to a Constant-Elasticity-of-Transformation function. The function maximizes revenues from wages after direct taxes, corrected for the consumer price index. Migrant labour leaves its home country along with some dependents and updates the population at home and abroad. Migrants send a share of the labour income as remittances back to the country they originate. Tax rates paid by firms and direct taxes paid by employees differ by country of origin of the migrants (respectively the home country).

#### 3.3.2 Equations

The key equation in the GMIG model links the number of workers of skill category  $l$  from region  $rNat$  who work in another region  $rNat1$  to differences in wages after direct taxes  $kappaFmig$  based on a CET function, i.e. it assumes that labour migration reflects a revenue maximization problem. To do so, we define purchasing power corrected wages  $v\_wagePP$  of workers from  $rNat$  received in  $rNat1$ , using the GDP price index  $pgdpmp$  to reflect changes in purchasing power:

$$v\_wagePP_{rnat,rnat1,l} = pft_{rnat1,l}(1 - v\_kappaFmig_{rnat,rnat1,l})/pgdpmp_{rnat1} \quad (38)$$

```
e_wagePP(rNat,rNat1,l,ts(t)) $ p_xftMIG(rNat,rNat1,l,t) ..
wagePP(rNat,rNat1,l,t) =E= pft(rNat1,l,t)/pgdpmp(rNat1,t) * (1 - kappaFMig(rNat,rNat1,l,t));
```

For all equations, the diagonal elements  $rNat,rNat$  are defined as well, they denote the workers who are employed in their home country. The average wage  $v\_wagePPA$  received for workers from  $rNat$  is defined from the dual price index at a transformation elasticity of  $p\_omegaMig$  and share parameters  $p\_gMIG$ :

$$v\_wagePPA_{rnat,l} = \left[ \sum_{rnat} p\_gMig_{rnat,rnat1,l} v\_wagePP_{rnat,rnat1,l}^{(1+p\_omegaMig)} \right]^{1/(1+p\_omegaMig)} \quad (39)$$

<sup>24</sup> Walmsley, Terrie L., Alan Winters, and Amer Ahmed. "The impact of the movement of labour: results from a model of bilateral migration flows." *Global Economy Journal* 11.4 (2011): 1850240



```
e_wagePPA(rNat,l,ts(t)) ..
```

```
wagePPA(rNat,l,t) =E= sum(rNat1 $ p_xftMIG(rNat,rNat1,l,t),
                           p_gMig(rNat,rNat1,l,t) * wagePP(rNat,rNat1,l,t)**(1+p_omegaMig(rNat,l))) ** (1/(1+p_omegaMig(rNat,l))));
```

The CET distribution determines the number of workers  $xfMig$ , measured in constant dollars at market prices, from  $rNat$  working in  $rNat1$  by distributing the given labour stock  $p\_xftMig$  of the originator country  $rNat$ . The total stock is updated by two terms: the real wage average, defined by the wage  $pft$  deflated by the price index of GDP  $pgdmp$ , exponent the labor supply elasticity of the total labor stock  $p\_etaf$  and the change in unemployment rate.

$$xfMig_{rnat,rnat1,l} = p\_XftMig_{rnat,rnat1,l} p\_gMig_{rnat,rnat1,l} \left( \frac{V_{wagePP}^{rnat,rnat1,l}}{V_{wagePPA}^{rnat,l}} \right)^{p\_omegaMig} \quad (40)$$

$$\left( 1 + \left( \frac{pft_{rnat1,l}}{pgdmp_{rnat1}} \right)^{p\_etaf_{rnat1}-1} + (unEmpl_{rnat1,l} - p\_unEmpl_{rnat1,l}) \right)$$

```
e_labDist(rNat,rNat1,l,ts(t)) $ p_xftMIG(rNat,rNat1,l,t) ..
xfMig(rNat,rNat1,l,t)s(xfMig,"rNat,rNat1,l,t")
=E= p_xftMIG(rNat,"tot",l,t)s(xfMig,"rNat,rNat1,l,t")
    * p_gMig(rNat,rNat1,l,t) * (wagePP(rNat,rNat1,l,t)/wagePPA(rNat,l,t))**p_omegaMig(rNat,l)

--- etaff effect (labor supply curve for labor participation rate)
* (1 + (pft(rNat1,l,t)/pgdmp(rNat1,t))**p_etaf(rNat1,l)-1

--- impact of real wages on unemployment
+ (unEmpl(rNat,l,t)-p_unEmpl(rNat,l,t))
  $ (p_unEmpl(rNat1,l,t) and p_unEmplMin(rNat1,l,t) and p_etafUnEmpl(rNat1,l))
);
```

The unemployment rate  $unEmpl$  is driven by the real wage exponent an elasticity  $p\_etafUnEmpl$ , but lower bounded by a minimal rate of unemployment  $p\_unEmplMin$ :

$$unEmpl_{rnat1,l} \quad (41)$$

$$= p_{unEmpl_{rnat1,l}} + \min \left( unEmpl_{rnat1,l} - p_{unEmplMin_{rnat1,l}} p_{unEmpl_{rnat1,l}} \left( \frac{pft_{rnat1,l}}{pgdmp_{rnat1}} \right)^{p_{etafUnEmpl_{rnat1}}-1} - 1 \right)$$

```
e_unEmpl(rNat,l,ts(t)) $ (p_unEmpl(rNat,l,t) and p_unEmplMin(rNat,l,t) and p_etafUnEmpl(rNat,l)) ..
unEmpl(rNat,l,t)
=E= m_smoothMin(p_unEmpl(rNat,l,t)-p_unEmplMin(rNat,l,t),p_unEmpl(rNat,l,t)*((pft(rNat,l,t)/pgdmp(rNat,t))**p_etafUnEmpl(rNat,l)-1)) + p_unEmpl(rNat,l,t);
```

Adding up over the workers from different regions  $rnat1$  currently working in  $rnat$  defines the total labour stock  $xft$  of skill categories in  $rnat$  and drives the prices for labour  $pft$ :

$$XFT_{rnat,l} = \sum_{rnat1} xfMig_{rnat1,rnat,l} \quad (42)$$

```
*
* --- labour use stock equation: labor use in rNat is sourced by workers from different countries
*
e_xftMig(rsNat(rNat),l,ts(t)) $ (rsNat(rNat) $ (xft.range(rNat,l,t) ne 0)) ..
xft(rNat,l,t)s(xft,"rNat,l,t") =E= sum(rNat1 $ p_xftMIG(rNat1,rNat,l,t), xfMig(rNat1,rNat,l,t))s(xft,"rNat,l,t");
*
```

As the data base reports bi-lateral tax rates from a buyers'  $fcttx$  and sellers'  $kappaf$  perspective, we need to derive average rates applied in a country to the labour stock by skill category  $l$ :

$$XFT_{rnat,l} pft_{rnat,l} kappa_{rnat,l} = \sum_{rnat1} v\_xfMig_{rnat,rnat1,l} pft_{rnat,l} kappa_{rnat,rnat1,l} \quad (43)$$

```
e_kappaMig(rNat,l,ts(t)) $ rsNat(rNat) ..
kappa(rNat,l,t) * xft(rNat,l,t)s(xft,"rNat,l,t")
=E= sum(rNat1 $ p_xftMIG(rNat1,rNat,l,t), xfMig(rNat1,rNat,l,t) * kappaMig(rNat1,rNat,l,t))s(xft,"rNat,l,t") ;
```

In case of ad-valorem factor taxes rates  $fcttx$ , the update process is somewhat more complex as factor taxes are defined specific per activity in the GTAP standard model whereas the GMIG data base seems to report only average differences across activities:

$$XFT_{rnat,l,a} pft_{rnat,l} fcttx_{rnat,l,a} = \sum_{rnat1} v\_xfMig_{rnat,rnat1,l} pft_{rnat,l} (fcttx^0_{rnat,l,a} + v\_fcttxMig_{rnat,rnat1,l} - v\_fcttxMig^0_{rnat,rnat1,l}) \quad (44)$$

```
e_fcttxMig(xfFlag(rNat,l,a),ts(t)) $ rsNat(rNat) ..
fcttx(rNat,l,a,t) * xft(rNat,l,t)s(xft,"rNat,l,t")
=E= sum(rNat1 $ p_xftMIG(rNat1,rNat,l,t), xfMig(rNat1,rNat,l,t)
* (fcttx.l(rNat,l,a,"%t0%")+fcttxMig(rNat,rNat1,l,t)-fcttxMig.l(rNat,rNat1,l,"%t0%"))s(xft,"rNat,l,t");
```

In both cases, the model equations drop the price weights  $pft$  which is found on both sides. They are shown in the equations above to underline that we use a value weighted average which fits the ad-valorem tax rate definitions used in the model.

Finally, we assume that the average relation between migrated people  $popMig$  and workers stays identical, using the head count  $p\_labMig$  by skill category reported in the data:

$$\frac{popMig_{rnat,rnat1}}{popMig^0_{rnat,rnat1}} = \frac{\sum_l p\_labMig_{rnat,rnat1,l} \frac{xfMig_{rnat,rnat1,l}}{xfMig^0_{rnat,rnat1,l}}}{\sum_l p\_labMig_{rnat,rnat1,l}} \quad (45)$$

The model equation considers that changes in the labor participation rate do not impact the population size:

```
.....
*
* --- update migrated population according to change in migrated labour stock
*
e_popMig(rNat,rsNat(rNat1),ts(t)) $ p_popMIG(rNat,rNat1) ..
popMig(rNat,rNat1,t)s(popMig,"rNat,rNat1,t")
=E= sum(l $ p_xftMIG0(rNat,rNat1,l),
*
* --- change in labour stock [measured in cnst dollar at market prices], per skill category
*
* xfMig(rNat,rNat1,l,t)/p_xftMig0(rNat,rNat1,l)
*
* --- take off etaff effect (a higher labor participation rate does not)
*
* /(1 + (pft(rNat1,l,t)/pgdpmp(rNat1,t))**p_etaf(rNat1,l)-1
*
* --- impact of real wages on unemployment
*
* + (unEmpl(rNat,l,t)-p_unEmpl(rNat,l,t))
* $ (p_unEmpl(rNat1,l,t) and p_unEmplMin(rNat1,l,t) and p_etafUnEmpl(rNat1,l))
*
* --- multiplied with # of workers = change in labour force in #
*
* p_labMig(rNat,rNat1,l)
* /
* ---- divided by # of workers at benchmark, multiplied with population in #
* sum(l, p_labMig(rNat,rNat1,l)) * p_popMig(rNat,rNat1)s(popMig,"rNat,rNat1,t");
```

From there, the current population in a country  $\text{pop}_{\text{rnat,hhsld}}$  is defined, taking a shifter in population size  $\text{popGmigShift}_{\text{rnat}}$  into account. That shifter is used in baseline construction to target a given population size.

$$\text{pop}_{\text{rnat,hhsld}} = \text{popGmigShift}_{\text{rnat}} \sum_{\text{rnat1}} \text{popMig}_{\text{rnat1,rnat}} \quad (46)$$

Changes in population of single households  $h$  follow the change in the total production:

```
e_popTot(rs(r),h,ts(t)) $ (hFlag(r,h) and (pop.range(r,h,t) ne 0))..
    pop(r,h,t)/p_pop(r,h,t) =E= pop(rNat,"hhsld",t)/p_pop(rNat,"hhsld",t);
```

The remittances  $v_{\text{remit}}$  are defined relative to base year level:

$$v_{\text{remit}}_{\text{rnat,rnat1}} = \sum_l \frac{\text{xfMig}_{\text{rnat,rnat1,l}}}{\text{xfMig}^0_{\text{rnat,rnat1,l}}} p_{\text{remit}}_{\text{rnat,rnat1,l}} \quad (47)$$

```
e_remit(rNat,rNat1,ts(t)) $ p_popMIG(rNat,rNat1) ..
    remit(rNat,rNat1,t)s(remit,"rNat,rNat1,t")
    =E= sum(l $ p_xftMig(rNat,rNat1,l,t),xfMig(rNat,rNat1,l,t)*pft(rNat1,l,t)/lcu(rNat1,t)
        *p_remit(rNat,rNat1,l)/p_xftMig0(rNat,rNat1,l))s(remit,"rNat,rNat1,t") ;
```

As the remittances are part of the original B.O.T., we would need some re-bookings to correct regional income and foreign savings to introduce them already at the benchmark where they must be part of the reported data set. We therefore introduce an absolute change in the remittances balance *finc* (= foreign income), a position which is already part of the myGTAP module:

$$\begin{aligned} & V_{\text{fincRnat}} \\ &= \sum_{\text{rnat1}} v_{\text{remit}}_{\text{rnat,rnat1}} - v_{\text{remit}}_{\text{rnat1,rnat}} - v_{\text{remit}}_{\text{rnat,rnat1}}^0 + v_{\text{remit}}_{\text{rnat1,rnat}}^0 \end{aligned} \quad (48)$$

```
e_fincGMIG(rsNat(rNat),ts(t)) $ (finc.range(rNat,t) ne 0) ..
    finc(rNat,t)s(finc,"rNat,t")
    =E= [ sum(rNat1, remit(rNat,rNat1,t) $ p_popMIG(rNat,rNat1) - remit(rNat1,rNat,t) $ p_popMIG(rNat1,rNat))
        -sum(rNat1, remit.l(rNat,rNat1,"%t0%") $ p_popMIG(rNat,rNat1) - remit.l(rNat1,rNat,"%t0%") $ p_popMIG(rNat1,rNat))
        ]s(finc,"rNat,t");
```

The change enters the B.O.T. = 0 identity and the regional income definitions.

### 3.3.3 Technical aspects

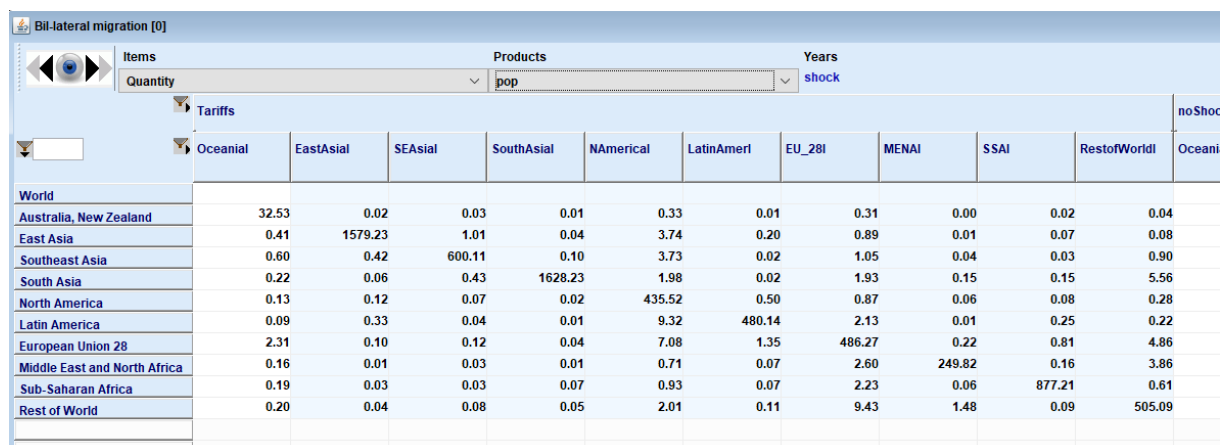
There are number of points to note with regard to the interaction with other modules in CGEBox

1. The GTAP standard model core has a supply equation for total factor supply  $e_{\text{xft}}$ . This equation is switched off for labour categories if the GMIG module is active as (42) above defines the labour stock.
2. The myGTAP module has a simplified equation to determine remittances from total labour use in the countries which is replaced by equation (47) above.
3. Some standard closures are changed. So are population and direct taxes on labour fixed by default in the GTAP standard model and become endogenous if the GMIG module is switched on.

The GMIG module is fully compatible with the other modules of CGEBox and had been tested in various combinations.

### 3.3.4 Result exploitation

The post-model processing adds a table which shows the bi-lateral population and labour force distribution and report the wages receives, at market prices and after direct taxation.



|                              | Oceania | EastAsia | SEAsia | SouthAsia | NAmerica | LatinAmerl | EU_28l | MENAI  | SSAI   | RestofWorldl | Oceani |
|------------------------------|---------|----------|--------|-----------|----------|------------|--------|--------|--------|--------------|--------|
| World                        |         |          |        |           |          |            |        |        |        |              |        |
| Australia, New Zealand       | 32.53   | 0.02     | 0.03   | 0.01      | 0.33     | 0.01       | 0.31   | 0.00   | 0.02   | 0.04         |        |
| East Asia                    | 0.41    | 1579.23  | 1.01   | 0.04      | 3.74     | 0.20       | 0.89   | 0.01   | 0.07   | 0.08         |        |
| Southeast Asia               | 0.60    | 0.42     | 600.11 | 0.10      | 3.73     | 0.02       | 1.05   | 0.04   | 0.03   | 0.90         |        |
| South Asia                   | 0.22    | 0.06     | 0.43   | 1628.23   | 1.98     | 0.02       | 1.93   | 0.15   | 0.15   | 5.56         |        |
| North America                | 0.13    | 0.12     | 0.07   | 0.02      | 435.52   | 0.50       | 0.87   | 0.06   | 0.08   | 0.28         |        |
| Latin America                | 0.09    | 0.33     | 0.04   | 0.01      | 9.32     | 480.14     | 2.13   | 0.01   | 0.25   | 0.22         |        |
| European Union 28            | 2.31    | 0.10     | 0.12   | 0.04      | 7.08     | 1.35       | 486.27 | 0.22   | 0.81   | 4.86         |        |
| Middle East and North Africa | 0.16    | 0.01     | 0.03   | 0.01      | 0.71     | 0.07       | 2.60   | 249.82 | 0.16   | 3.86         |        |
| Sub-Saharan Africa           | 0.19    | 0.03     | 0.03   | 0.07      | 0.93     | 0.07       | 2.23   | 0.06   | 877.21 | 0.61         |        |
| Rest of World                | 0.20    | 0.04     | 0.08   | 0.05      | 2.01     | 0.11       | 9.43   | 1.48   | 0.09   | 505.09       |        |

## 3.4 Sub-national dis-aggregation of production and factor markets in CGEBox

### 3.4.1 Background and motivation

Especially questions around the use of immobile natural resources such as land, minerals or water often ask for an analysis at sub-national level, particularly if environmental consequences are to be evaluated and externalities have a regional character. It is therefore not uncommon to find sub-national detail in CGE analysis at single country level. We present in here a rather basic approach to add such detail to the GTAP model. For Europe, ready-to-use data at the level of NUTS2 regions are automatically added during the data preparation step if single European countries are present in the data set. For other countries or model regions, the user must supply these data to employ the module. The interface allows selecting during simulation for which countries the module is active.

### 3.4.2 Methodology

We implement a rather basic dis-aggregation as a starting point for further work:

1. No intra-national trade margins are introduced. There is one uniform price for the domestic origin faced by intermediate demand in all sub-regions of the model region.
2. Commodity demand is accordingly modeled at national level.
3. Factor markets however are depicted by a CET structure which hence allows for sluggish or even no factor mobility across regions.
4. The tax system is not regionally differentiated.
5. All production functions are depicted at sub-national level.

As a consequence, the data requirements are relatively low. Basically, data on output quantities and cost shares net of taxes are required for the framework.

The *myGTAP* module can change the assumption above and introduce sub-national final demand with preferences for products produced in the same sub-national unit.

### 3.4.3 Integration into the modeling framework

Limited changes are necessary to introduce regional detail for the supply side into the model and there were mostly realized via macros. They use the cross-set  $r\_r(r, rp)$  that depicts the relation between sub-region  $r$  and nation  $rp$ . In case of model regions without sub-national detail, the cross-set depicts a diagonal relation, i.e.  $r\_r(rNat, rNat) = \text{yes}$ .

In order to avoid problems with a simply linear aggregation from regions to nations, a CES-aggregator is used in the code. This assumes quality differences for that same product  $i$  outputted in different sub-regional units  $subr$  such that production costs need not be identical. The equation  $xsNat$  defines the average price at national level, the second equation defines the demand from sub-region  $subr$  depending on its production costs relative to the national average.

```

--- Equations active for sub-regional units, not part of GTAP V7 model

(a) CES price aggregator from sub-regional offer prices to national average
e xsNat(disr,iIn(i),ts) $ ( sum(a,xFlag(disr,a,i)) $ diag(i) $ (xs.range(disr,i,ts) ne 0)) ..
ps(disr,i,ts) =E= sum(r_r(subr,disr) $ (p_gfi(subr,i,disr,ts) gt eps),
  p_gfi(subr,i,disr,ts)*ps(subr,i,ts)**(1-p_sigmaSubr(disr,i))**(1/(1-p_sigmaSubr(disr,i))));

(b) distribution of supply
e psSubr(subr,iIn(i),ts) $ ( sum(a,xFlag(subr,a,i)) $ (xs.range(subr,i,ts) ne 0) $ diag(i)) ..

(xs(subr,i,ts)
  $iftheni.hsubr "%hSubr%"=="on"
  - sum(h $ (p_alphad(subr,i,h,ts) gt eps), xd(subr,i,h,ts))
  $endif.hsubr
  - sum(aaDm,p_subrDem(subr,i,aaDm,ts))
)s(xs,'subr,i,ts')
=E= sum(r_r(subr,disr) $ p_gfi(subr,i,disr,ts),
  xs(disr,i,ts) * p_gfi(subr,i,disr,ts) * (ps(disr,i,ts)/ps(subr,i,ts))**p_sigmaSubr(disr,i))s(xs,'subr,i,ts');

```

In case of a non-diagonal make, the following equations are used:

```

(c) same as (a), above in case of non-diagonal activity: average costs of the activity
e xpNat(disr,aIn(a),ts) $ ( xpFlag(disr,a) $ (not diag(a)) ) ..

px(disr,a,ts) =E= sum(r_r(subr,disr) $ p_gfa(subr,a,disr,ts),
  p_gfa(subr,a,disr,ts)*px(subr,a,ts)**(1-p_sigmaASubr(disr,a))**(1/(1-p_sigmaASubr(disr,a))));

(d) same as (a), above in case of non-diagonal activity: distribution of activity level
e pxSubr(subr,aIn(a),ts) $ ( xpFlag(subr,a) $ (not diag(a)) ) ..

xp(subr,a,ts)s(xp,'subr,a,ts')
=E= sum(r_r(subr,disr),
  xp(disr,a,ts) * p_gfa(subr,a,disr,ts) * (px(disr,a,ts)/px(subr,a,ts))**p_sigmaASubr(disr,a))s(xp,'subr,a,ts') ;

```

Where  $p\_sigmaASubr$  is set to 5. It defines the substitution elasticities when aggregating production output from sub-region to nation. This implies that the national composition of output from the regions reacts rather flexibly.

There are various equations in the model that ensure that firm demand is aggregated over regions to country level, while it is assumed that all regions face the same prices for intermediate inputs. Factor prices clearly differ reflecting the assumption with regard to factor mobility across regions and sectors.

In order to ease reading the code and speeding up execution, the set  $disr$  depicts those nations which feature regional detail while the set  $subr$  depicts sub-regions. A further equation adds up intermediate in each sector over regions to national level:

```

--- aggregate firm demand from sub-region to regional level, using the Armington prices define above
e_xapNat(rsNat,iIn(i),aIn(a),ts) $ (xaiFlag(rsNat,i,a) $ disr(rsNat)) ..

xa(rsNat,i,a,ts)s(xa,"rsNat,i,a,ts")
=e= [
    sum(tNest_i_a(iotNest,i,a),
        sum(r_r(subr,rsNat) $ p_ioNestT(subr,iotNest,a,i),m_xaIoNest(subr,i,a,ts))
    ) $ iInTnest(i)
]s(xa,"rsNat,i,a,ts");

--- define Armington prices for intermediate demand at national level for dis-aggregated
model regions (normally, the price is substituted out

e_paNat(disr(rsNat),iIn(i),aIn(a),ts) $ (xaiFlag(disr,i,a)
    $ $ifi not "%modulesAggregate_firm_demand%"=="ON" and (not iMel(i))
) ..
pa(rsNat,i,a,ts) =E= m_paintr(rsNat,i,a,ts);

```

An additional equation was added which distributes the total national factor supply  $xft(disr)$  to the regions:

```

--- Factor market clearing from sub-regions to nation, not part of standard GTAP
e_xftRes(subr,fm,ts) $ (xftEqFlag(subr,fm) $ (pft.range(subr,fm,ts) ne 0)) ..
0 =E=
    --- "sluggish" factor supply based on CET function, factor is part of top level nest
    (xft(subr,fm,ts)s(xft,'subr,fm,ts')
        - sum(r_r(subr,disr),p_gfr(subr,fm,disr,ts)*xft(disr,fm,ts)s(xft,'subr,fm,ts')
            * (pft(subr,fm,ts)/pft(disr,fm,ts))**p_omegafr(disr,fm))
        ) $ (sum(r_r(subr,disr),p_omegafr(disr,fm)) ne inf)
    --- fully mobile factor: prices are equal across regions
    + (pft(subr,fm,ts)s(pft,'subr,fm,ts') - sum(r_r(subr,disr),pft(disr,fm,ts)s(pft,'subr,fm,ts')) $ (sum(r_r(subr,disr),p_omegafr(disr,fm)) eq inf);

```

If the transformation elasticity across regions  $p\_omegafr$  is set to zero for factors, the regional factor stock  $xft$  is fixed. The last line in the equation depicts the case of infinite transformation which leads to uniform prices. A matching dual price aggregator defines the price for mobile factors at national level:

```

--- factor supply from nation to sub-regions, not part of standard GTAP
e_pftNat(disr,fm,ts) $ (xftEqFlag(disr,fm) and (pft.range(disr,fm,ts) ne 0)) ..
0 =E=
    --- sluggish factor mobility across regions, dual price aggregator
    [ pft(disr,fm,ts)s(pft,'disr,fm,ts') - sum(r_r(subr,disr) $ xftEqFlag(subr,fm),
        p_gfr(subr,fm,disr,ts)*pft(subr,fm,ts)**(1+p_omegafr(disr,fm)))*(1/(1+p_omegafr(disr,fm)))s(pft,'disr,fm,ts')
    ] $ (p_omegafr(disr,fm) ne inf)
    --- fully mobile factors across regions, uniform prices => physiscal aggregation
    + [
        xft(disr,fm,ts)s(xft,'disr,fm,ts') - sum(r_r(subr,disr), xft(subr,fm,ts)s(xft,'disr,fm,ts'))
    ] $ (p_omegafr(disr,fm) eq inf);

```

Not that the last line provides the physical linear aggregation in case that the transformation elasticity is infinite.

In order to avoid that the tax income equations are changed when sub-national regions are introduced, factor use for each activity and matching prices are defined in two additional equations at national level:

```

--- Average factor returns at national level, aggregated from sub-regions, exhaustion
(Not part of standard GTAP)

= pfNat(disr,f,aIn(a),ts) $ (xfFlag(disr,f,a) $ (pf.range(disr,f,a,ts) ne 0)) ..

pf(disr,f,a,ts)*xf(disr,f,a,ts)s(xf,'disr,f,a,ts')
=E= sum(r_r(subr,disr) $ xfFlag(subr,f,a),
    pf(subr,f,a,ts)*xf(subr,f,a,ts))s(xf,'disr,f,a,ts');

--- Total national factor use, added up over sub-regions
(Not part of standard GTAP)

= xfNat(disr,f,aIn(a),ts) $ xfFlag(disr,f,a) ..

xf(disr,f,a,ts)s(xf,'disr,f,a,ts')
=E= sum(r_r(subr,disr) $ xfFlag(subr,f,a), xf(subr,f,a,ts))s(xf,'disr,f,a,ts');

```

The code comprises at various places if \$ *rNat* conditions to avoid that data transformations for sub-regional data are executed which do not relate to production or factor markets. During variable initialization and calibration, it is assumed that the regional data on intermediate input and primary factor use match the national totals (see next section for an implementation). Factor, production and immediate taxes rates are taken from the national SAMs.

```

fctts.l(r,f,a,t0) $ fcttsV0(r,f,a) = fcttsV0(r,f,a)/sam0(r,f,a) ;
fcttx.l(r,f,a,t0) $ fcttxV0(r,f,a) = fcttxV0(r,f,a)/sam0(r,f,a) ;

fctts.l(subR,f,a,t0) = sum(r_r(subR,r),fctts.l(r,f,a,t0));
fcttx.l(subR,f,a,t0) = sum(r_r(subR,r),fcttx.l(r,f,a,t0));

sam0(subR,"fctts",a) = sum(f, sam0(subR,f,a)*fctts.l(subR,f,a,"%t0%"));
sam0(subR,"fcttx",a) = sum(f, sam0(subR,f,a)*fcttx.l(subR,f,a,"%t0%"));

```

The intermediate taxes are implicitly taken over by using the relation between the national Armington intermediate demands and the national SAM entries net of taxes:

```

*
* --- Armington aggregator quantity in benchmark
* (calculated from value exhaustion)
*
xa.l(r,i,aa,t0) $ pa.l(r,i,aa,t0)
= (pdp.l(r,i,aa,t0)*xd.l(r,i,aa,t0)
+ pmp.l(r,i,aa,t0)*xm.l(r,i,aa,t0))/pa.l(r,i,aa,t0) ;

xa.l(subR,i,a,t0) $ sam0(subR,i,a)
= sam0(subR,i,a) * sum(r_r(subR,r), xa.l(r,i,a,t0)/sam0(r,i,a));

```

There were otherwise very limited changes to the calibration code.

### 3.4.4 Implementation for European countries at NUTS2 level

The NUTS (Nomenclature des unités territoriales statistiques) system provides a classification of administrative units for the European Union and some further regions. After NUTS0 (= nation) and NUTS1 (=federal state or similar), NUTS2 provides already relatively small regional units for economy wide assessment and beyond. In the context of the EU funded research project CAPRI-RD ([http://www.ilr.uni-bonn.de/agpo/rsrch/capri-rd/capri\\_rd\\_e.htm](http://www.ilr.uni-bonn.de/agpo/rsrch/capri-rd/capri_rd_e.htm)), regional SAMs at NUTS2 level and matching national ones for the EU member and candidates countries were compiled (Ferrari et al. 2012<sup>25</sup>). These SAMs featured somewhat limited sector detail (Agriculture, Forestry, Other primary,

<sup>25</sup> Ferrari, E., Himics, M. and Mueller M. (2010): WP2.2 Databases – Regional Social Accounting Matrices Deliverable: D2.2.4, Procedure for the compilation of regional SAMs based on national



Food processing, Other Manufacturing, Energy, Construction, Trade and transport, Hotels and restaurants, Education, Other Services) which reflects both data availability and the aim to model rural development policies under the Common Agricultural Policy (CAP). A specific feature of the SAMs is a dis-aggregation of intermediate demand to regional, national and imported origin. A matching single country CGE was developed (Törmä et al. 2010<sup>26</sup>, 2010, Britz 2012<sup>27</sup>) with a rather detailed driver to map individual rural development measures from the CAP into shocks for the model and an interface to the regional supply models and the market model of the partial equilibrium CAPRI modeling system. So far, the application of that model is rather limited (Schröder et al. 2015<sup>28</sup>, Britz et al. 2015<sup>29</sup>). The first version of the NUTS2 module used the production data of these regionals SAMs which also breaks down demand to the regional level and provides also details on the national and regional government accounts.

As the data were never updated, later version processed Eurostat regional economic statistics to estimate production value for each sector and sub-nation. Cost shares at sub-national level are assumed to be identical to the ones found in the GTAP data base.

The introduction of the list of sub-regions into the static set  $r$  (= all regions) which is used as the domain for various variables, equations and parameters requires that the list is available when the data from GTAPAGG are read. Accordingly, the data set program is partly read rather early in the program sequence:

```
$ifi %modulesGTAP_NUTS2%==on          $batinclude 'gtapNUTS2\gtapNUTS2.gms' decl1
$batinclude 'util\title.gms' '''Load data base for %dataSet%, %shock_withoutPath%'''
$include "%datdir%\gtap8Set.gms"
$include "build\load_gtapAgg.gms"
```

To load the following list with the NUTS2 regions:

---

SAMs and available regional datasets: dataset and documentation; CAPRI-RD deliverable D2.2.4, <http://www.ilr.uni-bonn.de/agpo/rsrch/capri-rd/docs/d2.2.4.pdf>

<sup>26</sup> Törmä, H., Zawalinska, K., Blanco-Fonseco M., Ferrari, E., Jansspon T. (2010): WP3.2 Model development and adaptation – Regional CGEs Deliverable: D3.2.1 Regional CGE model layout with a focus on integration with the partial equilibrium models and modelling of RD measures; CAPRI-RD deliverable D3.2.1, <http://www.ilr.uni-bonn.de/agpo/rsrch/capri-rd/docs/d3.2.1.pdf>

<sup>27</sup> Britz W. (2012) WP3.2 Model development and adaptation – Regional CGEs Additional Deliverable: D3.2.4 RegCgeEU+ in GAMS, documentation including the Graphical User Interface, CAPRI-RD deliverable D3.2.4, <http://www.ilr.uni-bonn.de/agpo/rsrch/capri-rd/docs/d3.2.4.pdf>

<sup>28</sup> Schroeder, L. A., Gocht, A., Britz, W. (2015): The Impact of Pillar II Funding: Validation from a Modelling and Evaluation Perspective, Journal of Agricultural Economics 66(2): 415–441

<sup>29</sup> Britz, W., Dudu, H., Ferrari, E. (2015): Economy-wide Impacts of Food Waste Reduction: A General Equilibrium Approach, selected paper presented at the International Conference of Agricultural Economists (ICAE 2015), 8-14 August 2015, Milan (Italy).



```

$iftheni.decl %1==decl

set s_nuts2SAM "NUTS2 labels used in SAM"/
BE100000 "Region de Bruxelles-Capitale"
BE210000 "Antwerpen (Arrondissement)"
BE220000 "Limburg (B)"
BE230000 "Oost-Vlaanderen"
BE240000 "Vlaams Brabant"
BE250000 "West-Vlaanderen"
BE310000 "Brabant Wallon"
BE320000 "Hainaut"
BE330000 "Verviers"
BE340000 "Luxembourg (B)"
BE350000 "Namur"
BG310000 "Severozapaden"
BG320000 "Severen tsentralen"
BG330000 "Severniztochen"

```

The actual data processing consists of the following steps:

1. Defining the regional hierarchy. Currently, the program only introduces sub-regions if the related country is a separate region in the GTAPAGG data set:

```

set rSam(r0); rSam(r0) $ sum(r0_msSam(r0,s_msSam),1) = yes;

check(s_msSam) = sum((mapr(r0,r),r0_msSam(r0,s_msSam)),1)-1;
set s_msSam_r(s_msSam,r);
*
* --- take out regions where countries with NUTS2 and without NUTS2 information are mixed
*
s_msSam_r(s_msSam,r) $ sum((mapr(r0,r),r0_msSam(r0,s_msSam)),1) = YES;
*
* --- take out regions where countries with NUTS2 and without NUTS2 information are mixed
*
s_msSam_r(s_msSam,r) $ (s_msSam_r(s_msSam,r) $ sum(mapr(r0,r) $(not sum(r0_msSam(r0,s_msSam),1)),1)) = NO;

```

From there, the regional hierarchy is defined:

```

set subReg_r(r2,r); subReg_r(r2,r) $ sum( (subReg_nuts0(r2,s_msSam),s_msSam_r(s_msSam,r)),1) = YES;

r_r(r2,r) $ sum(subreg_r(r2,r),1) = yes;
r_r(r,r) $ sum(subreg_r(r2,r),1) = no;

```

And the list of regions which are dis-aggregated:

```

disR(rNat) $ sum(subreg_r(r2,rNat),1) = yes;

sigmav(subReg,a) = sum(subReg_r(subReg,rs), sigmav(rs,a));
sigmand(subReg,a) = sum(subReg_r(subReg,rs), sigmand(rs,a));

omegafr(rs,fn) = 2.0;
omegafr(rs,"capital") = inf;
omegafr(rs,"land") = 0.00;

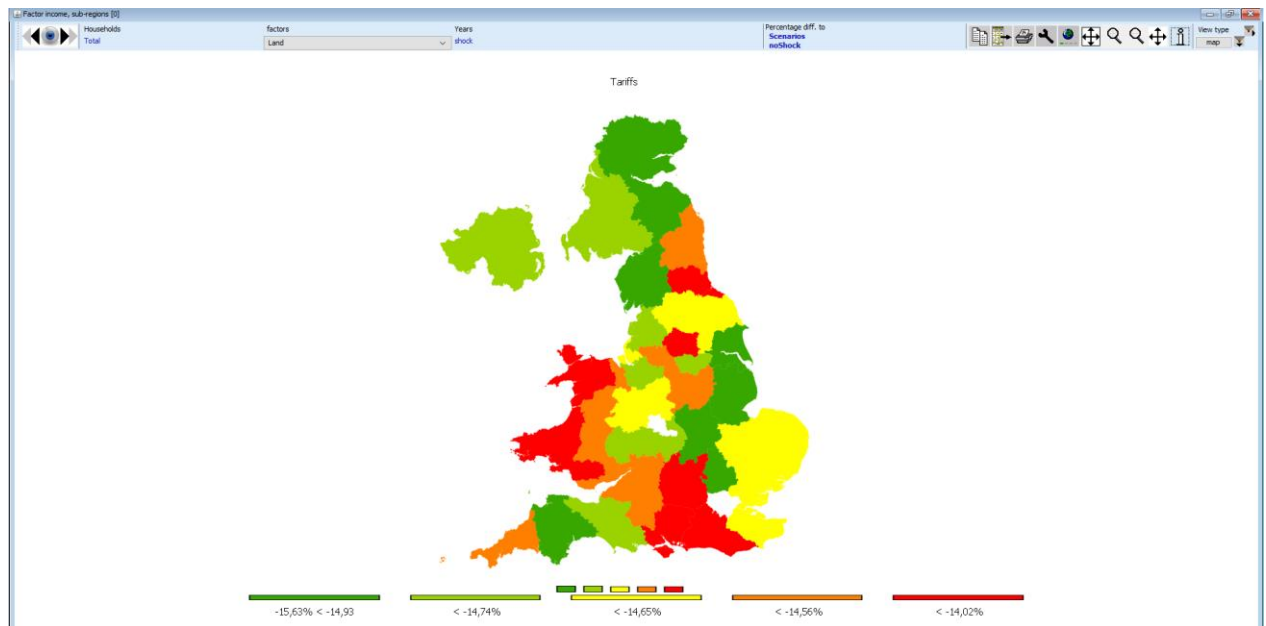
```

i.e.:

2. The substitution elasticities at national level are also used for the regions, and
3. Land cannot move between regions, capital is assumed as fully mobile and labor as sluggish. Note that natural resources are immobile at sector level.

### 3.4.5 Result exploitation

So far, there is only one specific view which uses a map with regional detail:



### 3.5 Armington with commitment terms

- Wolfgang Britz, 2021 –

#### 3.5.1 Background

This module allows introducing constant terms similar to a LES function in the bi-lateral import demand equations driven otherwise by a CES function. This allows dampening or increases the price responsiveness of selected bi-lateral trade flows, for instance, to counteract the “small shares stay small” problem of the standard CES approach.

As seen from the following graphic with equations in mathematical notation, the main differences to the usual CES aggregator are that (1) the Marshallians comprise constant terms  $c_i$ , (2) the exhaustion price is different from dual price CES index, (3) non-committed expenditure need to be calculated, which defines utility via exhaustion, and (4) the average import price used in the remainder of the model needs to be defined from expenditure exhaustion:

Utility  
(import demand  
quantity index)

Marshallians

$$x_i = u \delta_i \left( \frac{\tilde{p}}{p_i} \right)^\sigma + c_i$$

CES price index

$$\tilde{p} = \sum_i \delta_i p_i^{1-\sigma}$$

Non-committed expenditure

$$nce = e - \sum_i c_i p_i$$

u from exhaustion

$$u = nce / \tilde{p}$$

Definition of exhaustion price

$$e = u \hat{p}$$

A negative commitment  
increases the price elasticity

Note:  $e$  defined in other parts of model at given  $\hat{p}$

#### 3.5.2 Benchmarking and GUI setting

In opposite to the standard CES case, the LES-CES solution has an additional vector of product specific parameters. One set of import demand observations at a given substitution elasticity is hence not sufficient to identify both the share parameters and these additional constant terms. A second set of “artificial” observations is introduced that describes the desired changes in import shares relative to the standard model at a 10% price drop on the bi-lateral trade link.

The benchmarking process uses separate NLP models, one for each nation and product, that comprises the variable and equations from the LES-CES system, namely, for two set of observations of import prices and quantities called *calPoints*. This small model can also be used as a didactic one to experiment with expanded CES Armington model outside of the CGE model itself.

During calibration the share parameter  $v_s$  and the commitment terms  $v_c$  are to be defined such that the differences between the given second Marshallians observations  $p_x$  and the endogenous one  $v_x$  become minimal. The demands for the point observed in the data base are fixed.

The benchmarking model comprises the following equations:

- (1) Definition of the CES price index  $v_{pindCes}$  based on the usual dual CES price aggregator, at given prices  $p_p$ :

```
*
* --- CES price index
*
e_imppCES(calPoints) ..
    v_pindCes(calPoints) =E= sum(rnat1 $ v_s.l(rnat1), v_s(rnat1) * p_p(rnat1,calPoints) ** (1-p_p_sigmaW))**(1/(1-p_p_sigmaW));
```

- (2) Definition of non-committed expenditure  $v_{nonComE}$ . Note that the given budget for imports  $v_{bud}$  is fixed during benchmarking.

```
*
* --- Non-Committed expenditure
*
e_nonComE(calPoints) ..
    v_nonComE(calPoints) =E= v_bud(calPoints) - sum(rnat1 $ v_s.l(rnat1), v_c(rnat1) * p_p(rnat1,calPoints));
```

- (3) Indirect definition of utility  $v_u$  (which differs from total import quantity by the commitment terms) from non-committed expenditure and CES price index based on expenditure exhaustion:

```
*
* --- Definition of u
*
e_uCesLes(calPoints) ..
    v_u(calPoints) * v_pindCes(calPoints) =E= v_nonComE(calPoints);
```

- (4) Marshallian demands  $v_x$ , the difference to the standard CES being the constant terms  $v_c$ :

```
*
* --- Marshallian demands
*
e_marshall(rnat1,calPoints) $ v_s.l(rnat1) ..
    v_x(rnat1,calPoints) =E= (v_u(calPoints) * v_s(rnat1) * (v_pindCes(calPoints)/p_p(rnat1,calPoints))**(p_p_sigmaW + v_c(rnat1)));
```

- (5) Exhaustion price index  $v_{impp}$

```
*
* --- Definition of price index (u is implicate in simulation model where budget is exogenous)
*
e_impP(calPoints) ..
    v_bud(calPoints) =E= v_u(calPoints) * v_impp(calPoints);
```

- (6) Adding up of endogenous share parameters  $v_s$ :

```
*
* --- Share parameters add up to unity
*
e_addUpShare .. sum(rnat1 $ v_s.l(rnat1), v_s(rnat1)) =E= p_sumShare;
```

- (7) Inequality to make sure that Marshallian demands exceed commitment terms (to avoid that  $v_x$  is driven into the negative domain during simulation):

```
*
* --- Bounds on commitment terms to prevent problems in later sims
*
e_xc(rnat1,calPoints) $ v_s.l(rnat1) .. v_x(rnat1,calPoints)*0.50 =G= v_c(rnat1);
```

- (8) The objection function is defined as follows:

```

*
* --- fit for expert point (benchmark is fixed)
*
e_objFit .. v_opt * card(rnat1)
           =E= sum(rnat1, sqr[ (v_x(rnat1,"expert") - p_x(rnat1,"expert"))/(p_x(rnat1,"expert")+1.E-8) ]);

```

An additional equation is needed if the MRIO extension is switched on, where tariffs and import shares are given for each MRIO agent:

```

$iftheni.MRIO "%modulesGTAP_Mrio%"=="on"
*
* --- depending on the differences in c.i.f. prices across importers, the MRIO extension can imply quite different
* shares of the value of the commitment terms on total import value by agent. This equation ensures that the
* implicate expenditure can not exceed spendings by more than 25%
*
e_nonComEN(calPoints,mrioA) $ (card(curMioI) $ sum( (rsNat,curMioI,rnat1), mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"cif"))) ..
    sum(rnat1 $ (v_s.l(rnat1) $ sum((rsNat,curMioI), mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"cif"))),
        v_c(rnat1) * p_p(rnat1,calPoints)
            * sum( (rsNat,curMioI), mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"cif")
                * (1+mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"m")/mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"cif"))))
    =G= -sum(rnat1, v_x(rnat1,"%t0%") * p_p(rnat1,calPoints)
        * sum((rsNat,curMioI) $ mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"cif"),
            mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"cif")
                * (1+mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"m")/mrioSplitFactors(rsNat,curMioI,rnat1,mrioA,"cif"))))
    ) * 0.25;
$endif.MRIO

```

For the benchmarking, the user sets via the GUI (1) a multiplicative factor which defines the change in import demand relative to the standard Armington for an exogenous drop of the c.i.f. price by 10% (which can be also understood as a drop in the ad-valorem tariff), (2) define a threshold for the flows to boost.

Accordingly, the benchmarking code first calculates the Marshallians under the standard Armington response without commitment terms:

```

*
* --- simulate standard response (i.e. without commitment terms) at given total demand and updated prices
*
v_x(rnat1,calPoints) $ v_s(rnat1) = v_u(calPoints) * v_s(rnat1) * (v_pindCes(calPoints)/p_p(rnat1,calPoints))**p_p_sigmaW;

```

and defines from there “expert” points for the two cases, i.e. expected unchanged response and boosted response:

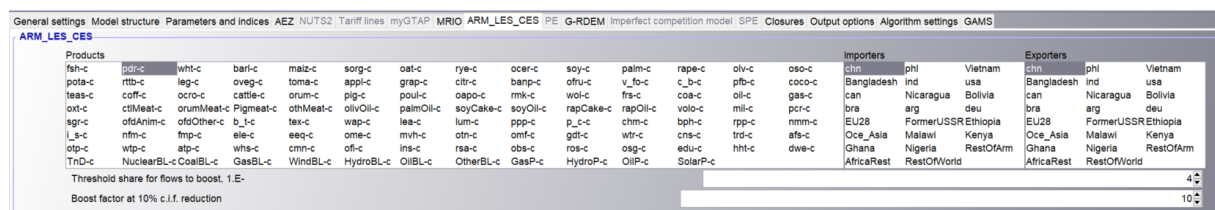
```

*
* --- for shares above the critical limit, we calibrated against the standard response
*
p_x(rnat1,"Expert") $ ((v_s(rnat1) gt 1.E-%arm_les_ces_boostShare%*p_sumShare) $ v_s(rnat1)) = v_x.l(rnat1,"expert");
*
* --- otherwise, we simulate the response at a share which is multiplied with the user chosen boostFactor,
* but does not exceed the critical share limit
*
p_x(rnat1,"Expert") $ ( (v_s(rnat1) le 1.E-%arm_les_ces_boostShare%*p_sumShare) $ v_s(rnat1) $ arm_les_ces_exp_sel(rnat1))
    = min(1.E-%arm_les_ces_boostShare%*v_u("expert") * (v_pindCes("expert")/p_p(rnat1,"expert"))**p_p_sigmaW,
        v_x(rnat1,"expert") * %arm_les_ces_boostFactor%);

```

The remainder code can be found in “[armLesCes\armLesCES\\_cal.gms](#)”.

The relevant settings on the interface are shown below:



i.e. the user can determine for which products, importers and exports the “boost” mechanism shall be activated, and defines the maximum share for flows to boost along with the boost factor.

The reader should note that it is typically impossible to introduce higher boost factor for larger trade shares, simply as boosting one flow implies a dampened response of others.

### 3.5.3 Equations

The  $e_{xw}$  equation for bi-lateral import demand in the GTAP Standard model is replaced by the following one called  $e_{xwArmLesCes}$  which considers simultaneously the standard case without and case with commitment terms. The cases where the commitment terms are considered as defined by a cross of importing countries and products  $arm\_LES\_CES(rnat,i)$ :

```
e_xwArmCom(rNat,i,rNat1,ts) $ (m_CombArmCom(rNat,i,rNat1) $ armComProCur(i)) ..
  xw(rNat,i,rNat1,ts) s(xw,"rNat,i,rNat1,ts")
  =e=
    [ ( p_amw(rNat,i,rNat1,ts)
      * ( xmt(rNat1,i,ts) $ (not arm_LES_CES(rNat1,i))
        + xmtUArmCom(rNat1,i,ts) $ arm_LES_CES(rNat1,i))
      * [(( pmt(rNat1,i,ts) $ (not arm_LES_CES(rNat1,i))
        + pmtUArmCom(rNat1,i,ts) $ arm_LES_CES(rNat1,i))
        /m_pmcifPlusTarArmCom(rNat,i,rNat1,ts))**p_sigmaW(rNat1,i)
      --- preference shifter (also termed iceberg cost term)
      * m_lambdam(rNat,i,rNat1,ts)**(p_sigmaW(rNat1,i) - 1) }
      %addCondXwArmCom1%
    ]
    --- add commitment term
    + p_amwCom(rNat,i,rNat1,ts)
  ) s(xw,"rNat,i,rNat1,ts") ]
```

The sole difference in here is that the utility aggregate from the import flows, aka total import demand by the agent, is captured by the variable  $xmt$  for the standard case and  $xmtU$  for the case with commitments. Negative commitment terms can provoke a situation where the adding the commitments to the usual share driven demands from a CES function implies a negative demand quantity. Therefore, the user can switch on a “fugding” mechanism which adds the following code to the equation:

```

    $$iftheni.fudg "%armComFudg%"=="on"
        $ (p_amwCom(rNat,i,rNat1,ts) ge 0)
*
*      -- fudging function to ensure non-negativity in case of negative commitment terms
*
    -(ncpVUSin(min(0,-xw.lo(rNat,i,rNat1,ts)),-[
        ( p_amw(rNat,i,rNat1,ts)
*
*          * ( xmt(rNat1,i,ts) $ (not arm_LES_CES(rNat1,i))
*            + xmtUArmCom(rNat1,i,ts) $ arm_LES_CES(rNat1,i))
*
*          * [( ( pmt(rNat1,i,ts) $ (not arm_LES_CES(rNat1,i))
*            + pmtUArmCom(rNat1,i,ts) $ arm_LES_CES(rNat1,i))
*
*            /m_pmcifPlusTarArmCom(rNat,i,rNat1,ts))**p_sigmaW(rNat1,i)
*
*      --- preference shifter (also termed iceberg cost term)
*
*          * m_lambdam(rNat,i,rNat1,ts)**(p_sigmaW(rNat1,i) - 1)}
*      %addCondXwArmCom1%
*
*      ]
*
*      --- add commitment term
*
*          + p_amwCom(rNat,i,rNat1,ts)
*
*      )
*      ],1.E-7)s(xw,"rNat,i,rNat1,ts")} $ (p_amwCom(rNat,i,rNat1,ts) lt 0)
    $$endif.fudg

```

It applies a Veelken-Ulbrich smoothing function which replaces a *min()* operator to ensure that the trade flows *xw* remain strictly positive. The average price of imports differs between the case with and without commitments, as also seen from the mathematical notation above:

```

*
* --- Average price of imported origin, equation (52) in VDM 2018
*
e_pmtArmCom(rNat1,iIn(i),ts) $ m_hasArmComImp(rNat1,i) ..
*
* --- standard case without commitments
*
    pmt(rNat1,i,ts) $ (not arm_LES_CES(rNat1,i))
*
* --- case with commitments for ARM with commitments
*
    + pmtUArmCom(rNat1,i,ts) $ arm_LES_CES(rNat1,i)
*
    =e=
*
*      --- CES case with p_sigmaW <> 1
*
    [sum(rNat $ (m_CombArmCom(rNat,i,rNat1)
        %addCondXwArmCom2%
    ),
        p_amw(rNat,i,rNat1,ts)
        * (m_pmcifPlusTarArmCom(rNat,i,rNat1,ts)
            /m_lambdam(rNat,i,rNat1,ts))**(1-p_sigmaW(rNat1,i))
        )**(1/(1-p_sigmaW(rNat1,i))) * p_axPmtCD(rNat1,i,ts)
    ] $ ((p_sigmaW(rNat1,i) ne 1) $ p_axPmtCD(rNat1,i,ts))
*
*      --- CD case with p_sigmaW == 1
*
    + [prod(rNat $ (m_CombArmCom(rNat,i,rNat1)
        %addCondXwArmCom2%
    ),
        (m_pmcifPlusTarArmCom(rNat,i,rNat1,ts)
            /((m_lambdam(rNat,i,rNat1,ts)*p_amw(rNat,i,rNat1,ts))
                **p_amw(rNat,i,rNat1,ts))
        ) * p_axPmtCD(rNat1,i,ts)
    ] $ ((p_sigmaW(rNat1,i) eq 1) $ p_axPmtCD(rNat1,i,ts))
*
    %addCondXwArmCom3%
*
;

```





### 3.6 Spatial Price Equilibrium instead of Armington

- Wolfgang Britz, 2021 -

#### 3.6.1 Background

The Armington assumption implies that quality differences are present according to where a product originates from. In many cases, this assumption makes sense already due to the fact that each product in the GTAP Data Base consists of many different tariff lines and that this composition differs across regions. However, there are cases, such as for raw sugar or when working with data bases with high product detail, where the assumption of a homogenous quality and the law of one price might be more appropriate. This assumption can be employed at the level of bi-lateral flows with the spatial arbitrage solution drawing the Spatial Price Equilibrium (SPE) framework. It allows for emerging and vanishing trade flows and overcomes the “small shares stay small” problem. Furthermore, it does not require any behavioral parameters to steer substitution between the different origins. The SPE solution was used in a range of partial equilibrium models, for instance:

- Anania, G., 2010. EU Economic Partnership Agreements and WTO negotiations. A quantitative assessment of trade preference granting and erosion in the banana market. *Food Policy*, 35(2), pp.140-153
- Nolte, S., Buysse, J. and Van Huylenbroeck, G., 2012. Modelling the effects of an abolition of the EU sugar quota on internal prices, production and imports. *European review of agricultural economics*, 39(1), pp.75-94
- Wieck, C., Schlüter, S.W. and Britz, W., 2012. Assessment of the impact of avian influenza–related regulatory policies on poultry meat trade and welfare. *The World Economy*, 35(8), pp.1037-1052
- Jansson, T., Heckeley, T. (2009): A new estimator for trade costs and its small sample properties, *Economic Modelling* 26: 489-498

The assumption of the SPE framework implies first the law of one price, i.e. prices at c.i.f. plus tariff across origins for realized trade flows must be equal, and second, a physical aggregation from bi-lateral flows to total imports. Unobserved trade flows provide a challenge, however. According to the first-order condition of the SPE model, a zero flow occurs if the f.o.b. price plus transaction costs (= trade margins in the GTAP Data Base) plus the tariff exceeds the c.i.f. price plus tariff of the lowest cost importer. Unfortunately, for unobserved trade flows, the size of this difference is unknown as trade data bases (TASTE, MacMaps) underlying most global CGE analysis do not report tariffs for non-observed flows and the GTAP Data Base does not comprise transport margins for such zero observation. This requires assumptions, such as using the largest reported applied tariffs, or some empirical relation between transport margins and transport distance to define benchmark prices in such cases.

Moreover, the application of SPE comes along with numerical challenges. Tests indicate that, for the framework to function, a (stronger) non-linear response in CIF prices to changes in bi-laterally traded quantities is required than found in the standard GTAP model layout. The underlying reasons are as follows. The FOC require that the c.i.f. price plus tariff of any importer is equal to the minimum c.i.f. plus tariff or the importer’s trade flow is zero. Accordingly, a drop in a single c.i.f. price and/or tariff at fixed other prices would imply that all other trade flows disappear. To keep some non-zero flows as a response to such a shock, the related production costs (defines f.o.b.) would need to fall to the point where the minimum bi-lateral price is reached, but this results in turn in new minimum prices for other

export destination of the supplying region where production cost crop. This mechanism also does not work well as production costs hardly react to a moderate change in a specific bi-lateral trade flow as long as this trade flow does not account for a high share of total output. This means that the production cost adjustment response is too low to bring all SPE conditions into equilibrium for an import product market (at least for smaller trade flows). A further reason for changes in import prices relates to the changes in trade cost, but the per unit transport margins and AVEs are independent of traded quantity on a flow (if there is no TRQ mechanisms). The transport margins are endogenous, but their changes depend on price feedback to changes in total global trade. Further, the SPE mechanism typically applies to many bi-lateral flows for a product as GTAP tends to have dense bi-lateral trade matrices.

The proposed implementation renders the per unit bi-lateral transport margin depending on the traded quantities. Potential reasons why this might occur in reality is that longer domestic transport distances are needed to add the supply of more firms to expand trade, or other impacts on marginal costs of international transport if more is traded on a trade link.

### 3.6.2 Equations and technical implementation

The first equation renders the per unit bi-lateral transport cost a function of traded quantities by changing the transport margin shifter *lambdamg*:

```
*
* --- definition of tech shifter for transport margin demand
*
e_lambdamg(m,rNat,iSpe,rNat1,ts) $ (p_amgm(m,rNat,iSpe,rNat1)
$ (lambdamg.range(m,rNat,iSpe,rNat1,ts) ne 0) $ m_combSpe(rNat,iSpe,rNat1)) ..

lambdamg(m,rNat,iSpe,rNat1,ts) =e= 1/(shiftLambdaMg(rNat,iSpe,rNat1,ts)+1);
```

Where the shifter *shiftLambdaMg* is defined by a linear function of the change in the trade flows *xw* relative to its benchmark level. In order to prevent that relative changes in tiny flows leads to large trade cost increases, the largest observed trade flow of an exporter *p\_largestXW* is added.

```
-
e_shiftLambdaMg(rNat,iSpe,rNat1,ts) $ (sum(m,p_amgm(m,rNat,iSpe,rNat1))
$ (sum(m,lambdamg.range(m,rNat,iSpe,rNat1,ts)) ne 0) $ m_combSpe(rNat,iSpe,rNat1)) ..

shiftLambdaMg(rNat,iSpe,rNat1,ts) =e=
([ ( xw(rNat,iSpe,rNat1,ts) + p_largestXW(rNat,iSpe))
/ (sum(t0,xw.l(rNat,iSpe,rNat1,t0)) + p_largestXW(rNat,iSpe))-1]*p_splopeTrsCostSpe);
```

The tariff inclusive bi-lateral imports prices are defined by the following equation and macro:

```
*
* --- definition of bi-lateral import prices
*
e_pmCifPlusTar(rNat,iSpe,rNat1,ts) $ (m_combSpe(rNat,iSpe,rNat1) $ (xw.range(rNat,iSpe,rNat1,ts) ne 0)) ..

pmCifPlusTar(rNat,iSpe,rNat1,ts) =E= m_pmCifPlusTarDef(rNat,iSpe,rNat1,ts);

' --- as per unit transport margins on each flows are endogenous, the definition of the c.i.f. price is more non-linear
' compared to the standard case. Combined with the (almost) convex wx x pmCIF =E= 0 relation, it has turned out that
' introducing the cif price as a variable helps the solver. Additionally, we need the ad-valorem tariff considered
'
$macro m_pmCifPlusTarDef(rNat,iSpe,rNat1,t0) %pmCif%(rNat,iSpe,rNat1,t0) \
* [1 + imptx(rNat,iSpe,rNat1,t0) + mtax(rNat,iSpe,t0)];
```

The condition that the flow must be zero if price differences are found and no price differences are allowed if the flow is non-zero is captured by the following equation paired to the trade flow *xw*. It is used if the model is solved as an MCP.

```

*
* --- spatial arbitrage condition: trade flows can only be non-zero if cif price is equal to import price.
* If the "manual" MCP solution method is used (speMultRef), an explicit slack is introduced
*
e_xwSpe(rNat,iSpe,rNat1,ts) $ ( m_combSpe(rNat,iSpe,rNat1) $ (xw.range(rNat,iSpe,rNat1,ts) ne 0)) ..
    pmCifPlusTar(rNat,iSpe,rNat1,ts)
    * p_speCal(rNat,iSpe,rNat1,ts)
*
* --- if the model is not solved by a MCP solver, the slack is handled by the e_xwSpeMultRef
*
- xwSpeSlack(rNat,iSpe,rNat1,ts) $ (p_solveType ne MCP)
  =E= pmt(rNat1,iSpe,ts);

```

The parameter  $p\_speCal$  is defined at the benchmark to render all tariff inclusive c.i.f. price equal. The fudging term  $xwSpeSlack$  captures the necessary re-definition of the equality into an inequality if the model is solved as an NLP.  $xwSpeSlack$  as defined by the equation  $e\_xwSpeSlack$ . If the slack is non-zero turning the equation above into an inequality, the related trade flow  $xw$  must be zero. A non-zero trade flow requires that  $xwSpeSlack$  is zero (very small). The model supports two different numerical approximation to implement this condition. The first block replaces  $a*b=0$  by  $(a-eps)*b(-eps)=0$ , where  $eps$  is chosen is very small. In this expression,  $a$  is the amount by which the import price of a trade flow exceeds the average (equal minimal) import price and  $b$  is the trade flows. The FOC require that  $a$  and  $b$  cannot be positive at the same time. Subtracting  $eps$  from both prevents that the first derivative become zero when  $a$  or  $b$  is equal to zero which will stall a gradient based solve.

The second block instead uses inbuilt smoothing functions.

```

*
* --- FOC from spatial arbitrage: xw * xwSpeSlack = 0. To improve numerical stability, this is slightly
* differently implemented.
*
e_xwSpeSlack(rNat,iSpe,rNat1,ts) $ ( m_combSpe(rNat,iSpe,rNat1) $ (p_solveType ne MCP)
    $ (xw.range(rNat,iSpe,rNat1,ts) ne 0)) ..
    $Siftheni.smoothingFunction "%smoothingFunctionSpe%"=="multRef"
*
* --- manual implementation of r * s = 0, with fudging term of 1.E-16 to yield non-zero derivatives at r=0 or s=0
* and scaling factor for trade flows. RHS defines the fixed fudging error at the benchmark
*
(xwSpeSlack(rNat,iSpe,rNat1,ts)-muSpe)
    * (xw(rNat,iSpe,rNat1,ts)s(xw,"rNat,iSpe,rNat1,ts")-muSpe)
    =E= sum(t0,(xwSpeSlack.l(rNat,iSpe,rNat1,t0)-muSpe)
        * (xw.l(rNat,iSpe,rNat1,t0)s(xw,"rNat,iSpe,rNat1,ts")-muSpe));
*
$Selse.smoothingFunction
*
* --- inbuilt smoothing functions for f(r,s) = 0 if r or s is zero
* RHS defines the fixed fudging error at the benchmark which is removed in the simulation point
* (the functions are not defined to work well to recover a condition such as f(r,s) = 1.E-5)
*
sum(t0,
    %smoothingFunctionSpe%(xwSpeSlack.l(rNat,iSpe,rNat1,t0)+muSpe,
        xw.l(rNat,iSpe,rNat1,t0)-muSpe,muSpe))
    =E= %smoothingFunctionSpe%(xwSpeSlack(rNat,iSpe,rNat1,ts)+muSpe,
        xw(rNat,iSpe,rNat1,ts)-muSpe,muSpe);
*
$Sendif.smoothingFunction

```

Next, we have the market clearing condition, i.e. adding up total imports from the non-zero flows:

```

*
* --- market clearance for imports, defines average import price
*
e_pmtSpe(rNat1,iSpe,ts) $ ( (pmt.Range(rNat1,iSpe,ts) ne 0) $ m_hasSpeImp(rNat1,iSpe) ) ..
    xmt(rNat1,iSpe,ts)*p_xmtScaleSpe(rNat1,iSpe)s(xmt,"rNat1,iSpe,ts")
    =E= sum(rNat $ m_combSpe(rNat,iSpe,rNat1), xw(rNat,iSpe,rNat1,ts))s(xmt,"rNat1,iSpe,ts");

```

The resulting sub-model is defined as follows:

```

model m_spe "Model equations for Spatial Arbitrage extension" /
*
* --- global provided by target model for additional equations provided by target model
* to update transport costs (or other variables) in the context of the SPE model
*
*$iftheni.addCet "%SPENonLinear%"=="Additive CET"
e_pmcifPlusTar.pmcifPlusTar
*$else.addCET
e_lambdaMg.lambdaMg
e_pmcifPlusTar.pmcifPlusTar
e_shiftLambdaMg.shiftLambdaMg
*$endif.addCET
e_xwSpeSlack.xwSpeSlack
e_xwSpe.xw
e_pmtSpe.pmt
/;
*

```

The additive CET solution is not discussed in here.

Solving these equations inserted in the overall model has proven quite challenging. Therefore, an own additional pre-solve mechanism was implemented. Whereas the standard pre-solve considers one model region, only, the pre-solves for SPE markets solve a global PE model for one product, only. In a loop, each solve takes on of the products treated with the SPE mechanism (*iSpe*) into the model and restrict price changes to +/-20% (changes of beyond this range can usually only occur if some part of the model becomes infeasible).

```

*
* --- reset
*
muSpeOri = muSpe;
muSpe = 0;

*
* --- Note don't write iSpe(kp) as otherwise, implicate loop in fix_for_pe will not work!
*
loop(kp $ iSpe(kp),

    option kill=iIn; iIn(kp) = yes;
    option kill=aIn; aIn(a) $ sum ( (rs,iIn), xFlag(rs,a,iIn)) = YES;
    $$include "model/fix_for_pe.gms"

    ps.lo(rs,iIn,tSim) = ps.l(rs,iIn,tSim) * 0.8;
    ps.up(rs,iIn,tSim) = ps.l(rs,iIn,tSim) * 1.2;

    gtap.optfile = standardOptfile;
    $$iftheni.modelType "%modelType%"=="MCP"
        solveType = MCP;
        solve gtap using MCP;
    $$else.modelType
        solveType = NLP;
        solve gtap using NLP minimizing dummy;
    $$endif.modelType
    $$batinclude "model/closures.gms" noIniVars

    if ((gtap.solveLink eq 6) and (not execError), handles(kp) = gtap.handle);
);

```

The code to collect and check the solutions is not shown in here. Equally, some additional fixing of variables is left out. Afterwards, prices and quantities in all markets updated, before the full model is solved.

### 3.6.3 Interface

The interface allows, as shown below:

- To select which products are subject to the SPE implementation (this excludes to use of other trade options, such as Melitz, Armington with commitments or the MRIO extension, for the chosen products)
- If the non-linearities stem from updating the transport margins or from using an additive CET (not discussed in here)

- If pre-solves for the SPE markets should be used
- and the Mu parameter used during pre-solves

Commodities

**SPE**  
commodities

|           |           |            |             |            |           |           |           |
|-----------|-----------|------------|-------------|------------|-----------|-----------|-----------|
| fsh-c     | pdr-c     | wht-c      | barl-c      | maiz-c     | sorg-c    | oat-c     | rye-c     |
| ocer-c    | soy-c     | palm-c     | rape-c      | olv-c      | oso-c     | pota-c    | rttb-c    |
| leg-c     | oveg-c    | toma-c     | appl-c      | grap-c     | citr-c    | banp-c    | ofru-c    |
| v_fo-c    | c_b-c     | pfb-c      | coco-c      | teas-c     | coff-c    | ocro-c    | cattle-c  |
| orum-c    | pig-c     | poul-c     | oapo-c      | rmk-c      | wol-c     | frs-c     | coa-c     |
| oil-c     | gas-c     | oxt-c      | ctilMeat-c  | orumMeat-c | Pigmeat-c | othMeat-c | olivOil-c |
| palmOil-c | soyCake-c | soyOil-c   | rapCake-c   | rapOil-c   | volc-c    | mil-c     | pcr-c     |
| sgr-c     | ofdAnim-c | ofdOther-c | b_t-c       | tex-c      | wap-c     | lea-c     | lum-c     |
| ppp-c     | p_c-c     | chm-c      | bph-c       | rpp-c      | nmm-c     | i_s-c     | nfm-c     |
| fmp-c     | ele-c     | eeq-c      | ome-c       | mvh-c      | otn-c     | omf-c     | gdt-c     |
| wtr-c     | cns-c     | trd-c      | afs-c       | otp-c      | wtp-c     | atp-c     | whs-c     |
| cmn-c     | ofi-c     | ins-c      | rsa-c       | obs-c      | ros-c     | osg-c     | edu-c     |
| hht-c     | dwe-c     | TnD-c      | NuclearBL-c | CoalBL-c   | GasBL-c   | WindBL-c  | HydroBL-c |
| OilBL-c   | OtherBL-c | GasP-c     | HydroP-c    | OilP-c     | SolarP-c  |           |           |

Non-linearities Transport margins ▾

Additional presolve for SPE markets ☒

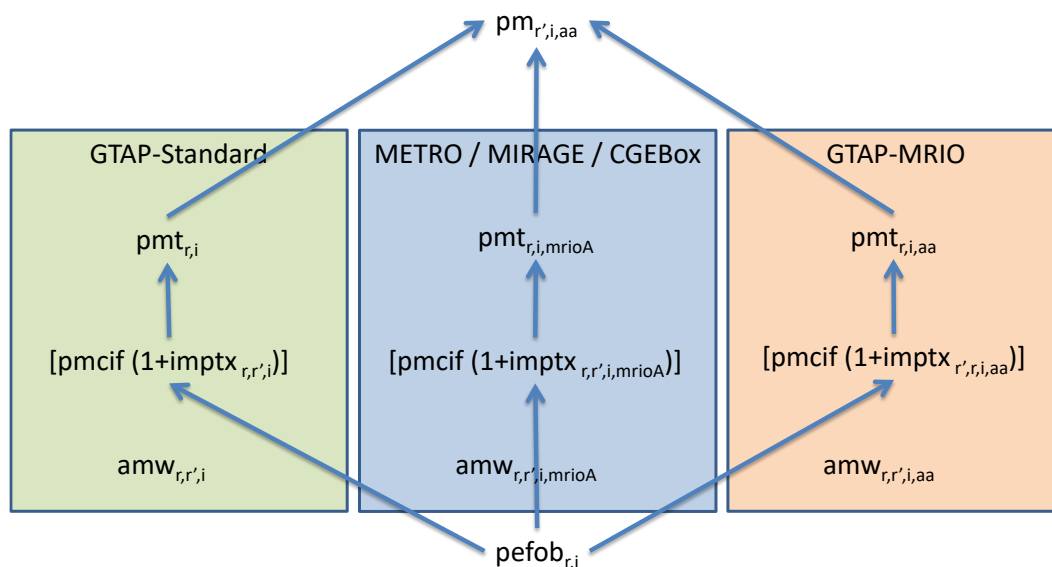
Mu, 1.E- 10 ▾

### 3.7 MRIO extension

- Wolfgang Britz, 2021 –

#### 3.7.1 Background

There is growing interest to better analysis global value chains and to conduct life-cycle assessment along such chains. This requires information on the bi-lateral sourcing of intermediate demand by industries and final demand agents. We discuss in here a module for CGEBox that is able to depict import demand differentiated by aggregate Armington agents, i.e. for aggregate intermediate use, household, government and investments. The module is set-up such that it can be used for specific product markets only and remains compatible with the other trade specification (standard Armington, Armington with commitment terms, Meltitz) but not with the other SPE specification. The selection of the markets is done via the GUI. The necessary data stem now from the GTAP MRIO data base (Carrico et al. 2020<sup>30</sup>) and can be added on demand during the data preparation step.



The graph above compares the chosen implementation (in the middle) with the GTAP-Standard approach and a fully-fledged MRIO implementation as in GTAP-MRIO (Corong 2020<sup>31</sup>). In the GTAP-Standard model, all Armington agents face the same import prices as tariffs and bi-lateral trade shares are not differentiated by agent. In the full MRIO implementation, tariffs and bi-lateral trade shares can differ for each Armington agent, whereas in the chosen implementation, they are identical across production sectors, but can differ between intermediate, household, government and final demand, according to the so-called BEC classification at the tariff line. The chosen implementation considers that the GTAP-MRIO Data Base uses proportionality assumption to split the total (i.e. aggregated over production activities) bi-lateral intermediate input demand of a product to production activities according to their total import demand (i.e. not differentiated by region of origin). Hence, the chosen model layout is aligned with the informational content of the underlying data base.

<sup>30</sup> Carrico, C., Corong, E., & van der Mensbrugghe, D. (2020). The GTAP version 10A Multi-Region Input Output (MRIO) Data Base. GTAP Research Memorandum No. 34, [https://www.gtap.agecon.purdue.edu/resources/res\\_display.asp?RecordID=6164](https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=6164)

<sup>31</sup> Corong, E. (2020). The GTAP Multi-Region Input Output (MRIO) Model. GTAP Research Memorandum No. 37.

A comparison exercise with the GTAP-MRIO model has shown that the extension in CGEBox and the GTAP-MRIO model deliver exactly the same results to shock if they use the same data base and parameterization and no additional modules are active in CGEBox.

### 3.7.2 Definition of consistent split factors

In order to define a consistent benchmark for the model, the total import demands of the MRIO agents need to be exhausted by their bi-lateral demands for imports while at the same time the given bi-lateral export flows must be exhausted. We apply a Linear Loss framework to derive consistent shares (see *gams/build/mrio.gms*) instead of simply copying data derived from the GTAP-MRIO Data Base. A dedicated consistency step is necessary, first, as the filter program might change the value of bi-lateral import demands, and second, as the GTAP-MRIO Data Base is not fully aligned with the standard version. A specific problem arises from the fact that the filter process is not aware of the GTAP-MRIO information. Assume for instance that the GTAP-MRIO reports a bi-lateral trade flow for household demand but that the overall import demand of households for this product is quite small in relative terms. The filter algorithm might remove the import demand by the aggregate household, but not the bi-lateral import flow of this product. As a consequence, there will a non-zero import flow in the SAM, but no import demand for a MRIO agent (here households) for which the GTAP-MRIO Data Base reports a bi-lateral import demand. This requires partly that the code is able to “invent” new MRIO flows as discussed below. Such cases are typical few and should relate to tiny flows anyhow. As same product detail is found in both data bases, a 1:1 mapping can be used. In case of additional product splits, currently, the same split factors are used.

As a first step, we aggregate the GTAP-MRIO Data Base into a parameter termed “*mrioSplitFactors*” and stored in GDX container. It comprises in its raw format information on the value of bi-lateral trade for each of the four MRIO agents (total intermediate demand *int*, households *hhsld*, government *gov* and investment demand *inv*) at c.i.f. and c.i.f. plus tariffs prices. Note that also tariffs can differ across MRIO agents according to tariff dispersion at HS6 level.

As a first step, we aggregate the value at c.i.f. and c.i.f. plus tariffs, as captured by the set *priceLevel* from GTAP region *rDat* to model region *rNat* according to the mapping set *rr*:

```
mrioSplitFactors(rNat1,i,rNat,mrioA,priceLevel)
$ ( (sam0(rNat1,rNat,i)+imptxY0(rNat,rNat1,i)) and sum(mrioA_aa(mrioA,aa), xmm0(rNat1,i,aa)))
= sum( (rr(rDat,rNat),rr1(rDat1,rNat1),mm(prod_comm,i)), p_splitFactorsG(prod_comm,mrioA,rDat,rDat1,priceLevel));
```

In case of a split, bi-lateral values at c.i.f. are used as the allocation key:

```
*
* --- scale MRIO factors for split up products according to share on bi-lateral imports
*
mrioSplitFactors(rNat1,spliti,rNat,mrioA,priceLevel) $ mrioSplitFactors(rNat1,spliti,rNat,mrioA,priceLevel)
= mrioSplitFactors(rNat1,spliti,rNat,mrioA,priceLevel)
  * sam0(rNat1,rNat,spliti)/sum(is_i(spliti,i) $ is_i(spliti,i),sam0(rNat1,rNat,spliti));
```

The balancing of this information against the given SAM and the information on the import demand at market prices for the production activities and final demand categories is based on a LP framework with a linear loss objective. The reason is that performant LP solvers can be used in this case (see Britz 2021<sup>32</sup> for a discussion).

---

<sup>32</sup> Britz, W. (2021): Comparing Penalty Functions in Balancing and Dis-aggregating Social Accounting Matrices, Journal of Global Economic Analysis 6(1): 34-81



A first constraint  $e\_xmMrio$  ensures that total import demand at market prices for each MRIO agent (i.e. at c.i.f. plus tariffs), when aggregated over the SAM agents linked to a MRIO agent (a 1:1 relation for the three final demand categories, for *int* it is mapped to all production activities) as defined on the LHS must be exhausted by the MRIO transactions on the RHS:

```
*
* --- Exhaust total import demand by MRIO agent by the bi-lateral flows at cif+tariff
*
e_xmMrio(rnat,curI(i),mrioA) $ sum(mrioA_aa(mrioA,aa), xmm0(rnat,i,aa)) ..
    sum(mrioA_AA(mrioA,aa), xmm0(rnat,i,aa))/max(1,sum(mrioA_AA(mrioA,aa), xmm0(rnat,i,aa)))
    =E= [sum(rnat1 $ mrioSplitFactors(rNat,i,rNat1,mrioA,"m"),
        mrioSplitFactors(rNat,i,rNat1,mrioA,"m")) * (1+v_corP(rNat,i,rNat1,mrioA,"m")-v_corN(rNat,i,rNat1,mrioA,"m"))
        -sum(rnat1, (mrioSplitFactors(rNat,i,rNat1,mrioA,"m")+mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")) * v_slack1(rNat,i,rNat1,mrioA))
    *
    * --- these are originally not observed entries. Their multiplier is pulled towards 0 => expected value is zero
    *
    + sum(rnat1 $ mrioSplitFactors(rNat,i,rNat1,mrioA,"m1"),
        mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")) * v_mult(rNat,i,rNat1,mrioA,"m1"))/max(1,sum(mrioA_AA(mrioA,aa), xmm0(rnat,i,aa)))
    ]
```

The RHS comprises three terms:

- The originally reported transactions from the GTAP-MRIO Data Base at market prices as indicated by “m” on the parameter *mrioSplitFactors*, multiplied with a correction factor which is split into a positive (*v\_corP*) and negative (*v\_corN*) relative correction.
- A slack which allows that values at c.i.f. and c.i.f. plus tariffs are not fully consistent, to capture potential small inaccuracies in the SAM.
- A term which allows introducing new MRIO flows, as motivated above.

Two equations exhaust bi-lateral trade as found in the SAM aggregated from the GTAP Data Base at c.i.f. and c.i.f. plus tariffs from the bi-lateral import demands of the MRIO agents.

```
*
* --- exhaustion of exports at c.i.f. value by the MRIO agents
*
e_cif(considerMrio(rnat,curI(i),rnat1)) $ sam0(rNat,rNat1,i) ..
    sam0(rnat,rnat1,i)/max(1,sam0(rnat,rnat1,i))
    =E= [ sum(mrioA $ mrioSplitFactors(rNat,i,rNat1,mrioA,"cif"),
        mrioSplitFactors(rNat,i,rNat1,mrioA,"cif")) * (1+v_corP(rNat,i,rNat1,mrioA,"cif")-v_corN(rNat,i,rNat1,mrioA,"cif"))
    *
    * --- these are originally not observed entries. Their negative correction factor is pulled toward -1 => expected value is zero
    *
    + sum(mrioA $ mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1"),
        mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1")) * v_mult(rNat,i,rNat1,mrioA,"cif1"))/max(1,sam0(rnat,rnat1,i));
    ]
```

For the case of c.i.f. plus tariffs, again a slack is introduced to reflect potential tiny imbalances in the SAM:

```
*
* --- exhaustion of exports at c.i.f. value + tariffs by the MRIO agents
*
e_cifM(considerMrio(rnat,curI(i),rnat1)) $ (sam0(rNat,rNat1,i)+imptxY0(rNat1,rNat,i)) ..
    [sam0(rnat,rnat1,i)+imptxY0(rNat1,rNat,i)]/max(sam0(rnat,rnat1,i)+imptxY0(rNat1,rNat,i),1)
    =E= [ sum(mrioA $ mrioSplitFactors(rNat,i,rNat1,mrioA,"m"),
        mrioSplitFactors(rNat,i,rNat1,mrioA,"m")) * (1+v_corP(rNat,i,rNat1,mrioA,"m")-v_corN(rNat,i,rNat1,mrioA,"m"))
        + sum(mrioA $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")+mrioSplitFactors(rNat,i,rNat1,mrioA,"m")),
            (mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")+mrioSplitFactors(rNat,i,rNat1,mrioA,"m")) * v_slack2(rNat,i,rNat1))
        + sum(mrioA $ mrioSplitFactors(rNat,i,rNat1,mrioA,"m1"),
            mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")) * v_mult(rNat,i,rNat1,mrioA,"m1"))
    ]/max(1,sam0(rnat,rnat1,i)+imptxY0(rNat1,rNat,i));
    ]
```

There is an additional inequality constraint which ensures that the tariff-inclusive value exceeds value “m” exceeds the value at c.i.f. as otherwise, negative ad-valorem tariffs would be estimated

```
*
* --- exports at c.i.f. value + tariffs must exceed exports at c.i.f. value for each agent to ensure positive import tariffs
* (slack not considered such it can contribute to feasibility)
*
e_cifLeM(considerMrio(rNat,curI(i),rnat1),mrioA) $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"m") or mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")) ..
    [ mrioSplitFactors(rNat,i,rNat1,mrioA,"m") * (1+v_corP(rNat,i,rNat1,mrioA,"m")-v_corN(rNat,i,rNat1,mrioA,"m"))
    + mrioSplitFactors(rNat,i,rNat1,mrioA,"m1") * v_mult(rNat,i,rNat1,mrioA,"m1"))
    /max(1, mrioSplitFactors(rNat,i,rNat1,mrioA,"m")+ mrioSplitFactors(rNat,i,rNat1,mrioA,"m1"))
    =G= [ mrioSplitFactors(rNat,i,rNat1,mrioA,"cif") * (1+v_corP(rNat,i,rNat1,mrioA,"cif")-v_corN(rNat,i,rNat1,mrioA,"cif"))
        + mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1") * v_mult(rNat,i,rNat1,mrioA,"cif1"))
        - (mrioSplitFactors(rNat,i,rNat1,mrioA,"cif")+mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1")) * v_slack(rNat,i,rNat1,mrioA)
    ]/max(1, mrioSplitFactors(rNat,i,rNat1,mrioA,"m")+ mrioSplitFactors(rNat,i,rNat1,mrioA,"m1"));
```



An additional inequality ensures that the corrections do not lead to implausible increases in tariffs by large differences between the correction on the “m” (= c.i.f. plus tariffs) and “cif” positions. The maximum allowed relative change is defined by the scalar  $p\_tarMaxDiff$ :

```
*
* --- plausible values for newly introduce tariffs
*
e_cifGeT(considerMrio(rNat,curI(i),rNat1),mrioA) $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"m") or mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")) ..
[ mrioSplitFactors(rNat,i,rNat1,mrioA,"m") * (1+v_corP(rNat,i,rNat1,mrioA,"m")-v_corN(rNat,i,rNat1,mrioA,"m"))
+ mrioSplitFactors(rNat,i,rNat1,mrioA,"m1") * v_mult(rNat,i,rNat1,mrioA,"m1")]
/max(1, mrioSplitFactors(rNat,i,rNat1,mrioA,"m")+ mrioSplitFactors(rNat,i,rNat1,mrioA,"m1"))

=L=
[ mrioSplitFactors(rNat,i,rNat1,mrioA,"cif") * (1+v_corP(rNat,i,rNat1,mrioA,"cif")-v_corN(rNat,i,rNat1,mrioA,"cif"))
+ mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1") * v_mult(rNat,i,rNat1,mrioA,"cif1")]
*
* --- less than maximum found in data
*
* [ smax(mrioA1 $ mrioSplitFactors(rNat,i,rNat1,mrioA1,"cif"),
mrioSplitFactors(rNat,i,rNat1,mrioA1,"m")/mrioSplitFactors(rNat,i,rNat1,mrioA1,"cif"))
$ sum(mrioA1, mrioSplitFactors(rNat,i,rNat1,mrioA1,"cif"))
+ ((sam0(rnat,rnat1,i)+imptxY0(rNat1,rNat,i))/sam0(rnat,rnat1,i))
$ (not sum(mrioA1, mrioSplitFactors(rNat,i,rNat1,mrioA1,"cif"))))
*
* -- plus some correction factors
*
+ max(p_tarMaxDiff,p_tarMaxDiff * ((sam0(rnat,rnat1,i)+imptxY0(rNat1,rNat,i))/sam0(rnat,rnat1,i)-1))
]
* 1/max(1, mrioSplitFactors(rNat,i,rNat1,mrioA,"m")+mrioSplitFactors(rNat,i,rNat1,mrioA,"m1"));
```

The objective function minimizes the linear loss:

```
*
* --- minimize linear loss
*
e_objMrio ..
v_objeMrio * sum( (considerMrio(rNat,curI(i),rNat1),mrioA,priceLevel1) $ mrioSplitFactors(rNat,i,rNat1,mrioA,priceLevel1),1)
=E=
*
* --- sum up pos/neg deviation relative to given MRIO trade flows
*
sum( (considerMrio(rNat,curI(i),rNat1),mrioA,priceLevel) $ mrioSplitFactors(rNat,i,rNat1,mrioA,priceLevel),
v_corP(rNat,i,rNat1,mrioA,priceLevel)+v_corN(rNat,i,rNat1,mrioA,priceLevel))
*
* --- pull newly introduce ones to zero (1 is added to the neg. dev = pulled against -1)
*
+ sum( (considerMrio(rNat,curI(i),rNat1),mrioA,priceLevelN) $ mrioSplitFactors(rNat,i,rNat1,mrioA,priceLevelN),
v_mult(rNat,i,rNat1,mrioA,priceLevelN))*p_slackWeight**(1/4)
*
* --- sum up pos/neg slacks against import totals (this should not happen!), high penalty
*
+ sum( (considerMrio(rNat,curI(i),rNat1),mrioA)
$ (mrioSplitFactors(rNat,i,rNat1,mrioA,"m") or mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")),
(mrioSplitFactors(rNat,i,rNat1,mrioA,"cif")+mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1"))
* (v_slack(rNat,i,rNat1,mrioA)+v_slack1(rNat,i,rNat1,mrioA)*p_slackWeight)*p_slackWeight)
+ sum( (considerMrio(rNat,curI(i),rNat1),v_slack2(rNat,i,rNat1))*p_slackWeight);
```

i.e. the positive and negative corrections are pulled towards zero while the two slacks receive a quite high value:

```
p_slackWeight=1E+6;
```

The problem is solved product by product in a loop:

```
loop(j $ sum( (considerMrio(rNat,j,rNat1)),1),
option kill=curI;curI(j) = yes;
m_mrio.optfile = 5;
m_mrio.limcol = %limcol%;
m_mrio.limrow = %limrow%;
$$iftheni.solprint "%solprint%"=="silent"
m_mrio.solprint = 2;
$$elseifi.solprint "%solprint%"=="off"
m_mrio.solprint = 0;
$$else.solprint
m_mrio.solprint = 1;
$$endif.solprint

if ( execerror,execError = 0);
p_mMrioStat(j,"nTries") = 1;

p_tarMaxDiff = 0.1;

solve m_mrio using LP minimizing v_objeMrio;
```

In case of infeasibilities:

(1) Missing split factors “cif1” are added assuming uniform tariffs:

```
mrioSplitFactors(rNat,j,rNat1,mrioA,"cif1") $ ((not mrioSplitFactors(rNat,j,rNat1,mrioA,"cif1")) $ mrioSplitFactors(rNat,j,rNat1,mrioA,"m1")
= mrioSplitFactors(rNat,j,rNat1,mrioA,"m1") * sam0(rnat,rnat1,j)/(sam0(rnat,rnat1,j)+imptxY0(rNat1,rNat,j));
mrioSplitFactors(rNat,j,rNat1,mrioA,"m1") $ ((not mrioSplitFactors(rNat,j,rNat1,mrioA,"m1")) $ mrioSplitFactors(rNat,j,rNat1,mrioA,"cif1"))
= mrioSplitFactors(rNat,j,rNat1,mrioA,"cif1")*(sam0(rnat,rnat1,j)+imptxY0(rNat1,rNat,j))/sam0(rnat,rnat1,j);
```

(2) The maximal tariff difference for newly introduced flows is increased:

```
p_tarMaxDiff = 0.25;
```

If the problem is still infeasible, this maximal difference is increased to 50% and then 100%, in the latter case, also the upper limit on negative corrections is removed. As expected, normally the problem will introduce almost no changes to the original MRIO data.

The code collects information on the final successful solve of each problem, i.e. product:

```

',
p_mMrioStat(j,"maxTariff") = p_tarMaxDiff;
p_mMrioStat(j,"sumInfes") = m_mrio.suminfes;
p_mMrioStat(j,"numInfes") = m_mrio.numinfes;
p_mMrioStat(j,"obj") = v_objeMrio.1;
```

It also reports the cases where MRIO flows had to be invented in order to derive a consistent set of MRIO data matching the filter GTAP Data Base:

```
parameter p_new;
p_new(considerMrio(rNat,i,rNat1),mrioA,"m") $ mrioSplitFactors(rNat,i,rNat1,mrioA,"m1")
= mrioSplitFactors(rNat,i,rNat1,mrioA,"m1") * v_mult.1(rNat,i,rNat1,mrioA,"m1");

p_new(considerMrio(rNat,i,rNat1),mrioA,"cif") $ mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1")
= mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1") * v_mult.1(rNat,i,rNat1,mrioA,"cif1");

p_new(considerMrio(rNat,i,rNat1),mrioA,"tar") $ p_new(rNat,i,rNat1,mrioA,"cif")
= p_new(rNat,i,rNat1,mrioA,"m")/p_new(rNat,i,rNat1,mrioA,"cif");

p_new(considerMrio(rNat,i,rNat1),mrioA,"tarAvg") $ p_new(rNat,i,rNat1,mrioA,"cif")
= (sam0(rnat,rnat1,i)+imptxY0(rNat1,rNat,i))/sam0(rnat,rnat1,i);
option p_new:2:4:1;
display p_new;
```

The maximum of these cases is also reported:

```
parameter p_maxNew;
p_maxNew("abs") = smax( (considerMrio(rnat,i,rnat1),mrioA,priceLeveln), v_mult.1(rnat,i,rnat1,mrioA,priceLeveln));
p_maxNew("rel") = smax( (considerMrio(rnat,i,rnat1),mrioA) $ (v_mult.1(rnat,i,rnat1,mrioA,"cif1") $ mrioSplitFactors(rNat,i,rNat1,mrioA,"cif")),
v_mult.1(rnat,i,rnat1,mrioA,"cif1")/mrioSplitFactors(rNat,i,rNat1,mrioA,"cif"));
p_maxNew("rel") = max( p_maxNew("rel"),
smax( (considerMrio(rnat,i,rnat1),mrioA) $ (v_mult.1(rnat,i,rnat1,mrioA,"m1") $ mrioSplitFactors(rNat,i,rNat1,mrioA,"m")),
v_mult.1(rnat,i,rnat1,mrioA,"m1")/mrioSplitFactors(rNat,i,rNat1,mrioA,"m")));
display p_maxNew;
```

After the solution, transactions are defined based on the equation structure and very tiny ones deleted:

```
mrioSplitFactors(considerMrio(rNat,i,rNat1),mrioA,"m") $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"m") or mrioSplitFactors(rNat,i,rNat1,mrioA,"m1"))
= mrioSplitFactors(rNat,i,rNat1,mrioA,"m") * (1+v_corP.1(rNat,i,rNat1,mrioA,"m") - v_corN.1(rNat,i,rNat1,mrioA,"m"))
+ mrioSplitFactors(rNat,i,rNat1,mrioA,"m1") * v_mult.1(rNat,i,rNat1,mrioA,"m1");
mrioSplitFactors(considerMrio(rNat,i,rNat1),mrioA,"m") $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"m") le 1.E-10) = 0;
```

The same for the value at c.i.f.:

```
mrioSplitFactors(considerMrio(rNat,i,rNat1),mrioA,"cif") $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"cif") or mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1"))
= mrioSplitFactors(rNat,i,rNat1,mrioA,"cif") * (1+v_corP.1(rNat,i,rNat1,mrioA,"cif") - v_corN.1(rNat,i,rNat1,mrioA,"cif"))
+ mrioSplitFactors(rNat,i,rNat1,mrioA,"cif1") * v_mult.1(rNat,i,rNat1,mrioA,"cif1");
mrioSplitFactors(considerMrio(rNat,i,rNat1),mrioA,"cif") $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"cif") le 1.E-10) = 0;
mrioSplitFactors(considerMrio(rNat,i,rNat1),mrioA,"m") $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"cif") le 1.E-10) = 0;
```

Afterwards, potential cases with negative tariffs are reported:

```
* --- recalculate slack variables necessary to ensure that m (=c.i.f. + tariffs) >= cif
*
option kill=v_slack.1;
v_slack.1(considerMrio(rnat,i,rnat1,mrioA),mrioA) $ (mrioSplitFactors(rNat,i,rNat1,mrioA,"cif") gt mrioSplitFactors(rNat,i,rNat1,mrioA,"m"))
= mrioSplitFactors(rNat,i,rNat1,mrioA,"cif") - mrioSplitFactors(rNat,i,rNat1,mrioA,"m");

option v_slack:7:3:1;
v_slack.1(considerMrio(rnat,i,rnat1,mrioA),mrioA) $ (v_slack.1(rnat,i,rnat1,mrioA) le 1.E-8) = 0;
display v_slack.1;

parameter p_maxSlack;
p_maxSlack("abs") = smax( (considerMrio(rnat,i,rnat1,mrioA), v_slack.1(rnat,i,rnat1,mrioA));
p_maxSlack("rel") = smax( (considerMrio(rnat,i,rnat1,mrioA) $ (v_slack.1(rnat,i,rnat1,mrioA) $ mrioSplitFactors(rNat,i,rNat1,mrioA,"m")),
v_slack.1(rnat,i,rnat1,mrioA)/mrioSplitFactors(rNat,i,rNat1,mrioA,"m"))*100;
display p_maxSlack;
```

Finally, the information so far in the total value of the transaction is mapped into shares:

```
*
mrioSplitFactors(rNat,i,rNat1,mrioA,priceLevel) $ { mrioSplitFactors(rNat,i,rNat1,mrioA,priceLevel) $ sam0(rNat,rNat1,i))
= mrioSplitFactors(rNat,i,rNat1,mrioA,priceLevel)/sum(mrioA1, mrioSplitFactors(rNat,i,rNat1,mrioA1,priceLevel));
```

### 3.7.3 Model equations

The model equations in the MRIO extension of CGEBox are equivalent to the second level Armington equations in the standard model, but are now differentiated by the MRIO agents (total intermediate demand *int*, households *hhsld*, government *gov* and investment demand *inv*). Note that the extension can be used with the “Armington with commitment terms” extension (indicated by the tests “%modulesArm\_Les\_ces%”==”on”).

Bi-lateral import demand found on the RHS of the equation *e\_xwAggMrio* is expressed via the usual share equations on the LHS for each MRIO agent *mrio*. Their sum defines the total import flow:

```
e_xwAggMrio(rNat1,iIn(iMrio),rNat,ts) $ (xwFlag(rNat1,iMrio,rNat) $ rrComb(rNat1,rNat)) ..

--- in the standard case, use the xw variable. If a negative commitment is found, define
xwToFudge which can become negative due to the commitment term

(xw(rNat1,iMrio,rNat,ts)
  $$iftheni.fudg "%modulesArm_Les_Ces%=="on"
  $ (p_amwCom(rNat1,iMrio,rNat,ts) ge 0)
  + xwToFudge(rNat1,iMrio,rNat,ts) $ (p_amwCom(rNat1,iMrio,rNat,ts) lt 0)
  $$endif.fudg
) s(xw,'rNat1,iMrio,rNat,ts')

=e=
```

```

sum(mrioA $ p_amwMrio(rNat1,iMrio,mrioA,rNat,ts),

--- equivalent to a standard CES, i.e. share parameter times total times price
    relation exponent the subs. elasticity

p_amwMrio(rNat1,iMrio,mrioA,rNat,ts)

    --- for the Armington with commitment terms, we distributed
        a differently determined utility aggregate

    * (
        xmtMrio(rNat,iMrio,mrioA,ts)
        $$iftheni.armLesCes "%ModulesArm_Les_Ces%"=="on"
        $ (not arm_LES_CES(rnat,iMrio))
        + xmtUMrio(rNat,iMrio,mrioA,ts) $ arm_LES_CES(rnat,iMrio)
        $$endif.armLesCes
    )s(xw,'rNat1,iMrio,rNat,ts')

    * [ (( pmtMrio(rNat,iMrio,mrioA,ts)
        $$iftheni.armLesCes "%ModulesArm_Les_Ces%"=="on"
        $ (not arm_LES_CES(rnat,iMrio))
        + pmtUMrio(rnat,iMrio,mrioA,ts) $ arm_LES_CES(rnat,iMrio)
        $$endif.armLesCes
    )

    -- note here the agent specific tariff rate
        (imptx is the weighted average defined below)

    /[(1 + p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts) + mtax(rNat,iMrio,ts))

    --- Note: cif price is defined via macro %pmCIF%, not a variable in model

    * %pmCIF(rNat1,iMrio,rNat,ts) ])**p_sigmax(rNat,iMrio)
    * m_lambdam(rNat1,iMrio,rNat,ts)**(p_sigmax(rNat,iMrio) - 1)

$$iftheni.smallShare "%hasSmallShare%"=="on"
    $ (not isSmallImpShare(rNat1,iMrio,rNat))
    + m_lambdam(rNat1,iMrio,rNat,ts) $ isSmallImpShare(rNat1,iMrio,rNat)
$$endif.smallShare
    ])

$$iftheni.fudg "%modulesArm_Les_Ces%"=="on"
    --- commitment term from Armington with Commitment term extension
    + [p_amwCom(rNat1,iMrio,rNat,ts)s(xw,'rNat1,iMrio,rNat,ts')] $ (p_amwCom(rNat1,iMrio,rNat,ts) ge 0)
$$endif.fudg
;

```

In mathematical notation without considering the special cases (commitment terms, small shares):

$$x_{w,r,i} = \sum_{mrioA} amw_{r,r',i,mrioA} x_{mt,r,i,mrioA} \left( \frac{p_{mt,r,i,mrioA}}{pmCIF_{r,r',i}(1 + impt_{r,r',i,mrioA})} \right)^{\sigma_{w,r,i}} \quad (49)$$

In order to consider the case of the CES with commitment terms, the import totals are either  $x_{mt}Mrio$  (standard case) or  $x_{mt}Mrio$  (if commitment terms are active). These additive commitment terms are added with the last term in the equation. The main differences to the equation in GTAP standard are share parameters  $p\_amwMrio$  and tariffs  $imptxMrio$  differentiated by MRIO agent. Note that the c.i.f. prices are not differentiated by MRIO agent, i.e. the exporters do not quality differentiated by MRIO destination (infinite transformation).

The total import demand  $x_{mt}Mrio$  which is distributed by the share equation above is aggregated over the individual agents linked to the MRIO agent, which is for intermediate input demand, only, different from a 1:1 relation:

```

*
* --- Aggregate import demand over MRIO agents to total import demand
*      (= identity for hhsld, gov, sav which are a single MRIO agent), equivalent to e_xmt in standard model
*
e_xmtMrio(rNat,iMrio,mrioA,ts) $ (xmtFlag(rNat,iMrio) $ sum(mrio_a(mrioA,aa), p_alpham(rNat,iMrio,aa,ts)
*      -- household at sub-regional level
*      $$if1 "%hsubr%"=="on" + sum(r_r(subr,rNat), p_alpham(subr,iMrio,aa,ts))
*      ) $ iIn(iMrio)) ..

xmtMrio(rNat,iMrio,mrioA,ts) s(xmtMrio,'rNat,iMrio,mrioA,ts')

=e= sum(mrio_a(mrioA,aa), xm(rNat,iMrio,aa,ts) $ (p_alpham(rNat,iMrio,aa,ts) gt eps)
*      $$if1 "%hsubr%"=="on" + sum(r_r(subr,rNat) $ (p_alpham(subr,iMrio,aa,ts) gt eps), xm(subr,iMrio,aa,ts))
*      ) s(xmtMrio,'rNat,iMrio,mrioA,ts');

```

In mathematical notation:

$$xmt_{r,i,mrioA} = \sum_{aa \in mrioA} xm_{r,i,aa} \quad (50)$$

The average import price of each MRIO agent is defined via the dual price aggregator and needs to differentiate between the case with (*pmtUMrio*) and without commitment terms (*pmtMrio*):

```

--- dual price aggregator by MRIO agent (considers case with commitment terms)

e_pmtMrio(rNat,iIn(iMrio),MrIoA,ts) $ ( (pmtMrio.range(rNat,iMrio,mrioA,ts) ne 0)
*      $ sum(rNat1, p_amwMrio(rNat1,iMrio,mrioA,rNat,ts))
*      $ xmtFlag(rNat,iMrio) $ (not p_pmtElas(rNat,iMrio)) ) ..

pmtMrio(rNat,iMrio,mrioA,ts)
*      $$ifthen1.armLesCes "%ModulesArm_Les_Ces%"=="on"
*      $ (not arm_LES_CES(rNat,iMrio))
*      + pmtUMrio(rNat,iMrio,mrioA,ts) $ arm_LES_CES(rNat,iMrio)
*      $$endif.armLesCes

=e=

--- CES case with p_sigmax <> 1

{ [sum(rNat1 $ (p_amwMrio(rNat1,iMrio,mrioA,rNat,ts)
*      $$ifthen1.smallShare "%hasSmallShare%"=="on"
*      $ (not isSmallImpShare(rNat1,iMrio,rNat))
*      $$endif.smallShare
*      ) , p_amwMrio(rNat1,iMrio,mrioA,rNat,ts)
*      * ((1 + p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts) + mtax(rNat,iMrio,ts))
*      * %pmcif%(rNat1,iMrio,rNat,ts) ]/m_lambdam(rNat1,iMrio,rNat,ts)) ** (1-p_sigmax(rNat,iMrio))
*      ** (1/(1-p_sigmax(rNat,iMrio))) * p_axPmtCDMrio(rNat,iMrio,mrioA,ts) )
*      $ $ifthen1.smallShare "%hasSmallShare%"=="on"
*      $ sum(rNat1 $ (xwFlag(rNat1,iMrio,rNat) $ (not isSmallImpShare(rNat1,iMrio,rNat))), 1)
*      $ $endif.smallShare
*      ) $ (p_sigmax(rNat,iMrio) ne 1)

*
*      --- CD case with p_sigmax == 1
*
+ [prod(rNat1 $ (p_amwMrio(rNat1,iMrio,mrioA,rNat,ts)
*      $$ifthen1.smallShare "%hasSmallShare%"=="on"
*      $ (not isSmallImpShare(rNat1,iMrio,rNat))
*      $$endif.smallShare
*      ) ,
*      ((1 + p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts) + mtax(rNat,iMrio,ts)) * %pmcif%(rNat1,iMrio,rNat,ts) )
*      / (m_lambdam(rNat1,iMrio,rNat,ts) * p_amw(rNat1,iMrio,rNat,ts)) ** p_amwMrio(rNat1,iMrio,mrioA,rNat,ts)
*      ) * p_axPmtCDMrio(rNat,iMrio,mrioA,ts) ]
*      $ ((p_sigmax(rNat,iMrio) eq 1)
*      $ $ifthen1.smallShare "%hasSmallShare%"=="on"
*      $ sum(rNat1 $ (xwFlag(rNat1,iMrio,rNat) $ (not isSmallImpShare(rNat1,iMrio,rNat))), 1)
*      $ $endif.smallShare
*      )

*      $ $ifthen1.smallShare "%hasSmallShare%"=="on"
*      + sum(rNat1 $ (amwMrio(rNat1,iMrio,mrioA,rNat,ts) $ isSmallImpShare(rNat1,iMrio,rNat)),
*      p_amwMrio(rNat1,iMrio,mrioA,rNat,ts)
*      * ((1 + p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts) + mtax(rNat,iMrio,ts))
*      * %pmcif%(rNat1,iMrio,rNat,ts) ]/m_lambdam(rNat1,iMrio,rNat,ts))
*      $ $endif.smallShare
*      ;

```

In mathematical notation, without considering the special cases:

(51)

$$pmt_{r,i,mrioA} = (\sum_r amw_{r,r',i,mrioA} [pmcif_{r,r',i} (1 + imptx_{r,r',i,mrioA})]^{1-\sigma_{w,r,i}})^{1/1-\sigma_{w,r,i}}$$

The MRIO agents' import prices are defined in the equation  $e\_pmpMrio$ :

```
*
* --- normally not in model, substituted out by macro
*
e_pmpMrio(rs(r),iMrio(iMrio),aa,ts) $ (p_alpham(r,iMrio,aa,ts) $ iIn(iMrio)) ..
pmp(r,iMrio,aa,ts) =e= sum(mrio_A(mrioA,aa), pmtMrio(r,iMrio,mrioA,ts)) * (1 + mintx(r,iMrio,aa,ts) + itxshft(r,iMrio,ts));
```

The average tariff  $imptx$  for a product used, for instance, to calculate import tax revenues, is defined by the equation  $e\_imptxMrio$ :

```
*
* --- Average import tariff definition on each trade flow
* (enters, for instance, calculation of tax revenues)
*
e_imptxMrio(rNat1,iIn(iMrio),rsnat(rNat),ts) $ ( xwFlag(rNat1,iMrio,rNat)
$ (imptx.range(rNat1,iMrio,rNat,ts) ne 0) $ rrComb(rNat1,rNat) ) ..

(
  $$iftheni.fudg "%modulesArm_Les_Ces%"=="on"
    (xw(rNat1,iMrio,rNat,ts) - p_amwCom(rNat1,iMrio,rNat,ts)) $ (p_amwCom(rNat1,iMrio,rNat,ts) ge 0)
    + xwToFudge(rNat1,iMrio,rNat,ts) $ (p_amwCom(rNat1,iMrio,rNat,ts) lt 0)
  $$else.fudg
    xw(rNat1,iMrio,rNat,ts)
  $$endif.fudg
) * (imptx(rNat1,iMrio,rNat,ts) + mtax(rNat,iMrio,ts))
%scale% / (xw.scale(rNat1,iMrio,rNat,ts) * max(0.01,imptx.1(rNat1,iMrio,rNat,"%t0%")))
=e=

sum(mrioA $ p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts),
* --- taxes to weight
(p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts) + mtax(rNat,iMrio,ts))

* --- weight: simulated bi-lateral import demand by MRIO agent
* p_amwMrio(rNat1,iMrio,mrioA,rNat,ts) * ( xmtMrio(rNat,iMrio,mrioA,ts)
  $$iftheni.armLesCes "%modulesArm_Les_Ces%"=="on"
    $ (not arm_LES_CES(rnat,iMrio))
    + xmtUMrio(rNat,iMrio,mrioA,ts) $ arm_LES_CES(rnat,iMrio)
  $$endif.armLesCes
)

* [ (( pmtMrio(rNat,iMrio,mrioA,ts)
  $$iftheni.armLesCes "%modulesArm_Les_Ces%"=="on"
    $ (not arm_LES_CES(rnat,iMrio))
    + pmtUMrio(rnat,iMrio,mrioA,ts) $ arm_LES_CES(rnat,iMrio)
  $$endif.armLesCes
) / ((1 + p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts) + mtax(rNat,iMrio,ts))

*
* --- Note: cif price is defined via macro %pmcif%, not a variable in model
*
* %pmcif%(rNat1,iMrio,rNat,ts) ])*p_sigmax(rNat,iMrio)
* m_lambdam(rNat1,iMrio,rNat,ts) * (p_sigmax(rNat,iMrio) - 1)

$$iftheni.smallShare "%hasSmallShare%"=="on"
$ (not isSmallImpShare(rNat1,iMrio,rNat))
+ m_lambdam(rNat1,iMrio,rNat,ts) $ isSmallImpShare(rNat1,iMrio,rNat)
$$endif.smallShare

])
%scale% / (xw.scale(rNat1,iMrio,rNat,ts) * max(0.01,imptx.1(rNat1,iMrio,rNat,"%t0%")))
;
```

In mathematical notation without considering the case of commitment terms and a specific treatment of small shares:

(52)

$$xw_{r,r',i} imptx_{r,r',i} = \sum_{mrioA} imptx_{r,r',i,mrioA} amw_{r,r',i,mrioA} xmt_{r,i,mrioA} \left( \frac{pmt_{r,i,mrioA}}{pmcif_{r,r',i}(1+imptx_{r,r',i,mrioA})} \right)^{\sigma_{w,r,i}}$$



In case the version with commitment terms is used, there are three additional equations that are structurally identically to what is described for the module in the chapter before for the case with one agent in the lower level Armington nest. The first defines the value of total import for each MRIO agent and product without the value of the commitments:

```

---- define non-committed expenditure on imports for nation/product combination in Arm-Les-Ces functional form
e_nonComImpVMrio(rNat,iMrio,mrioA,ts) $ ( (pmtMrio.range(rNat,iMrio,mrioA,ts) ne 0) $ (not p_pmtElas(rNat,iMrio))
$ sum(rNat1 $ (p_amwMrio(rNat1,iMrio,mrioA,rNat,ts) $ xmtFlag(rNat,iMrio)
$ (not isSmallImpShare(rNat1,iMrio,rNat))),1) $ arm_LES_CES(rNat,iMrio) ) ..

nonComImpVMrio(rnat,iMrio,mrioA,ts)s(xmtMrio,'rnat,iMrio,mrioA,ts')
=E=

--- total import expenditures
xmtMrio(rNat,iMrio,mrioA,ts)s(xmtMrio,'rnat,iMrio,mrioA,ts') * pmtMrio(rNat,iMrio,mrioA,ts)

--- minues values of commitment at c.i.f. prices
- sum(rnat1 $ p_amwCom(rNat1,iMrio,rnat,ts),
p_amwCom(rNat1,iMrio,rnat,ts)*mrioSplitFactors(rNat,iMrio,rNat1,mrioA,"cif")
*(1 + p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts) + mtax(rNat,iMrio,ts))
* %pmcif%(rNat1,iMrio,rNat,ts))s(xmtMrio,'rnat,iMrio,mrioA,ts');

```

The second equation indirectly defines the utility aggregate *xmtUMrio* from exhaustion:

```

--- utility under Arm-Les-Ces from prices index and non-committed expenditure on imports
e_xmtUMrio(rNat,iMrio,mrioA,ts) $ ( (xmtUMrio.range(rNat,iMrio,mrioA,ts) ne 0) $ (not p_pmtElas(rNat,iMrio))
$ xmtFlag(rNat,iMrio)
$ sum(rNat1 $ (p_amwMrio(rNat1,iMrio,mrioA,rNat,ts)
$ (not isSmallImpShare(rNat1,iMrio,rNat))),1) $ arm_LES_CES(rNat,iMrio) ) ..

xmtUMrio(rNat,iMrio,mrioA,ts)s(xmt,'rnat,iMrio,ts') * pmtUMrio(rnat,iMrio,mrioA,ts)
=E= nonComImpVMrio(rnat,iMrio,mrioA,ts)s(xmt,'rnat,iMrio,ts');

```

And the last equation derives from the two variables defined above the exhaustion price index *pmtMrio*:

```

*
*
* --- this defines the exhaustion price pmt for the imports for the case with commitments
*
e_pmtUMrio(rNat,iMrio,mrioA,ts) $ ( (pmtUMrio.range(rNat,iMrio,mrioA,ts) ne 0) $ (not p_pmtElas(rNat,iMrio))
$ xmtFlag(rNat,iMrio)
$ sum(rNat1 $ (p_amwMrio(rNat1,iMrio,mrioA,rNat,ts)
$ (not isSmallImpShare(rNat1,iMrio,rNat))),1) $ arm_LES_CES(rNat,iMrio) ) ..

(pmtMrio(rnat,iMrio,mrioA,ts)-pmtUMrio(rnat,iMrio,mrioA,ts))
* xmtUMrio(rnat,iMrio,mrioA,ts)s(xmtMrio,'rnat,iMrio,mrioA,ts')
=E= sum(rnat1 $ p_amwCom(rNat1,iMrio,rnat,ts),
p_amwCom(rNat1,iMrio,rnat,ts)*mrioSplitFactors(rNat,iMrio,rNat1,mrioA,"cif")
*(1 + p_imptxMrio(rNat1,iMrio,mrioA,rNat,ts) + mtax(rNat,iMrio,ts))
* %pmcif%(rNat1,iMrio,rNat,ts))s(xmtMrio,'rnat,iMrio,mrioA,ts');

```

### 3.7.4 Graphical User Interface for MRIO extension

#### Data base generation

In order to generate the split factors when the data base is set-up, the following checkbox must be activated:

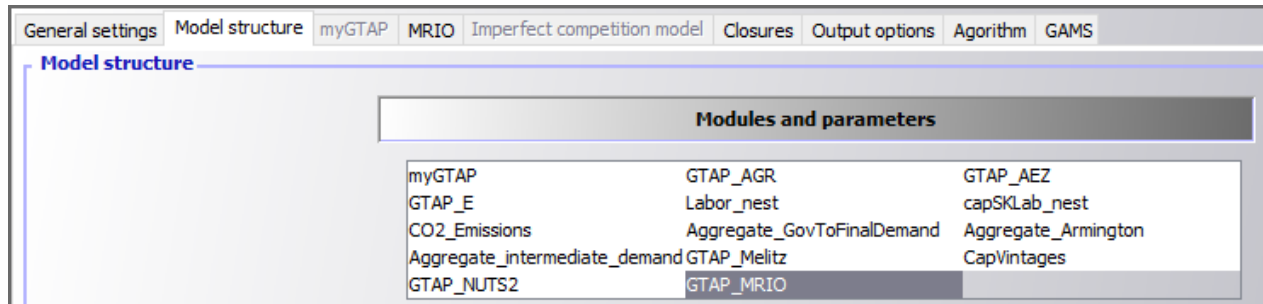
Generate MRIO split factors ☒

Fix import demand during split ☐

Furthermore, the user might activate the second check box to fix the import demands during the generation of the MRIO split factors.

### Simulation

When running the model, the MRIO module must be switched on:



In which case a tab allows selecting the products where the split is introduced:



The code will throw an error if the module is switch on but no MRIO shares are found in the data base.



### 3.8 Partial equilibrium closure

The partial equilibrium closure fixed prices and quantities of products not in the current solve. It also treats income as exogenous. Technically, that is implemented by three major approaches:

1. In “*model\closures.gms*”, the list of including products *iIn* and activities *aIn* which produce them is defined:

```
iIn(i) $ sum(sameas(i,pe_sel),1) = YES;
aIn(a) $ sum( (r,iIn), xFlag(r,a,iIn)) = YES;
```

Where *pe\_sel* are the user-selected products in the partial equilibrium version of the model.

2. Variables referring to products or activities not in the model are fixed, for instance:

```
xTNest.fx(r,tNest,a,tSim) $ (xTNest.l(r,tNest,a,tSim) $ (not aIn(a))) = xTNest.l(r,tNest,a,tSim);
pTNest.fx(r,tNest,a,tSim) $ (xTNest.l(r,tNest,a,tSim) $ (not aIn(a))) = pTNest.l(r,tNest,a,tSim);
va.fx(r,a,tSim) $ (va.l(r,a,tSim) $ (not aIn(a))) = va.l(r,a,tSim);
nd.fx(r,a,tSim) $ (nd.l(r,a,tSim) $ (not aIn(a))) = nd.l(r,a,tSim);
```

3. Equations defining these variables are not included in the model, for instance:

```
vaeq(rs(r),aIn(a),ts(t)) $ vaFlag(r,a) ..
```

These fixings need also to be reflected during the pre-solve (*solve\presolve.gms*) and during calibration and initialization of the different modules.

The user should note that shock files should not update tax rates etc. belonging to products not in the current model as some initialization statements might update for instance prices based on tax rates under the shock. Not all modules and options will work with the PE closures, such as for instance the MACC mechanism for endogenous process emissions.

### 3.9 Tariff Line Module

#### 3.9.1 Summary

The tariff line module allows to dis-aggregate selected bi-lateral trade flows to a lower detail and to explicitly model Tariff Rate Quotas. The user needs to provide the necessary data in a GAMS file in a specific format as detailed below, the file is selected from the GUI.

#### 3.9.2 Background

The GTAP data base offers bi-lateral protection and matching trade flow data based on MacMaps at the level of 57 respectively 65 GTAP sectors which are often further aggregated in model applications. That implies that information on protection measures at tariff line, which might also comprise complex instruments such as Tariff Rate Quotas (TRQs), needs to be aggregated to GTAP sector level and beyond. That might provoke aggregation bias (cf. Himics and Britz 2016<sup>33</sup>). Analysis especially of (potential) Free Trade Agreements requires often detail beyond the sectors of GTAP. The module for CGEBox described in here allows detail at tariff line for selected bi-lateral trade links and commodities while maintaining otherwise the chosen default trade structure for that sector (Armington or Armington plus CET, a combination with the Melitz module is currently not supported). If a MCP solver is available, multiple TRQs at bi-lateral tariff line can be introduced such that applied tariff rates become endogenous. At the same time, the structure provides a fully consistent aggregation from tariff line to commodity level.

The demand representation is based on a four-tier representation: (1) top level demand function such as the CDE in the GTAP standard model for final demanders and the FOC from cost minimization under CES technologies for intermediate demand, (2) where applicable, CES final or intermediate demand nests that ultimately define the total demand for single products by each Armington agent, (3) a first Armington nest that distributes the total demand by each Armington to domestic sales and aggregate imports, (4) Bi-lateral Armington nest, applied to the sum of the import demands over all Armington agents. The MRIO extension will use dedicated nests for each of the four MRIO agents in (4) and solely aggregate in case of intermediate demand.

A similar structure is found for supply transformation when the CET extensions is used: (1) first CET nest between domestic sales and aggregate exports, (2) CET nest between bi-lateral exports). Splitting up to tariff line level at the bi-lateral trade flow level as the lowest level of the demand system and supply transformation tree might seem odd at first look. One would perhaps rather like to split demand, for instance of dairy, first to tariff lines such as butter, cheese etc., next determine how much butter etc. is stemming from the home market and from imports, and finally how butter imports are composed from different importers. However, introducing the dis-aggregation and thus dis-aggregates nests with tariff line information directly below the top-level nests requires firstly introducing the tariff line split up for all bi-lateral trade flows of a region, and if a CET application is used, also for all its exporters and thus most probably for all regions in the model. That quickly renders the model quite large. Secondly, data on domestic sales at tariff line level must be available, data which are not generally available.

The proposed approach that applies the dis-aggregation at the lowest level of the demand system thus dramatically reduces data needs as it requires tariff line information on trade flows and protection data

---

<sup>33</sup> Himics, M., Britz, W. (2016): Flexible and welfare-consistent tariff aggregation over exporter regions, *Economic Modelling* 53: 375-387

only for some bi-lateral links, namely those in the focus of the analysis. Data on domestic sales at tariff lines are not needed at all. Furthermore, it keeps the model at a manageable size even if some bi-lateral trade flows are depicted with rich dis-aggregated tariff line information. That is especially important if TRQs are modelled which requires a MCP solver which typically is slower compared to solving a simple non-linear equation system.

### 3.9.3 Approach

The user has to provide the following information:

- The set of tariff lines considered and their mapping to the GTAP sectors in the data base used for the current model run
- Information on trade volumes for some bi-lateral trade flows at tariff line level. A bi-lateral trade flow for a commodity will only be depicted at tariff line level in the model if at least one non-zero flows for one tariff line is found. Otherwise, the bi-lateral commodity flow will continue to use the chosen default trade representation for that commodity without a dis-aggregation to tariff line detail.
- Matching applied tariff rate or TRQs with in- and out-of-quota tariff rates.

Based on that information, the code will add the necessary equations and variables for the selected trade links, fit the data to the given benchmark and introduce matching parameters.

Note again that Melitz extension cannot be combined with the tariff line module for a product.

### 3.9.4 Model equations

The code for the tariff line sub-module can be found in the sub-directory “*gams\tariffLines*”. The equations and core definition are found in “*tariffLines\_model.gms*”.

The code of the model comprises four equations:

- Two equations which define the bi-lateral exports and imports at tariff line level
- And two matching dual price aggregators which define the bi-lateral export and import prices at the level of GTAP commodities at bi-lateral level.

The bilateral export supply equation  $e\_xwstl$  based on CET transformation function is shown below. Note that finite transformation is required. The parameter  $p\_scaleFobX$  aligns the price level reported at tariff line with the average f.o.b. price.

```
*
* --- Bilateral export supply at tariff line level
*
e_xwstl(rNat,tl,rNat1,ts(t)) $ (p_gwtl(rNat,tl,rNat1,t) $ rrComb(rNat,rNat1) $ sum(tl_i(tl,Iin),1)) ..
    xwtl(rNat,tl,rNat1,t) $ xwtl,"rNat,tl,rNat1,t"
    =e= sum[tl_i(tl,i), p_gwtl(rNat,tl,rNat1,t) * xws(rNat,i,rNat1,t)/p_scaleFobX(rNat,i,rNat1) $ xwtl,"rNat,tl,rNat1,t"
        * (ptls(rNat,tl,rNat1,t) * p_scaleFobP(rNat,i,rNat1)/pe(rNat,i,rNat1,t)) ** p_omegawtl(rNat,i) ];
```

In mathematical notation:

$$\begin{cases} XWTL_{r,tl,d}^S = \gamma_{r,tl,d}^w XW_{r,i,d}^S \left( \frac{PETL_{r,tl,d}}{PE_{r,i,d}} \right)^{\omega_{r,tl}^w} & \text{if } \omega_{r,tl}^w \neq \infty \\ PETL_{r,tl,d} = PE_{r,i,d} & \text{if } \omega_{r,tl}^w = \infty \end{cases} \quad (1)$$

The parameter `p_wgtl` depicts the share of the tariff line `tl` in total bi-lateral trade of commodity `i` shipped from exporter `rNat` to importer `rNat1`. The equation is hence only introduced if a share parameter was defined during calibration, i.e. for such bi-lateral trade flows where information on that tariff line was provided.

Note that the variable `pe`, the bi-lateral export price, is part of the standard model when non-infinite CET is used while the price at the tariff line level `ptls` is not. Note equally that in the standard layout, bi-lateral export and import quantities at commodity level `xw` are by definition equal as they are no quality differences considered at that level without tariff line detail. That is no longer the case when the additional CES and CET nests at tariff line level are active such that the bi-lateral export quantity is now depicted by `xws` as a CET aggregate over multiple tariff lines. It is only identically to `xw` in the benchmark, in a simulation, the (implicit) non-linear aggregation based on a CES function on the demand side and a CET function on the supply side will let the aggregate supply and demand quantities deviate.

The average import price is defined as follows based on CET function in the equation `e_pe`:

```
e_petl(rNat,iIn(i),rNat1,ts(t)) $ (p_gw(rNat,i,rNat1,t) $ sum[tl_i(tl,i),p_gwtl(rNat,tl,rNat1,t)] $ rrComb(rNat,rNat1)) ..
pe(rNat,i,rNat1,t) =e= sum[tl_i(tl,i), p_gwtl(rNat,tl,rNat1,t) * (ptls(rNat,tl,rNat1,t)*p_scaleFobP(rNat,i,rNat1))**(1+p_omegawtl(rNat,i)))
**((1/(1+p_omegawtl(rNat,i))))];
```

Where `gw` is the share parameter in the second level CET nest.

In mathematical notation:

$$\begin{cases} PE_{r,i,d} = \left[ \sum_{tl \in i} \gamma_{r,tl,d}^w PETL_{r,tl,d}^{1+\omega_{r,tl}^w} \right]^{1/(1+\omega_{r,tl}^w)} & \text{if } \omega_{r,tl}^w \neq \infty \\ XW_{r,i,d}^s = \sum_{tl} XWTL_{r,tl,d}^s & \text{if } \omega_{r,tl}^w = \infty \end{cases} \quad (2)$$

In the original layout, the export price `pe` is implicitly defined by the equation `e_pe` which defines `xw` from total export supply. To keep that equation untouched, it is replaced here by a variant which is used if NO tariff line equation is given. Beside the related \$ condition, it is identical to the equation in the standard model:

```
* --- That is the original pe equation from ENVISAGE, only active if there no tariff line detail for that product
*
e_peMod(rNat,iIn(i),rNat1,ts(t)) $ (m_xwFlag(rNat,i,rNat1) $ psFlag(rNat,i,"pe") $ rrComb(rNat,rNat1) $ (not sum[tl_i(tl,i),p_gwtl(rNat,tl,rNat1,t)]) $ fullModel) ..
0 =e= ( xw(rNat,i,rNat1,t)s(xw,"rNat,i,rNat1,t") - p_gw(rNat,i,rNat1,t)*xet(rNat,i,t)s(xw,"rNat,i,rNat1,t")
* [1 $ isSmallExpShare(rNat,i,rNat1)
+ { (pe(rNat,i,rNat1,t)/pet(rNat,i,t))*p_omegaw(rNat,i) } $ (not isSmallExpShare(rNat,i,rNat1)) ] ) $ (p_omegaw(rNat,i) ne inf)
+ (pe(rNat,i,rNat1,t) - pet(rNat,i,t) $ (p_omegaw(rNat,i) ne inf) - ps(rNat,i,t) $ (p_omegaw(rNat,i) eq inf)) $ (p_omegaw(rNat,i) eq inf) );
```

The bi-lateral export supply at commodity level `xws` is defined identically, however neglects the case of small import shares or infinite transformation:

```
*
* -- same as above (infinite transformation case is however skipped, xw replaced by xws)
*
e_xws(rNat,iIn(i),rNat1,ts(t)) $ (m_xwFlag(rNat,i,rNat1) $ psFlag(rNat,i,"pe") $ rrComb(rNat,rNat1) $ sum[tl_i(tl,i),p_gwtl(rNat,tl,rNat1,t)]) ..
xws(rNat,i,rNat1,t)/xw.scale(rNat,i,rNat1,t)
=e= gw(rNat,i,rNat1,t)*xet(rNat,i,t)/xw.scale(rNat,i,rNat1,t)
* [1 $ isSmallExpShare(rNat,i,rNat1)
+ { (pe(rNat,i,rNat1,t)/pet(rNat,i,t))*p_omegaw(rNat,i) } $ (not isSmallExpShare(rNat,i,rNat1)) ];
```

The demand side comprises similar equations, first to define bi-lateral import demand at tariff line level `xwtl` (which by definition is equal to export supply):

```

*
* --- Bilateral import demand at tariff line level (= equal to bilateral export supply at tariff line level)
*
e_xwtl(rNat,tl,rNat1,ts(t)) $ (p_amwtl(rNat,tl,rNat1,t) $ rrComb(rNat1,rNat) $ sum(tl_i(tl,iin),1)) ..
xwtl(rNat,tl,rNat1,t)
s(xwtl,"rNat,tl,rNat1,t")
  =e= sum(tl_i(tl,i),
    p_amwtl(rNat,tl,rNat1,t)*xw(rNat,i,rNat1,t)/p_scaleFobX(rNat,i,rNat1)
    * ((%pmcif(rNat,i,rNat1,t)) * (1 + imptx(rNat,i,rNat1,t) + mtax(rNat1,i,t))) /
    ((m_pmCifTl(rNat,tl,rNat1,t))*p_scaleCifImptx(rNat,i,rNat1)*(1 + imptxTl(rNat,tl,rNat1,t) + mtax(rNat1,i,t))))**p_sigmtl(rNat1,i)
  )s(xwtl,"rNat,tl,rNat1,t");

```

In mathematical equation:

$$XWTL_{s,tl,r}^d = \alpha_{s,tl,r}^w XW_{s,i,r}^d \left( \frac{PM_{s,i,r}}{PMTL_{s,tl,r}} \right)^{\sigma_{r,tl}^w} \quad (3)$$

The macro `m_pmcifTl` defines the c.i.f. price, to which a tariff line specific ad-valorem tariff `imptxTL` is added as well as potentially a general shifter for import taxation of product `i`. Note that `imptxTL` is defined as a variable to render it endogenous under a TRQ. The macro is based itself on a macro to define the f.o.b. price:

```

$macro m_pefobTl(r,tl,rp,t) [ ptls(r,tl,rp,t) * (1 + exptx(r,i,rp,t) + etax(r,i,t)) ]

$macro m_pmcifTl(r,tl,rp,t) { [ m_pefobTl(r,tl,rp,t)/lcu(r,t) \
+ (sum(m $ amgm(m,r,i,rp), amgm(m,r,i,rp)*ptmg(m,t)/lambdamg(m,r,i,rp,t))*tmarg(r,i,rp,t)) \
]*lcu(rp,t) }

```

Both macros are identically structured as the macros in the standard model. Note that export taxes are currently not differentiated at tariff line level and that the trade margins are also assumed to be identical for all tariff lines belonging to the same commodity.

The average bi-lateral import price `pm` for a commodity `i`, normally substituted out in the standard model with the `m_pmcif` macro, is in here explicitly defined as an aggregation over the tariff-line prices:

```

e_pm(rNat,iIn(i),rNat1,ts(t)) $ (sum(tl_i(tl,i) $ p_amwtl(rNat,tl,rNat1,t),1) $ rrComb(rNat1,rNat)) ..
(%pmcif(rNat,i,rNat1,t))*(1 + imptx(rNat,i,rNat1,t) + mtax(rNat1,i,t)) =e=
sum(tl_i(tl,i) $ p_amwtl(rNat,tl,rNat1,t), p_amwtl(rNat,tl,rNat1,t)
* ((m_pmCifTl(rNat,tl,rNat1,t))*p_scaleCifImptx(rNat,i,rNat1)
* (1 + imptxTl(rNat,tl,rNat1,t) + mtax(rNat1,i,t)) )**(1-p_sigmtl(rNat1,i))))
** (1/(1-p_sigmtl(rNat1,i)));

```

In mathematical notation:

$$PM_{s,i,r} = \left[ \sum_{tl \in i} \alpha_{s,tl,r}^w PMTL_{s,tl,r}^{1-\sigma_{r,tl}^w} \right]^{1/(1-\sigma_{r,tl}^w)} \quad (4)$$

To keep the standard model simple, the modified xw equation `e_xwMod` which distinguishes between the case of dis-aggregation to tariff line or not is only comprised in the sub-module model:

```

*
* --- bilateral import demand at commodity level, note that the CEGBox default equation
* is the first block (not tariff line information) and uses pmCIF plus tariffs,
* whereas the second block (tariff line information) uses the pm variable define by the dual
* price aggregator above
*
e_xwMod(rNat1,iIn(i),rNat,ts(t)) $ (xwFlag(rNat1,i,rNat) $ rrComb(rNat1,rNat) $ (not iMrio(i))) ..
  xw(rNat1,i,rNat,t)/xw.scale(rNat1,i,rNat,t) =e= amw(rNat1,i,rNat,t)*xmt(rNat,i,t)/xw.scale(rNat1,i,rNat,t)

          *( [ (pmt(rNat,i,t)/[(1 + imptx(rNat1,i,rNat,t) + mtax(rNat,i,t))

*
* Note WB: cif price is defined via macro %pmCIF%, not a variable in model
*
          * %pmCIF(rNat1,i,rNat,t)/chipm(rNat1,i,rNat) ])**sigmaw(rNat,i)
          * lambdam(rNat1,i,rNat,t)**(sigmaw(rNat,i) - 1) $ (not isSmallImpShare(rNat1,i,rNat))

          + lambdam(rNat1,i,rNat,t) $ isSmallImpShare(rNat1,i,rNat)
          ] $ (not sum[t1_i(t1,i) $ p_amwtl(rNat1,t1,rNat,t),1])

          + [ (pmt(rNat,i,t)*chipm(rNat,i,rNat1)/pm(rNat1,i,rNat,t))**sigmaw(rNat,i)
          * lambdam(rNat1,i,rNat,t)**(sigmaw(rNat,i) - 1) $ (not isSmallImpShare(rNat1,i,rNat))

          + lambdam(rNat1,i,rNat,t) $ isSmallImpShare(rNat1,i,rNat)
          ] $ sum[t1_i(t1,i) $ p_amwtl(rNat1,t1,rNat,t),1]
          );

```

Finally, there is an equation for the TRQ mechanism:

```

*
* --- TRQ bound equation, paired with tariff
*
e_trq(rNat,t1,rNat1,ts(t)) $ (p_trq(rNat,t1,rNat1,t) $ rrComb(rNat1,rNat)) ..

p_trq(rNat,t1,rNat1,t) =E= xwtl(rNat,t1,rNat1,t);

```

The TRQ equation clearly requires that the related tariff at tariff line level `imptxTL` is endogenous and not fixed. With a MCP solver, bounds on that tariff variable model can depict the regime switch from under-filled quota (= tariff at lower bound) to binding quota (= tariff between lower and upper bound, the tariff variable is equal to the per unit quota rent plus the in-quota tariff) and over-quota imports (= tariff at upper bound). Note that introducing TRQs with a CNS solver will only work for the cases where the quota is binding both at the benchmark and in the simulation, i.e. the implicit tariff is between its lower and upper bounds in the final solution.

For a more expanded discussion, see Jafari, Y., Himics, M., Britz, W., Beckman, J. (2021): It is all in the details: A bilateral approach for modelling trade agreements at the tariff line, *Canadian Journal of Agricultural Economics* 69(3): 415-442 and Jafari, Y., Britz, W., Guimbard, H., Beckman, J. (2021): Properly capturing tariff rate quotas for trade policy analysis in computable general equilibrium models, *Economic Modelling* 104(105620).

### 3.9.5 Entering the data

The user provided data are stored in files in the directory “gams\scen\tariffLine”. A trial implementation to model EU->CAN exports at tariff line level can be found there as “ceta.gms”. It fits any data base with at least EU28, USA and ROW as labels for regions and “mil” and “wht” as commodities.

Note that the information provided by such a file to the module needs to be broken up in two blocks:

1. The **declarations**, i.e. the set of considered tariff lines and how they are mapped to the GTAP commodities in the current aggregation:

```

$iftheni.decl "%1"=="decl"

    set tl "Tariff line active in model"
        / cheese, butter,SMP,WMP,othDairy,
          baking_wht,feed_wht /;
    set tl_i(tl,i) "Link between tariff lines and GTAP commodities"
        / (cheese,butter,SMP,WMP,othDairy).mil-c
          (baking_wht,feed_wht).wht-c
        /;

```

Note that the set of tariff lines `tl` and their link to the commodity `tl_i` must be defined in the `decl` block in a `$ifthen-else-endif` structure as shown above.

## 2. The actual **data and parameter** to use

```

$else.decl
TABLE p_importTL (rp,tl,r) "Import Export Matrix of Cheese, importer first"
----- 19 line(s) not displayed
CAN.Baking_wht      0      5      47.97
EU_28.Baking_wht    0      0      24.46
ROW.Baking_wht      0.04    61.42    0
CAN.Feed_Wht        0      2      47.97
EU_28.Feed_Wht      0      0      24.46
ROW.Feed_Wht        0.04    61.42    0
;

*
* -- copy data to bi-lateral trade flow variable at tariff line,
*    will be used to calibrate parameters in tariff line model
*    and initialize further variables
*
xwtl.l(rNat,tl,rNat1,t0) = p_importTL(rNat1,tl,rNat);
*
* --- bi-lateral protection data
*
option kill=p_TRQ;
imptxTL.l(rNat,tl,rNat1,t0) $ p_importTL (rNat1,tl,rNat)
    = sum(tl_i(tl,i),imptx.l(rNat,i,rNat1,t0)) * uniform(0.5,1.5);
*
* --- set transformation and substitution elasticity at tariff line level
*
p_omegawtl(rnat,i) $ sum((tl_i(tl,i),rNat1), xwtl.l(rNat,tl,rNat1,"%t0%")) = 5;
p_sigmaxtl(rNat1,i) $ sum((tl_i(tl,i),rNat), xwtl.l(rNat,tl,rNat1,"%t0%")) = 5;

$endif.decl

```

Note that the data above are invented and the tariff randomly drawn to provide a test framework for the module. The data to use need to be stored in “*gams\scen\tariffline*” and can from there selected via the GUI (see below).

### 3.9.6 Calibration

The bi-lateral trade flow and protection data are *not* scaled to match the information in the global SAM. Instead, three scaling factor are defined that align (1) the bi-lateral trade volumes *p\_scaleFobX*, the (2) f.o.b. price level *p\_scaleFobP*, and (3) the tariffs to match the reported totals at product level as found in the model data base. This solution allows to define TRQs in “natural” definitions and to use specific tariffs as used in the extension with the endogenous quota allocation mechanism

```

*
* --- aggregate from tariff line to total bi-lateral exports
*
p_scaleFobX(rNat,i,rNat1) = sum(tl_i(tl,i),xwtl.1(rNat,tl,rNat1,"%t0%"));
p_scaleFobX(rNat,i,rNat1) $ p_scaleFobX(rNat,i,rNat1) = xw.1(rnat,i,rNat1,"%t0%")/p_scaleFobX(rnat,i,rnat1);

p_scaleFobP(rNat,i,rNat1) = sum(tl_i(tl,i),xwtl.1(rNat,tl,rNat1,"%t0%")*ptls.1(rNat,tl,rNat1,"%t0%")*p_scaleFobX(rnat,i,rnat1));
p_scaleFobP(rNat,i,rNat1) $ p_scaleFobP(rNat,i,rNat1) = xw.1(rnat,i,rNat1,"%t0%")*pe.1(rNat,i,rNat1,"%t0%")/p_scaleFobP(rnat,i,rnat1);
*
* --- scale bi-lateral tariffs at tariff line to exhaust given tariffs at commodity level
*
p_scaleCifImptx(rNat,i,rNat1) = 1;
p_scaleCifImptx(rNat,i,rNat1) $ p_scaleFobX(rnat,i,rnat1)
= xw.1(rNat,i,rNat1,"%t0%")*p_mcif(rNat,i,rNat1,"%t0%")*(1+imptx(rNat,i,rNat1,"%t0%"))
/sum(tl_i(tl,i), xwtl.1(rNat,tl,rNat1,"%t0%")*p_scaleFobX(rnat,i,rnat1)*(m_pmCifTl(rNat,tl,rNat1,"%t0%"))*(1+imptxTl.1(rNat,tl,rNat1,"%t0%"))));

```

Share parameters can be derived the user-supplied price and quantities at tariff line level, considering the relevant scaling factors:

```

*
* --- calculate share parameter for CET: prices are equal,
* simply quantity share
*
p_gwtl(rNat,tl,rNat1,t0) $ xwtl.1(rNat,tl,rNat1,t0)
= xwtl.1(rNat,tl,rNat1,t0)/sum(tl_i(tl,i),xws.1(rnat,i,rNat1,t0)/p_scaleFobX(rnat,i,rnat1))
*sum(tl_i(tl,i),(pe.1(rNat,i,rNat1,t0)/(ptls.1(rNat,tl,rNat1,t0)*p_scaleFobP(rnat,i,rnat1)))*p_omegawtl(rnat,i));
*
* --- bilateral import prices are tariff levels are not equal,
* price relation needed for calibration of share parameter
*
p_amwtl(rNat,tl,rNat1,t0) $ xwtl.1(rNat,tl,rNat1,t0)
= sum(tl_i(tl,i), xwtl.1(rNat,tl,rNat1,t0)*p_scaleFobX(rnat,i,rnat1)/xw.1(rnat,i,rNat1,t0)

* ((m_pmCifTl(rNat,tl,rNat1,t0))*(1 + imptxTl(rNat,tl,rNat1,t0))*p_scaleCifImptx(rNat,i,rNat1)
/({p_mcif(rNat,i,rNat1,"%t0%")*(1+imptx(rNat,i,rNat1,"%t0%"))}))*p_sigmaxtl(rNat1,i));

```

### 3.9.7 Post model processing

The code comprising the model equations also comprises a section for post-model processing:

```

p_results(rNat,"Q",tl,rNat1,"dom",%arg1%,"%version%") $ xwtl.1(rNat,tl,rNat1,%arg2%) = xwtl.1(rNat,tl,rNat1,%arg2%) * %scale%;
p_results(rNat,"P",tl,rNat1,"dom",%arg1%,"%version%") $ xwtl.1(rNat,tl,rNat1,%arg2%) = ptls.1(rNat,tl,rNat1,%arg2%);
p_results(rNat,"T",tl,rNat1,"dom",%arg1%,"%version%") $ xwtl.1(rNat,tl,rNat1,%arg2%) = sum(tl_i(tl,i),exptx.1(rNat,i,rNat1,%arg2%) + etax.1(rNat,i,%arg2%));

p_results(rNat,"TRQ",tl,rNat1,"imp",%arg1%,"%version%") $ p_trq(rNat1,tl,rNat,%arg2%) = p_trq(rNat1,tl,rNat,%arg2%);
p_results(rNat,"LowTar",tl,rNat1,"imp",%arg1%,"%version%") $ p_trq(rNat1,tl,rNat,%arg2%) = imptxTl.lo(rNat1,tl,rNat,%arg2%) + eps;
p_results(rNat,"UppTar",tl,rNat1,"imp",%arg1%,"%version%") $ p_trq(rNat1,tl,rNat,%arg2%) = imptxTl.up(rNat1,tl,rNat,%arg2%);

p_results(rNat,"Q",tl,rNat1,"imp",%arg1%,"%version%") $ p_amwtl(rNat1,tl,rNat,%arg2%) = xwtl.1(rNat1,tl,rNat,%arg2%) * %scale%;
p_results(rNat,"P",tl,rNat1,"imp",%arg1%,"%version%") $ p_amwtl(rNat1,tl,rNat,%arg2%)
= sum(tl_i(tl,i),m_pmCifTl(rNat1,tl,rNat,%arg2%)*(1 + imptxTl(rNat1,tl,rNat,%arg2%) + mtax(rNat,i,%arg2%)));

p_results(rNat,"T",tl,rNat1,"imp",%arg1%,"%version%") $ xwtl.1(rNat1,tl,rNat,%arg2%)
= imptxTl(rNat1,tl,rNat,%arg2%) + sum(tl_i(tl,i),mtax(rNat,i,%arg2%));

p_results(rNat,"V",tl,rNat1,domImp,%arg1%,"%version%") $ p_results(rNat,"Q",tl,rNat1,"dom",%arg1%,"%version%")
= p_results(rNat,"Q",tl,rNat1,domImp,%arg1%,"%version%") * p_results(rNat,"P",tl,rNat1,"dom",%arg1%,"%version%");

p_results(rNat,"G",tl,rNat1,domImp,%arg1%,"%version%") $ p_results(rNat,"Q",tl,rNat1,domImp,%arg1%,"%version%")
= p_results(rNat,"Q",tl,rNat1,domImp,%arg1%,"%version%") * p_results(rNat,"T",tl,rNat1,domImp,%arg1%,"%version%");

p_results(rNat,"VG",tl,rNat1,domImp,%arg1%,"%version%") $ p_results(rNat,"Q",tl,rNat1,domImp,%arg1%,"%version%")
= p_results(rNat,"V",tl,rNat1,domImp,%arg1%,"%version%") + p_results(rNat,"G",tl,rNat1,domImp,%arg1%,"%version%");

p_results(rNat,"PT",tl,rNat1,domImp,%arg1%,"%version%") $ p_results(rNat,"VG",tl,rNat1,domImp,%arg1%,"%version%")
= p_results(rNat,"VG",tl,rNat1,domImp,%arg1%,"%version%") / p_results(rNat,"Q",tl,rNat1,domImp,%arg1%,"%version%");

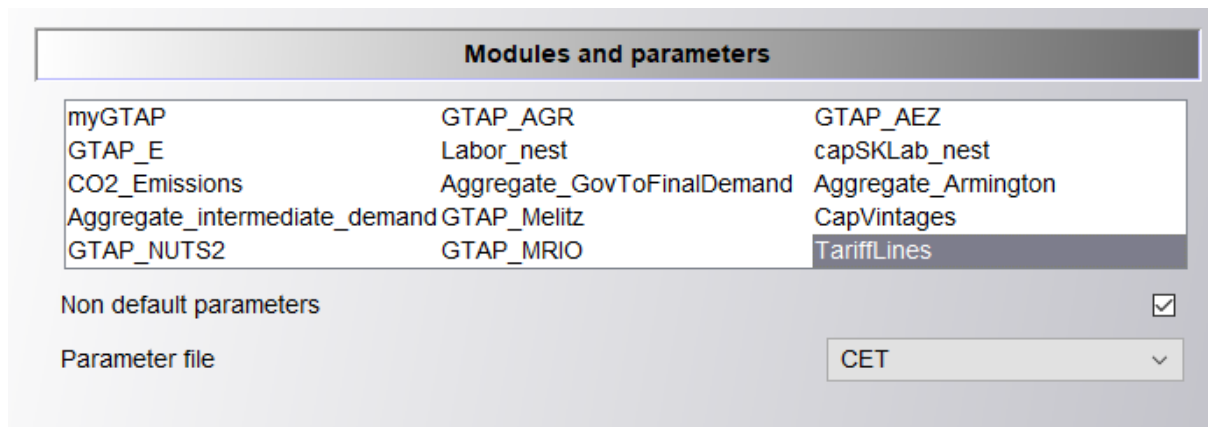
```



### 3.9.8 Integration in the GUI

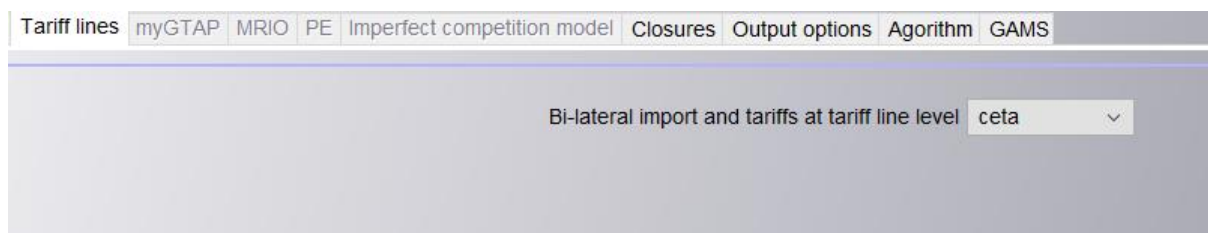
#### Model configuration

The module can as usually be switched on under the “Model structure” tab:



Note again that a non-default parameter file must be chosen which introduces transformation elasticities. It is not possible to use the model with an infinite transformation elasticity between export destinations.

The file with the set of considered tariff lines and related data can be chosen on the “Tariff lines” tab:



#### Result exploitation

The results at tariff line level are comprised in the two tables “Exports, trade matrix, Tariff lines” and “Imports, trade matrix, Tariff lines”:

|               |   |                                          |
|---------------|---|------------------------------------------|
| Trade         | > | Exports, trade matrix                    |
| CO2           | > | Exports, trade matrix, Tariff lines      |
| No table      |   | Exports, trade matrix flow map, detailed |
| Security      |   | Exports, by product and destination      |
| Price (net of |   | Imports, trade matrix                    |
| Value (net of |   | Imports, trade matrix, Tariff lines      |

A screen shot of a part of such a table is shown below:

## Tariff Line Module

Imports, trade matrix, Tariff lines [0]

Importer: Canada Years: shock

|        |                      | ceta  |     |
|--------|----------------------|-------|-----|
|        |                      | EU_28 | Row |
| cheese | Quantity             | 0.19  |     |
|        | Price (net of taxes) | 1.05  |     |
|        | Value (net of taxes) | 0.21  |     |
|        | Tax rate             | 0.27  |     |
|        | Tax income           | 0.05  |     |
|        | Price, tax inclusive | 1.39  |     |
|        | Value, tax inclusive | 0.27  |     |
|        | TRQ                  | 0.19  |     |
|        | LowTar               | 0.00  |     |
|        | UppTar               | 1.44  |     |
| butter | Quantity             | 0.01  |     |
|        | Price (net of taxes) | 0.89  |     |
|        | Value (net of taxes) |       |     |
|        | Tax rate             | 1.91  |     |
|        | Tax income           | 0.02  |     |
|        | Price, tax inclusive | 1.91  |     |
|        | Value, tax inclusive | 0.02  |     |

### 3.10 CO<sub>2</sub>, Non-CO<sub>2</sub> Accounting and air pollution emissions, energy use accounting

The CO<sub>2</sub>, Non-CO<sub>2</sub>, air pollution and energy use data are by now a default part of newer GTAP data base releases. They provide first additional results which can be exploited with the GUI. Second, the CO<sub>2</sub> and Non-CO<sub>2</sub> emissions can be used to define total CO<sub>2</sub> equivalent emissions which can be taxed with a fixed carbon tax or an endogenous one under an emission ceiling.

Parameters and indices | AEZ | NUTS2 | Tariff lines | myGTAP | CO2 emissions | MRIO | ARM\_LES\_CES | PE | G-RDEM | Imperfect competition

| Modules and parameters     |                     |                       |             |
|----------------------------|---------------------|-----------------------|-------------|
| myGTAP                     | GTAP_AGR            | GTAP_AEZ              | GTAP_E      |
| CO2_Emissions              | Non_CO2_Emissions   | Air_Pollution         | MTOE        |
| Aggregate_GovToFinalDemand | Aggregate_Armington | Aggregate_firm_demand | GTAP_Melitz |
| ARM_LES_CES                | SPE                 | GTAP_MRIO             | TariffLines |
| CapVintages                | GTAP_NUTS2          | Parameters            | Nestings    |
| G-RDEM                     | GMIG                |                       |             |

#### 3.10.1 GHG emissions in CO<sub>2</sub> equivalents

The CO<sub>2</sub> accounting defines reported emissions in the model based on the following equations (see *gams/co2eq/co2eq\_model.gms*):

```
*
* --- emissions at regional level
*
e_emisWOCarbon(rsNat,ts(tRun)) ..
    emisWOCarbon(rsNat,tRun) s(emisWOCarbon,'rsNat,tRun')
    =E= [ sum( (i,aa) $ (m_alphad(rsNat,i,aa,tRun) $ emidCo2eq(rsNat,i,aa,tRun)),
        emidCo2eq(rsNat,i,aa,tRun) * (1-v_redEmiaCo2eq(rsNat,aa,tRun) * p_nCo2ShareD(rsNat,i,aa) + p_emiShift(rsNat,i,aa,tRun))
        * xd(rsNat,i,aa,tRun))
        + sum( (i,aa) $ (m_alpham(rsNat,i,aa,tRun) $ emiiCo2eq(rsNat,i,aa,tRun)),
        emiiCo2eq(rsNat,i,aa,tRun) * (1-v_redEmiaCo2eq(rsNat,aa,tRun) * p_nCo2ShareI(rsNat,i,aa) + p_emiShift(rsNat,i,aa,tRun))
        * xm(rsNat,i,aa,tRun))
        + sum( a $ (xpFlag(rsNat,a) $ emiaCo2eq(rsNat,a,tRun)),
        emiaCo2eq(rsNat,a,tRun) * (1-v_redEmiaCo2eq(rsNat,a,tRun) * p_emiShift(rsNat,"",a,tRun))
        * m_xp(rsNat,a,tRun))
        + sum( (f,a) $ (xfFlag(rsNat,f,a) $ emifCo2eq(rsNat,f,a,tRun)),
        emifCo2eq(rsNat,f,a,tRun) * (1-v_redEmiaCo2eq(rsNat,a,tRun) * p_emiShift(rsNat,f,a,tRun))
        * xf(rsNat,f,a,tRun))
    ] s(emisWOCarbon,'rsNat,tRun');
```

The equations above comprise four vectors of emission factors defined per unit of USD and not on physical basis, all expressed on CO<sub>2</sub> equivalents:

1. *emidCO2eq* are emission factors per unit of domestic demand of product *i* and Armington agent *aa*.
2. *emiiCO2eq* are emission factors per unit of imported demand of product *i* and Armington agent *aa*.
3. *emiaCo2eq* are emission factors per unit of output.
4. *emifCo2eq* are emission factors per unit of factor use.

The emission factors *emidCO2eq* and *emiiCO2eq* comprise both emissions from combustion of fossil fuels as various process emissions. The share of process emission is stored on the parameters *p\_nCo2ShareD* and *p\_nCo2ShareI*, respectively. The model can use a mechanism where all process emissions of an activity *a* are subject to a uniform endogenous relative reduction factor *v\_redEmiaCo2eq* depending on the carbon price and points on Marginal Abatement Cost Curves (MACCs).

To these emissions, CO2 changes from carbon stock changes can be added as calculated by the AEZ module:

```

--- emissions at regional level
e_emis(rsNat,ts(tRun)) ..
emis(rsNat,tRun) s(emis,'rsNat,tRun') =E= [emisWOCarbon(rsNat,tRun)
$ $if1 "%modulesGTAP_AEZ%"=="on" - dCarbon(rsNat,tRun)*0.001*3.67 $ ((not sameas(tRun,"%t0%"))
$ sum((r_r(rs,rsNat),aez,land_Cover) $ aezFlag(rs,aez),p_carbon(rsNat,aez,land_cover,"carbon"))) $ hasGtapAez
]
s(emis,'rsNat,tRun');

```

The reported emissions can differ from the taxed ones, see equation  $e\_emisTaxedWoCarbon$ . This allows to define carbon taxation scheme that cover not all emissions or emitters.

```

e_emisTaxedWoCarbon(rsNat,ts(tRun)) ..
emisTaxedWoCarbon(rsNat,tRun) s(emisTaxedWoCarbon,'rsNat,tRun')
=E= [ sum( (i,aa) $ (m_alphad(rsNat,i,aa,tRun) $ emidCo2eqTaxed(rsNat,i,aa,tRun)),
emidCo2eqTaxed(rsNat,i,aa,tRun) * (1-v_redEmiaCo2eq(rsNat,aa,tRun) * p_nCo2ShareD(rsNat,i,aa) + p_emisShift(rsNat,i,aa,tRun))
* xd(rsNat,i,aa,tRun))
+ sum( (i,aa) $ (m_alpham(rsNat,i,aa,tRun) $ emiiCo2eqTaxed(rsNat,i,aa,tRun)),
emiiCo2eqTaxed(rsNat,i,aa,tRun) * (1-v_redEmiaCo2eq(rsNat,aa,tRun) * p_nCo2ShareI(rsNat,i,aa) + p_emisShift(rsNat,i,aa,tRun))
* xm(rsNat,i,aa,tRun))
+ sum( a $ (xpFlag(rsNat,a) $ emiaCo2eqTaxed(rsNat,a,tRun)),
emiaCo2eqTaxed(rsNat,a,tRun) * (1-v_redEmiaCo2eq(rsNat,a,tRun) + p_emisShift(rsNat,"",a,tRun))
* (m_xp(rsNat,a,tRun) $ (not disr(rsNat))
$ $if1 "%modulesGTAP_NUTS2%"=="ON" + sum(r_r(subr,rsNat) $ xpFlag(subr,a), m_xp(subr,a,ts))
))
+ sum( (f,a) $ (xfFlag(rsNat,f,a) $ emifCo2eqTaxed(rsNat,f,a,tRun)),
emifCo2eqTaxed(rsNat,f,a,tRun) * (1-v_redEmiaCo2eq(rsNat,a,tRun) + p_emisShift(rsNat,f,a,tRun))
* xf(rsNat,f,a,tRun))
] s(emisTaxedWoCarbon,'rsNat,tRun');

```

Note that carbon stock changes can be taxed, only, if the GTAP-AEZ module is active and the attached emission price is not fixed to zero.

The definition which emissions are taxed is determined by the selection of the user on the tab “CO2 emissions”:

| id                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | indices   | AEZ       | Sub-regions | Tariff lines | myGTAP | CCDR        | CO2 emissions | MRIO    | ARM_LES_CES | PE | G-RDEM | CC Damages | Imperfect competition model | SPE | C |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|-----------|-------------|--------------|--------|-------------|---------------|---------|-------------|----|--------|------------|-----------------------------|-----|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--|--|--|--|--|--|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------------|---------|---------|----------|----------|-----------|-----------|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--|--|--|--|--|--|--|
| <div> <div>CO2 Taxed commodities</div> <table border="1"> <tr><td>wht-c</td><td>gro-c</td><td>pdr-c</td><td>v_f-c</td><td>osd-c</td><td>c_b-c</td><td>pfb-c</td><td>ocr-c</td><td>ctl-c</td></tr> <tr><td>oap-c</td><td>rmk-c</td><td>wol-c</td><td>frs-c</td><td>fsh-c</td><td>coa-c</td><td>oil-c</td><td>gas-c</td><td>oxl-c</td></tr> <tr><td>cmt-c</td><td>omt-c</td><td>vol-c</td><td>mil-c</td><td>pcr-c</td><td>sgr-c</td><td>ofd-c</td><td>b_t-c</td><td>tex-c</td></tr> <tr><td>wap-c</td><td>lea-c</td><td>lum-c</td><td>ppp-c</td><td>p_c-c</td><td>chm-c</td><td>bph-c</td><td>rpp-c</td><td>nmn-c</td></tr> <tr><td>i_s-c</td><td>nfm-c</td><td>fmp-c</td><td>ele-c</td><td>eeq-c</td><td>ome-c</td><td>mvh-c</td><td>otn-c</td><td>omf-c</td></tr> <tr><td>ely-c</td><td>TnD-c</td><td>gdt-c</td><td>wtr-c</td><td>cns-c</td><td>trd-c</td><td>afs-c</td><td>otp-c</td><td>wtp-c</td></tr> <tr><td>atp-c</td><td>whs-c</td><td>cmn-c</td><td>ofi-c</td><td>ins-c</td><td>rsa-c</td><td>obs-c</td><td>ros-c</td><td>osg-c</td></tr> <tr><td>edu-c</td><td>hht-c</td><td>dwe-c</td><td></td><td></td><td></td><td></td><td></td><td></td></tr> </table> </div> <div> <div>CO2 Taxed activities</div> <table border="1"> <tr><td>wht-a</td><td>gro-a</td><td>pdr-a</td><td>v_f-a</td><td>osd-a</td><td>c_b-a</td><td>pfb-a</td><td>ocr-a</td><td>ctl-a</td><td>oap-a</td></tr> <tr><td>rmk-a</td><td>wol-a</td><td>frs-a</td><td>fsh-a</td><td>coa-a</td><td>oil-a</td><td>gas-a</td><td>oxl-a</td><td>cmt-a</td><td>omt-a</td></tr> <tr><td>vol-a</td><td>mil-a</td><td>pcr-a</td><td>sgr-a</td><td>ofd-a</td><td>b_t-a</td><td>tex-a</td><td>wap-a</td><td>lea-a</td><td>lum-a</td></tr> <tr><td>ppp-a</td><td>p_c-a</td><td>chm-a</td><td>bph-a</td><td>rpp-a</td><td>nmn-a</td><td>i_s-a</td><td>nfm-a</td><td>fmp-a</td><td>ele-a</td></tr> <tr><td>eeq-a</td><td>ome-a</td><td>mvh-a</td><td>otn-a</td><td>omf-a</td><td>TnD-a</td><td>NuclearBL-a</td><td>GASBL-a</td><td>OILBL-a</td><td>WindBL-a</td></tr> <tr><td>CoalBL-a</td><td>HydroBL-a</td><td>OtherBL-a</td><td>SolarP-a</td><td>gdt-a</td><td>wtr-a</td><td>cns-a</td><td>trd-a</td><td>afs-a</td><td>otp-a</td></tr> <tr><td>wtp-a</td><td>atp-a</td><td>whs-a</td><td>cmn-a</td><td>ofi-a</td><td>ins-a</td><td>rsa-a</td><td>obs-a</td><td>ros-a</td><td>osg-a</td></tr> <tr><td>edu-a</td><td>hht-a</td><td>dwe-a</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr> </table> </div> <div> <div>Include Non-Co2 emissions in taxation</div> <input checked="" type="checkbox"/> </div> <div> <div>Include emissions from final consumption in taxation</div> <input checked="" type="checkbox"/> </div> <div> <div>MACCS for process emissions</div> <input checked="" type="checkbox"/> </div> <div> <div>Share of abatement cost collected as production taxes</div> <input type="text" value="1.0"/> </div> <div> <div>Maximum abatement option to consider</div> <input type="text" value="ep5000"/> </div> <div> <div>File with MACC information</div> <input type="text" value="maccs_dummy"/> </div> |           |           |             |              |        |             |               |         |             |    |        |            |                             |     |   | wht-c | gro-c | pdr-c | v_f-c | osd-c | c_b-c | pfb-c | ocr-c | ctl-c | oap-c | rmk-c | wol-c | frs-c | fsh-c | coa-c | oil-c | gas-c | oxl-c | cmt-c | omt-c | vol-c | mil-c | pcr-c | sgr-c | ofd-c | b_t-c | tex-c | wap-c | lea-c | lum-c | ppp-c | p_c-c | chm-c | bph-c | rpp-c | nmn-c | i_s-c | nfm-c | fmp-c | ele-c | eeq-c | ome-c | mvh-c | otn-c | omf-c | ely-c | TnD-c | gdt-c | wtr-c | cns-c | trd-c | afs-c | otp-c | wtp-c | atp-c | whs-c | cmn-c | ofi-c | ins-c | rsa-c | obs-c | ros-c | osg-c | edu-c | hht-c | dwe-c |  |  |  |  |  |  | wht-a | gro-a | pdr-a | v_f-a | osd-a | c_b-a | pfb-a | ocr-a | ctl-a | oap-a | rmk-a | wol-a | frs-a | fsh-a | coa-a | oil-a | gas-a | oxl-a | cmt-a | omt-a | vol-a | mil-a | pcr-a | sgr-a | ofd-a | b_t-a | tex-a | wap-a | lea-a | lum-a | ppp-a | p_c-a | chm-a | bph-a | rpp-a | nmn-a | i_s-a | nfm-a | fmp-a | ele-a | eeq-a | ome-a | mvh-a | otn-a | omf-a | TnD-a | NuclearBL-a | GASBL-a | OILBL-a | WindBL-a | CoalBL-a | HydroBL-a | OtherBL-a | SolarP-a | gdt-a | wtr-a | cns-a | trd-a | afs-a | otp-a | wtp-a | atp-a | whs-a | cmn-a | ofi-a | ins-a | rsa-a | obs-a | ros-a | osg-a | edu-a | hht-a | dwe-a |  |  |  |  |  |  |  |
| wht-c                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | gro-c     | pdr-c     | v_f-c       | osd-c        | c_b-c  | pfb-c       | ocr-c         | ctl-c   |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| oap-c                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | rmk-c     | wol-c     | frs-c       | fsh-c        | coa-c  | oil-c       | gas-c         | oxl-c   |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| cmt-c                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | omt-c     | vol-c     | mil-c       | pcr-c        | sgr-c  | ofd-c       | b_t-c         | tex-c   |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| wap-c                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | lea-c     | lum-c     | ppp-c       | p_c-c        | chm-c  | bph-c       | rpp-c         | nmn-c   |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| i_s-c                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | nfm-c     | fmp-c     | ele-c       | eeq-c        | ome-c  | mvh-c       | otn-c         | omf-c   |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| ely-c                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | TnD-c     | gdt-c     | wtr-c       | cns-c        | trd-c  | afs-c       | otp-c         | wtp-c   |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| atp-c                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | whs-c     | cmn-c     | ofi-c       | ins-c        | rsa-c  | obs-c       | ros-c         | osg-c   |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| edu-c                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | hht-c     | dwe-c     |             |              |        |             |               |         |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| wht-a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | gro-a     | pdr-a     | v_f-a       | osd-a        | c_b-a  | pfb-a       | ocr-a         | ctl-a   | oap-a       |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| rmk-a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | wol-a     | frs-a     | fsh-a       | coa-a        | oil-a  | gas-a       | oxl-a         | cmt-a   | omt-a       |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| vol-a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | mil-a     | pcr-a     | sgr-a       | ofd-a        | b_t-a  | tex-a       | wap-a         | lea-a   | lum-a       |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| ppp-a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | p_c-a     | chm-a     | bph-a       | rpp-a        | nmn-a  | i_s-a       | nfm-a         | fmp-a   | ele-a       |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| eeq-a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | ome-a     | mvh-a     | otn-a       | omf-a        | TnD-a  | NuclearBL-a | GASBL-a       | OILBL-a | WindBL-a    |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| CoalBL-a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | HydroBL-a | OtherBL-a | SolarP-a    | gdt-a        | wtr-a  | cns-a       | trd-a         | afs-a   | otp-a       |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| wtp-a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | atp-a     | whs-a     | cmn-a       | ofi-a        | ins-a  | rsa-a       | obs-a         | ros-a   | osg-a       |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |
| edu-a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | hht-a     | dwe-a     |             |              |        |             |               |         |             |    |        |            |                             |     |   |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |             |         |         |          |          |           |           |          |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |  |  |  |  |  |  |  |

Emissions of groups of model regions can be summed up under common emission ceilings. These groups are stored on the set *ets*, the cross-set *ets\_rnat(ets,rNat)* indicates which model regions *rNat* are part of a common emission trading scheme *ets*:

```

--- trade in certificates must add up to zero, defines common price in ETS

e_emisEtsClear(ets,ts) $ ((card(rsNat) gt 1) $ sum(ets_rnat(ets,rnat),p_permitCo2Eq(rNat,ts))) ..
    sum(ets_rnat(ets,rnat), emisEtsTrd(rNat,ts))/sum(ets_rnat(ets,rnat),p_permitCo2Eq(rNat,ts)) =E= 0;

```

The emission prices *emisP* are either fixed exogenously or are defined under an emission ceiling in the equation *e\_emisP*:

```

*
*   --- regional emission ceiling equations, paired to emission price (=tax)
*
e_emisP(rsNat,ts) $ ( [(emisP.range(rsNat,ts) ne 0) $ (emisTaxed.range(rsNat,ts) ne 0)] ) ..
    0 =E=
*
*   --- case with regional ceiling
*
    + ( (emisTaxed(rsNat,ts) - p_permitCo2Eq(rsNat,ts))/p_permitCo2Eq(rsNat,ts) )
        $ p_permitCo2Eq(rsNat,ts)
*
*   --- neither, price is fixed or zero
*
    + emisP(rsNat,ts) $ ( (not p_permitCo2Eq(rsNat,ts)) );

```

The taxed emissions are defined from the taxed emission without carbon stock changes *emisTaxWoCarbon* as defined in the equation above, plus taxed carbon changes, plus sold permits *emisEtsTrd* under an emission trading scheme covering multiple model regions:

```

--- emissions at regional level

e_emisTaxed(rsNat,ts(tRun)) ..

emisTaxed(rsNat,tRun)s(emisTaxed,'rsNat,tRun')

=E= [ emisTaxedWoCarbon(rsNat,tRun)

    $if1 "%modulesGTAP_AEZ%"=="on"      - (dCarbon(rsNat,tRun)-dCarbonBuiltup(rsNat,tRun)) * 3.67*0.001
    $ ((not sameas(tRun,"%t0%")) $ sum(a $ ((emisPLnd.range(rsNat,a,tRun) ne 0) $ xfFlag(rsNat,lnd,a),1))

    + sum(ets_rnat(ets,rsnat) $ p_permitCo2Eq(rsNat,ts), emisEtsTrd(rsNat,ts))
    + sum(ets_rnat(ets,rsnat) $ ((not p_permitCo2Eq(rsNat,ts)) $ (emisEtsTrd.1(rsNat,ts) ne 0)),emisEtsTrd.1(rsNat,ts))

    ]s(emisTaxed,'rsNat,tRun') ;

```

Examples to define emission taxes or caps are found in the directory “*scen/gams/scen/GHG*”.

The changes in taxed carbon changes are calculated in the following equation, using the carbon stock changes defined in the AEZ model:

```

--- only losses are taxed

e_dCarbonTaxedN(rsNat(rNat),ts(tRun)) $ ((not sameas(ts,"%t0%")) ) ..

dCarbonTaxedN(rNat,tRun)s(dCarbon,"rNat,tRun")
    =E= -ncpVUSin(dCarbon(rnat,ts)- dCarbonBuiltup(rNat,ts) ,0,1.E-5)s(dCarbon,"rNat,tRun-1");

```

The emission price for each land category (cropland, pasture, forest) is defined from the economy wide emission price *emisP* converted from CO<sub>2</sub> into carbon, times the share of the land category in the *carbonTaxBase* divided by the value of the factor demand for the land category *xfLandCat*. This value is calculated from the value of the nests in the AEZ model linked to the land category:

```

e_xfLandCat(rNat(rNat), landCat, ts(tRun)) $ ( (not sameas(tRun, "%t0%"))
$ sum((r_r(rs, rNat), aez) $ p_carbon(rNat, aez, landCat, "carbon"), p_landUse(rs, aez, landCat, "ha"))
$ sum(a $ xfFlag(rNat, lnd, a), (emispLnd.range(rNat, a, tRun) ne 0)) ..

xfLandCat(rNat, landCat, ts) / sum((r_r(rs, rNat), aez, aezNest) $ (landCat_aezNest(landCat, aezNest)
$ ( sum(aezNest_a(aezNest, aezA), p_gAez(rs, aez, aezA, ts))
or sum(aezNest_n(aezNest, aezNest1), p_gAez(rs, aez, aezNest1, ts)) ),
xaezNest.scale(rs, aez, aezNest, ts) * paezNest.scale(rs, aez, aezNest, ts))

=E=
sum((r_r(rs, rNat), aez, aezNest) $ (landCat_aezNest(landCat, aezNest)
$ ( sum(aezNest_a(aezNest, aezA), p_gAez(rs, aez, aezA, ts))
or sum(aezNest_n(aezNest, aezNest1), p_gAez(rs, aez, aezNest1, ts)) ),
xaezNest(rs, aez, aezNest, ts) * paezNest(rs, aez, aezNest, ts))

/ sum((r_r(rs, rNat), aez, aezNest) $ (landCat_aezNest(landCat, aezNest)
$ ( sum(aezNest_a(aezNest, aezA), p_gAez(rs, aez, aezA, ts))
or sum(aezNest_n(aezNest, aezNest1), p_gAez(rs, aez, aezNest1, ts)) ),
xaezNest.scale(rs, aez, aezNest, ts) * paezNest.scale(rs, aez, aezNest, ts))

```

In order to map this into a tax per unit of land used, the following equation is added:

```

--- map (positive) tax per constant USD from land cover nest to related individual land use sectors

e_emispLndA(rsNat(rNat), a, ts(tRun)) $ ( (not sameas(tRun, "%t0%"))
$ xfFlag(rNat, lnd, a)
$ (emispLnd.range(rNat, a, tRun) ne 0)) ..

emispLnd(rNat, a, tRun) =E= sum(landCat_a(landCat, a), emispLndN(rNat, landCat, tRun));

```

Note that carbon taxation is still an experimental feature.

### 3.10.2 Marginal abatement curves for GHG emissions

Without the MACC mechanism, emission abatement of process emission per unit of output based on the equations shown above is depicted in CGEBox by changes in the input and primary factor mix as emission factors are linked to these inputs and can be taxed. As discussed above, the model makes a distinction between reported emissions and taxed emissions which allows to exclude certain emission sources fully or partially from taxation or pricing. Substitution possibilities in the technologies, as for instance depicted by the structures and parameters taken from GTAP-E and GTAP-Power, allow firms to change their input mix according to the price signals which stem from the emission per unit of input use and carbon prices/taxes. Similarly, final demanders can have CES-sub-nests under the final demand system which allow adjustments in response to price changes stemming from carbon prices or taxes, again following GTAP-E respectively GTAP-Power, on top of the price responsiveness from the demand system.

Accordingly, depending on substitution possibilities in product and final demand, the different emission intensities of factors and products combined with carbon taxing or pricing let to structural changes (output, product and factor demand composition) in the economy towards decarbonization, including changes in imports and exports. Generally, the differentiated updates of use taxes due to carbon taxing or pricing imply allocative losses by forcing firms away from the cost minimal cost mix at the benchmark and consumers away from their preferred consumption bundles. In recursive-dynamic analysis, there are also potentially effects from lower capital accumulation arising from reduced income or changes in saving rates. The model needs also to ensure B.O.P=0. In the default closure, trade balances are flexible and driven by the so-called global bank agent. Carbon taxes in this setup have an impact on foreign savings by changing implicitly expected returns to capital.

Abatement options currently not covered without the MACC mechanism are those which reduce process emissions per unit of output, factor or intermediate input use from non-fossil fuels. These matter clearly mostly in case of production activities with considerably emissions per unit of output or input use (for instance, cement, paddy rice, ruminants). The taxation of benchmark emission factors can imply massive changes in their per unit costs. Not considering cost effective abatement options to reduce emissions per unit of output or factor use, will overexaggerate the impact of taxation. Equally,

if demand for products stemming from activities with high process emission is price inelastic, emission levels might be overestimated.

Kiula et al. 2019 describe an approach where they define one economy wide additional production activity which serves as input into all others, based on Rutherford and Kiula 2013. Rutherford and Kiula 2013 describe also a so-called sector-specific approach where capital can substitute against abatable emissions. We follow their ideas in the following.

### **Basic concept**

The basic concept of the implementation of the AC module is summarized as follows:

1. It is assumed that (points on) Marginal Abatement Cost Curves (MACC)s from bottom-up studies provide the backbone of the module, used to define the abated emissions per unit of input or factor use or per unit of output, and related cost at distinct points defined by carbon tax levels.
2. Total emissions are calculated from process emissions minus abated emissions.
3. The costs of abatement are depicted by a Leontief bundle of updated intermediate input demands per unit of abated emissions, financed from the savings account. If a separate government account is available, it is assumed that the government spending on savings is expanded to cover (a larger part of the abatement costs). The government can collect these partly or in a total by additional production taxes.
4. The integration in the overall model is based on an additional Armington agent which demands products to realize the abatement technologies.

The layout above will imply that the realization of abatement technologies leads to a new source of final demand which is linked to the emission sources of the activities, but not added to their production function nesting. This helps to keep impacts on the overall modeling structure clean and also avoids constructions such as “unproductive capital”. Instead, abatement technologies require specific inputs (for instance, construction, machinery or educational services) which need to be produced or imported, and thus compete with the given intermediate and final demand in the economy. This also allows to track in a transparent manner the impacts of realizing the abatement options.

The implicit assumption here are that the abatement efforts need to be repeated in each period. This implies that, for instance, a one-time capital demand for an abatement technology reported in literature needs to be mapped into the annualized (investment) demand for specific products by the abatement agent.

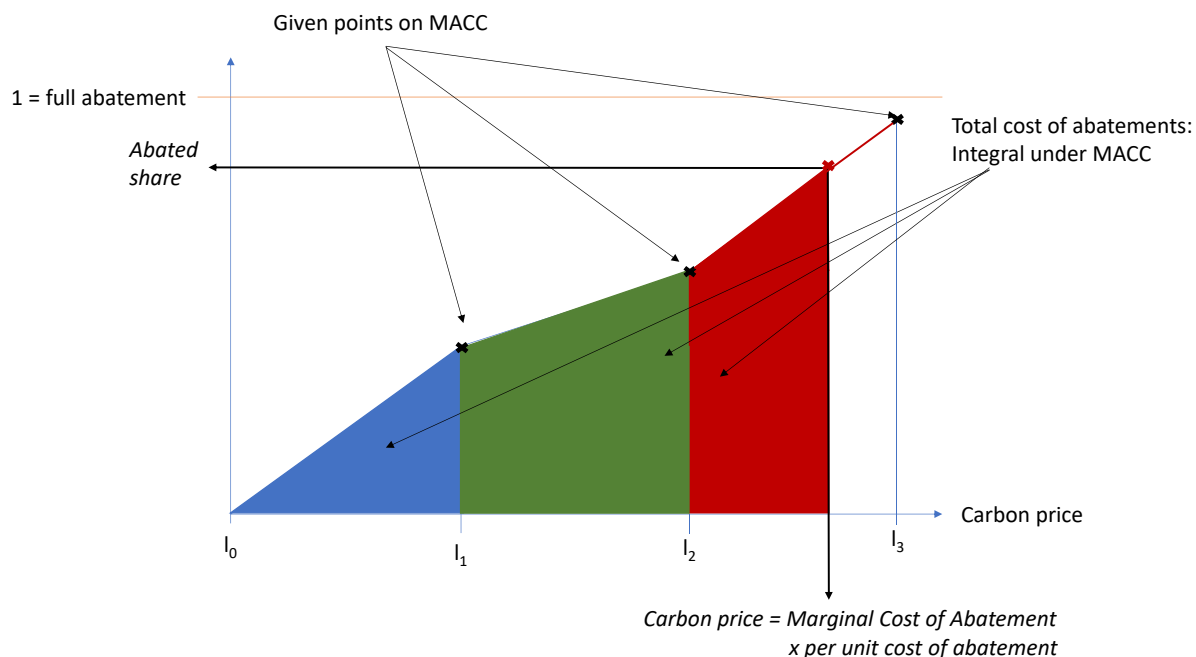
### **Equation structure**

The key variable is  $v\_redEmiaCo2eq$  which expresses the relative reduction in process emissions for a production activity  $a$  and region  $r$ . It can therefore take on values between 0 (= technology not used) and 1 (= technology removes emissions completely). It is driven by the carbon price  $emisp$ , corrected for the price index of abatement cost  $pindMaccs$ :

$$v\_redEmiaCo2eq_{r,a} = f\left(\frac{emisp_r}{pindMaccs_r}\right) \quad (53)$$

In the current code realization, no differences between different emitting sources (factors, intermediate, output levels) or emitted gases (NH<sub>4</sub>, NO<sub>x</sub>, FGAS) exists. The function in (53) is a step-wise linearization of points on MACC curves, see Figure 6. It is important to understand that the

assumed abatement technology is not returns-to-scale, such that average costs of abatement differ from marginal ones.



**Figure 6: Step-wise linearization of points on a MACC**

Let  $emisRed$  define the set of points on the marginal abatement curve for activity  $a$  and region  $r$ , assuming that the same cost points  $l$  are used for all activities and regions while related abatement levels can differ. These points are ordered by increased level of abatement. In the example of Figure 6, three such points  $l_1$  until  $l_3$  are depicted by black crosses.

The related equation is shown next:

$$v_{redEmiaCo2eq_{r,a}} = \sum_l \max[0, \min(cost_l, \frac{emispr_r}{pindMaccs_r}) - cost_{l-1}] \frac{emisRed_{r,a,l} - emisRed_{r,a,l-1}}{cost_l - cost_{l-1}} \quad (54)$$

In the expression above, the emission price is corrected for the price index of abatement services.

In order to calculate the level of abated emissions as a share of the total ones at a given carbon price  $emispr$ , we need to add over the given points on the MACC. For the example above, up to point  $l_2$ , the summation will add up the total differences in emission abatement as the term in the max operator will be equivalent to the denominator  $cost_l - cost_{l-1}$ . As the carbon price is below point  $l_3$ , only a share of the term  $\frac{emisRed_{r,a,l} - emisRed_{r,a,l-1}}{cost_l - cost_{l-1}}$  will be applied for the additional emission reduction achieved between  $l_2$  and  $l_3$  as it is multiplied with  $\max[0, \min(cost_l, \frac{emispr_r}{pindMaccs_r}) - cost_{l-1}]$ . Accordingly, the expression will work as an interpolation which selects the point with the red cross in figure above, by a share of the additional emission abatement achieved from  $l_2$  to  $l_3$  according to given carbon price.

The total cost of abatement services as the integral under the MACC curve are calculated accordingly as shown in the formula below. The first term multiplied  $\frac{1}{2}$  would be equivalent to the blue triangle for



the first step. For follow up steps, the reached emission reduction from the previous step must be taken into account as well:

$$qAbat_{r,a} = \sum_l \max[0, \min\left(\frac{cost_l}{pindMaccs_r}, \frac{emisRed_{r,a,l-1}}{cost_l - cost_{l-1}}\right) - cost_{l-1}] \quad (55)$$

In (55),  $emis$  are the total Co2 equivalent process emissions linked to activity  $a$  in region  $r$ , i.e. emissions per unit of factor use times factor use, emissions per unit of intermediate use times intermediate input use – differentiated by domestic and imported use - and emissions per process level times production output:

$$emis_{r,a} = \sum_{emits} \left( \begin{aligned} &xf_{r,f,a} emifCo2EqTaxed_{r,f,a} \forall f = emits \\ &+xd_{r,i,a} emidCo2EqTaxed_{r,i,a} \forall i = emits \\ &+xm_{r,i,a} emiiCo2EqTaxed_{r,i,a} \forall i = emits \\ &+xp_{r,a} emiaCo2EqTaxed_{r,a} \forall o = emits \end{aligned} \right) \quad (56)$$

The total cost of abatement must be exhausted in a CGE by demand for products and services. To do so, a new final demand agent  $maccs$  is defined which minimizes abatement costs under a Leontief technology. Its demand for products and services  $i$  is defined by summing up over the total abatement services demand of the production activities  $a$  times a Leontief coefficient  $p\_maccsShare_{r,i,a}$ :

$$xa_{r,i,"maccs"} = \sum_a qAbat_{r,a} p\_maccsShare_{r,i,a} \quad (57)$$

The total value of these demands is:

$$valMaccs_r = \sum_a xa_{r,i,"maccs"} pa_{r,i,"maccs"} \quad (58)$$

This allows to define a price index for the  $maccs$  agent:

$$pindMaccs_r = \frac{\sum_a xa_{r,i,"maccs"} pa_{r,i,"maccs"}}{\sum_a xa_{r,i,"maccs"}} \quad (59)$$

This price index is used in two equations above.

### Abatement curves

The functional forms of MACCs (53) are not defined above. Figure 6 depicts a typical case. Typically, certain technologies will only remove a part of the specific emissions and will be chosen according to their cost, i.e. the choice of the abatement technologies to use will depend on the carbon price faced by the firm. The higher the carbon price, the costlier options to abatement will be chosen. As the min and max operator used above are not differentiable, the Veelken-Ulbrich smoothed min function available from GAMS is used instead.

The layout defined above would support multiple abatement technologies which differentiated input demands as shown above.

Finding the data on changes in which inputs are underlying specific MACCs can probably prove quite demanding or even impossible. Possible fallbacks options:

- Using the economic wide composition of investment demands instead.
- Defining some broad types (such as 80% based on new structures and 20% training) etc.

### Technical test

We set up a test framework where we define the input by the emission reduction achieved by a set of emission prices or CO2 tax rates per unit of CO2 equivalent:

```
*
set emisPLevels "Potential steps on abatement cost curve"
/ ep0,ep1,ep5,ep10,ep50,ep100,ep500,ep1000 /;
alias(emisPLevels,emisPLevels1);

parameter p_emisPLevels "CO2 tax levels linked to the steps "
/ ep0 1.E-6,ep1 1,ep5 5,ep10 10,ep50 50,ep100 100,ep500 500,ep1000 1000/;
```

The prices are deemed to cover the relevant range up to a tax of 1000 USD per unit. The emission reduction potential of the different technologies must be defined at these prices:

```
table p_emisRed(rWor,ae,emisPLevels) "Relative reduction in process emissions reduction"

      ep1      ep5      ep10      ep50      ep100      ep500      ep1000
wor.nfm-a      1.E-6      1.E-4
wor.ctl-a      1.E-6      1.E-4
;
```

In the next step, we linearly interpolate between given tax points to define relative abatement levels for missing ones:

```
*
* --- Set (almost) ( zero reduction at tax of zero
*
p_emisRed(rWor,ae,"ep0") $ sum(emisPLevels,p_emisRed(rWor,ae,emisPLevels))
= 1.E-10;

*
* --- make sure that last step is always set
*
p_emisRed(rWor,ae,"ep1000") $ (sum(emisPLevels,p_emisRed(rWor,ae,emisPLevels))
$ (not p_emisRed(rWor,ae,"ep1000")))
= smax(emisPLevels,p_emisRed(rWor,ae,emisPLevels));

*
* --- interpolate missing steps
*
loop(emisPLevels,

p_emisRed(rWor,ae,emisPLevels) $ (not p_emisRed(rWor,ae,emisPLevels))
= [smin(emisPLevels1 $ ((emisPLevels1.pos gt emisPLevels.pos) $ p_emisRed(rWor,ae,emisPLevels1))
- p_emisRed(rWor,ae,emisPLevels-1))
/ [smin(emisPLevels1 $ ((emisPLevels1.pos gt emisPLevels.pos) $ p_emisRed(rWor,ae,emisPLevels1)),
p_emisPLevels(emisPLevels1)) - p_emisPLevels(emisPLevels-1)]
* [p_emisPLevels(emisPLevels) - p_emisPLevels(emisPLevels-1)] + p_emisRed(rWor,ae,emisPLevels-1);

);
```

The outcome of these processes is a completely filled table:

```
-----      80 PARAMETER p_emisRed Relative reduction in process emissions reduction

      ep0      ep1      ep5      ep10      ep50      ep100      ep500      ep1000
wor.nfm-a 1.000000E-10 1.000000E-6 1.000000E-4      0.026      0.237      0.500      0.750      0.750
wor.ctl-a 1.000000E-10 1.000000E-6 1.000000E-4      0.005      0.047      0.100      0.278      0.500
```

Next, we set up a dummy partial equilibrium model for one region and the two activities shown above. The model has downward sloping demand curve  $e\_dem$  which defines the level of production output (supply=demand)  $v\_xp$ . Demand is driven by the per unit production cost  $v\_cost$  which are equal to unity if no CO2 tax is active. The cost function  $e\_cost$  subtract the base process emissions per unit  $p\_emisCoeff$  times the CO2eq price  $emisP$ , considering that per unit process emission drop with increasing carbon prices. The change in per unit emissions is also reflected in the adding up of emissions to regional totals in the equation  $e\_emis$ .

The most demanding part from a numerical perspective is the step-wise function for the emission reduction per unit of output in  $e\_emisRed$ . The term

```
* ( p_emisRed(rWor,ae,emisPLevels)-p_emisRed(rWor,ae,emisPLevels-1))
/ ( p_emisPLevels(emisPLevels)-p_emisPLevels(emisPLevels-1))
```

Defines the changes in relative emission reductions per unit of change in the emission price between two neighboring point on the MACC curve. This term must be multiplied with

```
max(0,min(v_emisP(rWor),p_emisPLevels(emisPLevels)) - p_emisPLevels(emisPLevels-1))
```

If the current carbon tax  $v\_emisP$  exceeds the carbon tax level  $p\_emisPLevels$  at point  $emisPLevels$ , the full change will be applied. The minimum operator will ensure that not more than the differences to the next lower point will be applied. For cases where the current carbon tax does not reach level  $p\_emisPLevels(emisPLevels-1)$ , the max operator is needed.

The reduction level is hence defined as:

```
sum[emisPLevels $ p_emisRed(rWor,ae,emisPLevels-1),
{max(0,
  min(p_check(rWor,ae,draws,"tax"),p_emisPLevels(emisPLevels)) - p_emisPLevels(emisPLevels-1))}

* ( p_emisRed(rWor,ae,emisPLevels)-p_emisRed(rWor,ae,emisPLevels-1))
/ ( p_emisPLevels(emisPLevels)-p_emisPLevels(emisPLevels-1))];
```

The min and max operators are not differentiable. We therefore use the ncpVUSin function from GAMS as smooth approximation of min:

|                    |                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ncpVUSin(r,s[,MU]) | <div style="border-left: 1px solid black; padding-left: 10px;"> <p><b>NCP Veelen-Ulbrich: smoothed min(r,s)</b></p> <math display="block">\text{result} = \begin{cases} \frac{(r+s- t )}{2} &amp; \text{if }  t  \geq \mu \geq 0 \\ \frac{(r+s-(\frac{2\mu}{\pi} \sin(\frac{\pi}{2\mu} + \frac{3\pi}{2}) + \mu))}{2} &amp; \text{otherwise} \end{cases}</math> <p>where <math>t = r - s</math>. Default <math>MU = 0</math>.</p> </div> |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Which leads to the term

```
sum[emisPLevels $ p_emisRed(rWor,ae,emisPLevels-1),
{-ncpVUSIN(0,
  -ncpVUSIN(p_check(rWor,ae,draws,"tax"),p_emisPLevels(emisPLevels),1.E-3) + p_emisPLevels(emisPLevels-1),1.E-3)}

* ( p_emisRed(rWor,ae,emisPLevels)-p_emisRed(rWor,ae,emisPLevels-1))
/ ( p_emisPLevels(emisPLevels)-p_emisPLevels(emisPLevels-1))];
```

This term is introduced in the model equation  $e\_emisRed$  below. At  $MU=1.E-3$ , no measurable error in the approximation occurs of the relevance range of carbon taxes [0,1000].

```

variable v_emisRed(rWor,aE) "Relative CO2 price driven reduction in process emissions"
v_emis(rWor) "Total regional emissions"
v_emisp(rWor) "CO2 price"
v_xp(rWor,aE) "Supply (= demand in single region test model)"
v_cost(rWor,aE) "Per unit production cost";

equation e_emisRed(rWor,aE),e_emis(rWor),e_dem(rWor,aE),e_Cost(rWor,aE);

parameter p_emisCoeff(rWor,aE);

p_emisCoeff(rWor,AE) = 0.001;

e_dem(rWor,aE) ..

    v_xp(rWor,aE) =E= 100 - 50 * v_cost(rWor,aE);

e_cost(rWor,aE) ..

    v_Cost(rWor,aE) =E= 1 + p_emisCoeff(rWor,aE) * (1 - v_emisRed(rWor,aE)) * v_emisP(rWor);

e_emis(rWor) ..

    v_emis(rWor) =E= sum(aE, v_xp(rWor,aE) * p_emisCoeff(rWor,aE) * (1 - v_emisRed(rWor,aE)));

e_emisRed(rWor,aE) $ (v_emisRed.range(rWor,aE) ne 0) ..

    v_emisRed(rWor,aE) =E=

        sum[emisPLevels $ p_emisRed(rWor,ae,emisPLevels-1),
        {-ncpVUSIN(0,
        -ncpVUSIN(v_emisP(rWor),p_emisPLevels(emisPLevels),1.E-3) + p_emisPLevels(emisPLevels-1),1.E-3))
        * ( p_emisRed(rWor,ae,emisPLevels)-p_emisRed(rWor,ae,emisPLevels-1))
        / ( p_emisPLevels(emisPLevels)-p_emisPLevels(emisPLevels-1))];

model m_emisRed / e_emisRed,e_emis,e_dem,e_cost /;

```

We test the approach by first solving the model with a fixed carbon prices of 1 and emission coefficients of 1.E-3:

```

model m_emisRed / e_emisRed,e_emis,e_dem,e_cost /;
m_emisRed.solveLink = 5;
m_emisRed.limCol = 0;
m_emisRed.limRow = 0;
m_emisRed.holdfixed = 1;

p_emisCoeff(rWor,AE) = 0.001;
v_emisRed.lo(rWor,AE) = 0;
v_emisRed.up(rWor,AE) = 1;
v_emisP.fx(rWor) = 0;
option CNS = CONOPT4;
solve m_emisRed using CNS;

```

Afterwards, we fixing the emissions to a ceiling (from experiment *r0* onwards) which is equal to 90% of the emissions under the previous iteration and render the carbon tax endogenous.

```

set redStep /r0*r10 /;
parameter p_res;

loop(redStep,
    v_emisP.up(rWor) = +inf;
    v_emisP.lo(rWor) = 0;
    v_emis.fx(rWor) = v_emis.l(rWor);
    solve m_emisRed using CNS;
    p_res(rWor,ae,"xp",redStep) = v_xp.l(rWor,ae);
    p_res(rWor,ae,"emis",redStep) = v_xp.l(rWor,ae) * p_emisCoeff(rWor,aE) * (1 - v_emisRed.l(rWor,aE));
    p_res(rWor,ae,"emisRed",redStep) = v_emisRed.l(rWor,aE);
    p_res(rWor,"","emis",redStep) = v_emis.l(rWor);
    p_res(rWor,"","emisP",redStep) = v_emisP.l(rWor);
    p_res(rWor,ae,"cost",redStep) = 1 + p_emisCoeff(rWor,aE) * (1 - v_emisRed.l(rWor,aE)) * v_emisP.l(rWor);
    p_res(rWor,"","solve",redStep) = m_emisRed.modelStat;
    v_emis.l(rWor) = v_emis.l(rWor) * 0.90;
    v_emisp.l(rWor) = v_emisp.l(rWor) + 10;
);

```

The first solve was however infeasible if not starting value for the emission prices is set as shown above.

This leads to the following results:

|               | r0     | r1     | r2     | r3     | r4     | r5      | r6      | r7      | r8      | r9      | r10     |
|---------------|--------|--------|--------|--------|--------|---------|---------|---------|---------|---------|---------|
| nfm-a.xp      | 50.000 | 48.740 | 48.022 | 47.622 | 47.498 | 46.314  | 45.229  | 44.438  | 43.931  | 43.695  | 43.711  |
| nfm-a.emis    | 0.050  | 0.043  | 0.036  | 0.030  | 0.024  | 0.021   | 0.019   | 0.017   | 0.015   | 0.013   | 0.011   |
| nfm-a.emisRed |        | 0.125  | 0.252  | 0.373  | 0.486  | 0.537   | 0.579   | 0.621   | 0.662   | 0.702   | 0.741   |
| nfm-a.cost    | 1.000  | 1.025  | 1.040  | 1.048  | 1.050  | 1.074   | 1.095   | 1.111   | 1.121   | 1.126   | 1.126   |
| ctl-a.xp      | 50.000 | 48.595 | 47.488 | 46.492 | 45.609 | 43.044  | 40.434  | 38.058  | 35.914  | 33.999  | 32.311  |
| ctl-a.emis    | 0.050  | 0.047  | 0.045  | 0.043  | 0.041  | 0.038   | 0.034   | 0.031   | 0.028   | 0.026   | 0.024   |
| ctl-a.emisRed |        | 0.025  | 0.050  | 0.075  | 0.097  | 0.126   | 0.156   | 0.186   | 0.215   | 0.244   | 0.271   |
| ctl-a.cost    | 1.000  | 1.028  | 1.050  | 1.070  | 1.088  | 1.139   | 1.191   | 1.239   | 1.282   | 1.320   | 1.354   |
| .emis         | 0.100  | 0.090  | 0.081  | 0.073  | 0.066  | 0.059   | 0.053   | 0.048   | 0.043   | 0.039   | 0.035   |
| .emisP        |        | 28.817 | 52.913 | 75.806 | 97.271 | 159.232 | 226.764 | 293.385 | 358.909 | 423.069 | 485.508 |
| .solve        | 16.000 | 16.000 | 16.000 | 16.000 | 16.000 | 16.000  | 16.000  | 16.000  | 16.000  | 16.000  | 16.000  |

As expected, with stricter emission ceilings, production quantities have to adjust. In order to reduce emissions (*emis*) to 35% of their benchmark levels, a tax of almost 500 USD is required (*emisP*).

The reduction of the cattle activity “ctl-a” is higher despite an equal base emission coefficient of 0.001. This reflects that abatement possibilities in cattle are more limited. Whereas the non-ferrous metal activity “nfm-a” is assumed to abate up to 75% of its per unit emissions, reached at an emission price of 500 USD per ton CO<sub>2</sub>eq, the maximal abatement level for cattle is 50% and requires a tax of 1000 USD. At a 500 USD tax rate, the cattle sector will reduce its per unit emissions by around 28%, as shown above. Accordingly, the per unit cost of cattle production including the emission taxes reach up to 1.35 at the lowest ceiling, compared to 1.12 for non-ferrous metals. As output levels at the benchmark and the slopes of the demand curves are assumed identical, the cattle sector shrinks more in absolute and relative terms, from 50 to around 33 units compared to non-ferrous metals where output is reduced from 50 to about 43 units.

The test model underlines that the approach is feasible. It was hence chosen the implementation in CGEbox as described next.

### Integration into CGEBox

The integration into the CGE model CGEBox requires to map the equations above into the existing structure of the CGE. The code is added to the CO<sub>2</sub> and Non-Co<sub>2</sub> emission module (*co2eq/coe2eq.gms*). It comprises four main components:

1. New equations which define the emission reduction, the related demand for intermediate and their value and the link into the overall CGE.
2. An initialization block where the necessary parameters and start values for variables are defined.
3. Controls on the GUI
4. And the integration into the reporting part.

### Equations

Based on concept, there are three new equations in the systems. The first defines the relative reduction in process emissions  $v\_redEmiaCo2Eq$  as a smoothed linear step-function of the emission price *emisP* as discussed above, depicting the abatement level according to the MACC curve:

```

e_redemiaCo2eq(rsNat(rNat),a,ts) $ ((redemiaCo2eq.range(rsNat,a,ts) ne 0) $ p_emisRed(rsNat,a,"ep0") $ xpFlag(rsNat,a)) ..
redemiaCo2eq(rsNat,a,ts) =E=
sum[(emisPLevels,emisPLevels1) $ ( p_emisRed(rsNat,a,emisPLevels) $ p_emisRed(rsNat,a,emisPLevels1)
$ (emisPLevels.pos gt emisPLevels1.pos)
$ (p_emisRed(rsNat,a,emisPLevels1) eq smax(emisPLevels2 $ (emisPLevels.pos gt emisPLevels2.pos), p_emisRed(rsNat,a,emisPLevels2))))),
[-ncpVUSIN(0,
[-ncpVUSIN(emisPLevels,emisPLevels1),max(1.E-4,1.E-4*p_emisPLevels(emisPLevels))]
- p_emisPLevels(emisPLevels1)],max(1.E-4,1.E-4*p_emisPLevels(emisPLevels1))))
* ( p_emisRed(rsNat,a,emisPLevels)-p_emisRed(rsNat,a,emisPLevels1))
/ ( p_emisPLevels(emisPLevels)-p_emisPLevels(emisPLevels1))
];

```

The equation is introduced for the combination of region *rsNat*, activity *a* and solution period *ts* where a process emission coefficient *emiaCo2eqTaxed* is given and an emission reduction coefficient *p\_emisRed* is defined. The initialization code ensures that all points on the MACC curve are present. It is therefore sufficient to test for the highest price level, only. The RHS comprises the step-wise linear function.

A second equation defines per unit emisison of each activity *emisA*:

```

e_emisA(rsNat(rNat),a,ts(tRun)) $ ((redemiaCo2eq.range(rsNat,a,ts) ne 0) $ p_emisRed(rsNat,a,"ep0") $ xpFlag(rsNat,a)) ..
emisA(rnat,a,ts)*m_xp(rsNat,a,ts)s(xp,"rnat,a,ts") =E=
{
--- quantity abated process emissions per unit of output
p_emisCo2eqTaxed(rsNat,a,ts)*(1+p_emisShift(rsNat,"",a,tRun))
* (m_xp(rsNat,a,tRun) $ (not disr(rsNat))
$ $ifi "%ModulesGTAP_NUTS2%"=="ON" + sum(r_r(subr,rsNat) $ xpFlag(subr,a) m_xp(subr,a,ts))
)

--- quantity of abated process emissions per unit of factor use
+ sum(f $ xfFlag(rsNat,f,a),
p_emisCo2eqTaxed(rsNat,f,a,ts)*(1+p_emisShift(rsNat,f,a,tRun))*xf(rsNat,f,a,ts))

--- quantity of abated process emissions per unit of intermediate product use
+ sum( (j) $ (m_alpha(rsNat,j,a,ts) $ p_emisCo2eqTaxed(rsNat,j,a,ts)),
p_emisCo2eqTaxed(rsNat,j,a,ts) * xd(rsNat,j,a,ts)*p_nCo2ShareD(rsNat,j,a) * (1+p_emisShift(rsNat,j,a,ts)))

+ sum( (j) $ (m_alpha(rsNat,j,a,ts) $ p_emisCo2eqTaxed(rsNat,j,a,ts)),
p_emisCo2eqTaxed(rsNat,j,a,ts) * xm(rsNat,j,a,ts)*p_nCo2ShareI(rsNat,j,a) * (1+p_emisShift(rsNat,j,a,ts)))
)s(xp,"rnat,a,ts")*1000;

```

The value of the abated emission is calculated from the points on the MACC, and the emis price *emisP* deflated with the per unit cost of emissions *pindMaccs*, multiplied with the per unit emissions *emisA*:

```

--- demand of emission services per unit of output (this is a value, not a quantity)
e_vAbat(rsNat(rNat),a,ts(tRun)) $ ((redemiaCo2eq.range(rsNat,a,ts) ne 0) $ p_emisRed(rsNat,a,"ep0") $ xpFlag(rsNat,a)) ..
vAbat(rsNat,a,ts)s(vAbat,"rsNat,a,ts") =E=
sum[(emisPLevels,emisPLevels1) $ ( p_emisRed(rsNat,a,emisPLevels) $ p_emisRed(rsNat,a,emisPLevels1)
$ (emisPLevels.pos gt emisPLevels1.pos)
$ (p_emisRed(rsNat,a,emisPLevels1) eq smax(emisPLevels2 $ (emisPLevels.pos gt emisPLevels2.pos), p_emisRed(rsNat,a,emisPLevels2))))),
[-ncpVUSIN(0,
[-ncpVUSIN(emisP(rsNat,ts)/pindMaccs(rsNat,ts),p_emisPLevels(emisPLevels),max(1.E-4,1.E-4*p_emisPLevels(emisPLevels)))
- p_emisPLevels(emisPLevels1)],max(1.E-4,1.E-4*p_emisPLevels(emisPLevels1))))
* [ 0.5 * (p_emisRed(rsNat,a,emisPLevels)-p_emisRed(rsNat,a,emisPLevels1))
+ p_emisRed(rsNat,a,emisPLevels1)
]
]*emisA(rnat,a,ts)s(vAbat,"rsNat,a,ts")
;

```

Note here that we calculate the costs of process emission abatement not based on the last unit abated which defines the carbon tax according to marginal costs of abatement per unit being equal to the tax rate. Rather, we integrate over the step-wise linear function and thus consider that the first units are cheaper to abate, such as average costs of abatement are below marginal ones.

The third equation introduces a new demand agent called “*maccs*” in the model. Its demand is financed from the saving account. It is hence treated as an investment activity, but can have its own

demand shares, defined by the parameter  $p\_macShare$ . As discussed below, the investment activity is not increasing the physical capital stock in follow-up period and competes with government demand. More on this later.

```

--- demand for intermediates for process emission reduction
e_xaMaccs(rsNat(rnat),i,"maccs",ts) $ xaiFlag(rsNat,i,"maccs") ..

    xa(rsNat,i,"maccs",ts)*m_pa(rsNat,i,"maccs",ts)
s(xa,'rsNat,i,"maccs",ts')
    =E= {sum(a $ ((redemiaCo2eq.range(rsNat,a,ts) ne 0) $ p_emisRed(rsNat,a,"ep0") $ xpFlag(rsNat,a)),
        vAbat(rsNat,a,ts)/1000*p_maccsShare(rsNat,i,a)*m_xp(rsNat,a,ts))+1.E-10}
s(xa,'rsNat,i,"maccs",ts');

```

The right-hand side distributes the total value of abated process emissions  $valAbat$  of each activity per unit of output to product demand based on the shares  $p\_macShare$ , multiplied with total output based on the macro  $m\_xp$ . The left-hand side defines the demand as the product of quantity and prices. The values on the parameter  $p\_macShare$  has to add up to unity when summed up over the products  $i$  for any activity  $a$  to ensure that the total abatement costs are exhausted by demand for services and products.

The last equation defines the total value of these demands over all products:

```

--- total value of intermediate demand linked to process emission reduction
e_valMaccs(rsNat(rnat),ts) ..

    valMaccs(rsNat,ts)s(valMaccs,"rsNat,ts")
    =E= sum(i $ xaiFlag(rsNat,i,"maccs"),
        xa(rsNat,i,"maccs",ts)*m_pa(rsNat,i,"maccs",ts))s(valMaccs,"rsNat,ts");

```

To reflect that changes in the prices of abatement services will impact their level, a prices index is defined:

```

--- attention: this formulation can be used with the current layout only, where all activities
    have the same Leontief shares p_macShare
e_pindMaccs(rsNat(rnat),ts(tRun)) ..

    pindMaccs(rsNat,ts) =E= sum(i $ sum(a,p_maccsShare(rsNat,i,a)),
        smax(a,p_maccsShare(rsNat,i,a)) * m_pa(rsNat,i,"maccs",ts))*p_calPindMaccs(rsNat);

```

This new variable is the one case where updates in the core model (*gams/model.gms*) are required. As it is assumed that savings finance the abatement technologies due to their investment character, we need to reduce the value of gross investments accordingly by the value of the related demand,  $valMaccs$ :

```

e_yi(rsNat,ts)
$iftheni.emp not "%modelType%"=="EMP"
    $ ( ((not rres(rsNat)) or singleCountry or (card(rsNat) eq 1) or (pnum.range(ts) gt 0)) )
$endif.emp
..

    yi(rsNat,ts)/m_var2Scale1(yi,rsNat,ts) =e= (
*
*      ---- value of depreciation
    valDep(rsNat,ts)*m_var2Scale(valDep,rsNat,ts)
*
*      ---- savings of the regional household
    + rsav(rsNat,ts)*m_var2Scale(rsav,rsNat,ts)
$if "%modelType%"=="EMP"
    + walras(ts) $ rres(rsNat)
*
*      ---- direct investments of firm
    + invFromCapInc(rsNat,ts)
$if "%MACCsProcessEmis%"=="on"
    - valMaccs(rsNat,ts)
*
*      ---- Value of foreign savings in local currency
    + valsavf(rsNat,ts)*m_var2Scale(valSavf,rsNat,ts))/yi.scale(rsNat,ts);
.

```

As gross investment are reduced, the MACC demand is not increasing capital accumulation. The term *valMaccs* is introduced only when the MACC mechanism is switched on, i.e. the global *MACCsProcessEmis* is set to “on”. This is done via a checkbox on the GUI as shown below.

The same term *valMaccs* must be also reflected in the WALRAS check equation which calculates the above relation for the residual region *rres* for which is equation *e\_yi* is not active:

```

$wlras(ts) ..
/
--- equivalent to savfeq (investments = savings), for residual region
wlras(ts) == sum(rres(rNat) $ rNat(rNat),
  (yi(rNat,ts)*m_var2Scale(yi,rNat,ts)
    - (valDep(rNat,ts)*m_var2Scale(valDep,rNat,ts)
      + rsav(rNat,ts)*m_var2Scale(rsav,rNat,ts)
      + invFromCapInc(rNat,ts)
    )
  )
$if1 %ModulesCO2_Emissions%==on + sum(ets_rnat(ets,rNat), emisEtsTrd.l(rNat,ts)*emisP(rNat,ts) $ (card(rNat) eq 1) + emisEtsTrd(rNat,ts)*emisP(rNat,ts) $ (card(rNat) gt 1))
$if1 %MACCsProcessEmis%==on - valMaccs(rNat,ts)
+ valSavf(rNat,ts)*m_var2Scale(valSavf,rNat,ts))/yi.scale(rNat,ts) ;;

```

During simulation, in order to avoid that process emission abatement cost crowd out productive investments, for instance, the saving share of the government can be adjusted in a recursive-dynamic fashion taken last period’s government spend and the value of the government subsidized cost of abatement into account:

```

*
* --- shift government consumptions towards more investments
*
betasG.fx(rsNat,ts) = min(0.8,betasG.l(rsNat,"t0%") + valMaccs.l(rsNat,ts)/(rsav.l(rsNat,ts)+yg.l(rsNat,ts)));

```

Furthermore, the macro which adds to the production taxes *prdtx* emission taxes linked to process emissions had to be expanded in order into reflect that emission coefficients now can be endogenously reduced based on the variable *v\_redEmiaCo2eq*:

```

$$macro m_prdtx(r,a,i,t) (1 + prdtx(r,a,i,t)) \
  $$if1 %ModulesCO2_Emissions%==on + emisP(r,t) * p_emiaCo2eqTaxed(r,a,t) * (1-redEmiaCo2eq(r,a,t) + p_emisShift(r,"",a,t))

```

Similarly, the macro for factor taxes has to consider now the share of abated process emissions per unit of factor use:

```

$$macro m_fctTaxes(r,f,a,t) ( 1 + fctts(r,f,a,t) + fcttx(r,f,a,t) + fcttxShift(r,t) $ fTaxShift(f,a)) \
  $$ifthen1.CO2 %ModulesCO2_Emissions%==on
    + emisP(r,t) * emisCo2eqTaxed(r,f,a,t) * (1-v_redEmiaCo2eq(r,a,t)) \
  $$if1.AEZ %ModulesGTAP_AEZ%=="on" + emisPLnd(r,a,t) $ lnd(f)
$$endif.CO2

```

And finally, the definitions of imported and domestic prices need to take the abated shares into account:

```

*
* --- the macro m_pdp substitutes out the domestic sales prices: with an infinite CET for exports, use the supply
* price ps, otherwise, use the domestic price
*
$$macro m_pdpDef(r,i,aa,t) ( m_domPrice(r,i,aa,t) \
  * (1 + dintx(r,i,aa,t) + itxshft(r,i,t))
*
* --- add emission taxes for domestic use
*
$if1 %ModulesCO2_Emissions%==on + emisP(r,t) * emisCo2eqTaxed(r,i,aa,t) * (1-v_redEmiaCo2eq(r,aa,t) * p_nCo2ShareD(r,i,aa) $ a(aa)) \
}

*
* --- the macro m_pmp substitutes out the imports' sales prices
*
$$macro m_pmpDef(r,i,aa,t) ( \
*
* ---- average import price ( pmt or pmtMrio )
*
m_pmt(r,i,aa,t) * (1 + mntx(r,i,aa,t) + itxshft(r,i,t)) \
*
* --- add emission taxes for imported use
*
$if1 %ModulesCO2_Emissions%==on + emisP(r,t) * emisCo2eqTaxed(r,i,aa,t) * (1-v_redEmiaCo2eq(r,aa,t) * p_nCo2ShareI(r,i,aa) $ a(aa)) \
}

```

This required to set fix this variable to zero if the MACC extension is not in use, in “co2Eq/coe2Eq\_setVars.gms”:



```

...
$$iftheni.maccs not "%MACCsProcessEmiss=="on"

    redemiaCo2eq.fx(rsNat,a,tSim) $ (redemiaCo2eq.range(rsNat,a,"%t0%") eq 0) = 0;

```

A share *abatCostSharePrdtx* of the MAC can be allocated to production taxes according to the following equation:

```

e_prdtxAbat(rsNat,a,i,ts(tRun)) $ ((prdtx.range(rsNat,a,i,ts) ne 0) $ (redemiaCo2eq.range(rsNat,a,ts) ne 0) $ p_emisRed(rsNat,a,"ep0") $ xFlag(rsNat,a,i)) ..
[prdtx(rsNat,a,i,ts)-prdtx.l(rsNat,a,i,"%t0%")]/max(1,prdtx.l(rsNat,a,i,tRun-1))
=e= (vAbat(rsNat,a,ts)*0.001 * %AbatCostSharePrdtx + p_addPrdtx(rsNat,a,ts))/max(1,prdtx.l(rsNat,a,i,tRun-1));

```

The left-hand side defined the value of the changed tax income (difference in tax rates time production quantity as the tax base), the right-hand side defines the total activity specific costs of abatement times the share which is charged to the activity itself. The remainder of the cost need to be subsidized by government.

An extension for of the myGTAP model with endogenous government borrowing considers that the MAC demand is financed by a mix of (increased) production taxes and government subsidies:

```

...
--- government account: gov savings as a share of GDP (benchmark relation)
s rsavg(rsNat(rNat),ts(t)) $ ((rsavg.range(rNat,t) ne 0) $ (newGovBond.range(rnat,t) ne 0)) ..
rsavg(rNat,t)s(rsavg,'rnat,t')
=e= [rsavg.l(rnat,"%t0%")-betaSG(rnat,t)*yg.l(rnat,"%t0%")]*s(rsavg,'rnat,t')
* [ regy(rnat,t)/regy.l(rnat,"%t0%")
* (1 - spendShrByGovInc(rnat,t)
+ govBondDamp(rNat,t) $ (newGovBond.range(rnat,t) ne 0)
)
+ (yTaxTot(rnat,t)+aidi(rnat,t)
$ $if "%MACCsProcessEmiss=="on" - valMaccs(rnat,t)
-sum((r_r(rNat,rs),h) $ hFlag(rs,h), p_payPastGovBonds(rs,h,t)))/yTaxTot.l(rnat,"%t0%")
* (spendShrByGovInc(rnat,t) + govBondDamp(rNat,t) $ (newGovBond.range(rnat,t) ne 0))
]
--- demand for MACC
$ $if "%MACCsProcessEmiss=="on" + valMaccs(rnat,t)s(rsavg,'rnat,t')
;

```

The last term will increase government saving by the value of the MACCs. The first term drives productive savings demand of the government partly by GDP and partly a tax income, corrected by the share of MAC which is paid via production taxes by the activities.

## Initialization

The initialization consists of two blocks:

### 1. Input of the MACCs

Currently, a dummy table is available (see *gams/scen/ghg/maccs\_dummy.gms*) for testing purposes:

```
table p_emisRed(*,*,emisPLevels) "Relative reduction in process emissions reduction"

      ep1      ep5      ep10      ep50      ep100      ep500      ep1000      ep5000
wor.nmm-a    1.E-6    1.E-4    0.75    0.95    1.00    0.75    0.95    1.00
wor.bph-a    1.E-6    1.E-4    0.75    0.95    1.00    0.75    0.95    1.00
wor.chm-a    1.E-6    1.E-4    0.75    0.95    1.00    0.75    0.95    1.00
wor.ele-a    1.E-6    1.E-4    0.50    0.95    1.00    0.50    0.95    1.00
wor.wtr-a    1.E-6    1.E-4    0.50    0.75    0.99    1.00    0.99    1.00
wor.ctl-a    1.E-6    1.E-4    0.10    0.50    0.90    1.00    0.90    1.00
wor.rmk-a    1.E-6    1.E-4    0.10    0.50    0.90    1.00    0.90    1.00
wor.rest    1.E-6    0.05    0.10    0.20    0.75    1.00    0.90    1.00
;
```

It comprises made-up default values for the world; for some sectors, explicit assumptions are invented for some sectors (non-minerals nec, chemicals, electric equipment, water, cattle and row milk) and for all others, one MACC curve is defined in the row “rest”. This data serves to test the functioning of the extension.

In a first step, we copy the defaults stored on “rest” to any activity with process emissions and no explicit MACC information:

```
set aProcEmis(a) "Activities with process emissions";
aProcEmis(a) $ sum(rNat, emiaCo2EqTaxed(rnat,a,"%t0%")) = YES;

*
* --- if no information on specific activity is given use information for "rest"
*
p_emisRed("wor",aProcEmis,emisPLevels) $ (not sum(emisPLevels1, p_emisRed("wor",aProcEmis,emisPLevels1)))
= p_emisRed("wor","rest",emisPLevels);
*
```

The next statements have been discussed above. They ensure that all points on the MACCs curves are present, based on linear interpolation:

```
*
* --- Set (almost) ( zero reduction at tax of zero
*
p_emisRed(rWor,aProcEmis,"ep0") $ sum(emisPLevels,p_emisRed(rWor,aProcEmis,emisPLevels))
= 1.E-10;

*
* --- make sure that last step is always set
*
p_emisRed(rWor,aProcEmis,"ep1000") $ (sum(emisPLevels,p_emisRed(rWor,aProcEmis,emisPLevels))
$ (not p_emisRed(rWor,aProcEmis,"ep1000")))
= smax(emisPLevels,p_emisRed(rWor,aProcEmis,emisPLevels));

*
* --- linearly interpolate missing steps
*
loop(emisPLevels,

  p_emisRed(rWor,aProcEmis,emisPLevels) $ ((not p_emisRed(rWor,aProcEmis,emisPLevels)) $ p_emisRed(rWor,aProcEmis,"ep0"))
  = [smin(emisPLevels1 $ ((emisPLevels1.pos gt emisPLevels.pos) $ p_emisRed(rWor,aProcEmis,emisPLevels1)),
    p_emisRed(rWor,aProcEmis,emisPLevels1))
    - p_emisRed(rWor,aProcEmis,emisPLevels-1)]
  / [smin(emisPLevels1 $ ((emisPLevels1.pos gt emisPLevels.pos) $ p_emisRed(rWor,aProcEmis,emisPLevels1)),
    p_emisPLevels(emisPLevels1)) - p_emisPLevels(emisPLevels-1)]

  * [p_emisPLevels(emisPLevels) - p_emisPLevels(emisPLevels-1)] + p_emisRed(rWor,aProcEmis,emisPLevels-1);
);
```

In case where no model region specific information is given (which is the case for the test data set for all regions), the world MACC is used:

```
*
* --- use world average if no information for country is given
*
p_emisRed(rNat,aProcEmis,emisPLevels) $ ((not p_emisRed(rNat,aProcEmis,emisPLevels)) $ emiaCo2EqTaxed(rnat,aProcEmis,"%t0%"))
= p_emisRed("wor",aProcEmis,emisPLevels);
```

## 2. Introduction of the new Armington agent

In order to introduce the demand for abatement into all relevant equations (domestic and import demand balances, tax revenues, definition of GDP, etc.), the set of SAM columns *is* and Armington agents *aa* in the model is extended (in *build/load\_gtapAgg.gms*):

```
$if not "%modulesCo2_emissions%"=="on" $setGlobal MACCsProcessEmis off

set is /
    set.is1
    set.set_hhslds
    $if "%MACCsProcessEmis%"=="on" maccs
/;

set aa(is) /
    set.aa_ori
    set.set_hhslds
    $if "%MACCsProcessEmis%"=="on" maccs
/;
```

To avoid that the global for the MACCs is on and variables such as *valMaccs* are introduced in the model despite the co2 emission accounting being switched off, the global is automatically set to off in this case as shown above.

For our test implementation, we assume that the abatement requires a mix of services, construction and machinery:

```
*
*   ---- assumed shares: 36% is better management, linked to demand for education, Business services nec
*                           and linked to services to implement (communication, other financial mediation)
*
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"obs-c") = 0.15;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"ofi-c") = 0.03;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"edu-c") = 0.15;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"cmn-c") = 0.03;
*
*   ---- rest goes partly to construction and partly to other machiney and equipment (filters, ...)
*   assume 2% insurance cost for new installations
*
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"cns-c") = 0.64 * .21;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"lum-c") = 0.64 * .03;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"ome-c") = 0.64 * .39;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"eeq-c") = 0.64 * .10;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"omf-c") = 0.64 * .10;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"trd-c") = 0.64 * .05;
p_maccsShare(rNat,i,aProcEmis) $ sameas(i,"ins-c") = 0.64 * .02;
```

Note that the demand for construction services “cns-c” is relatively small to avoid that indirectly a higher demand for cement is generated.

As we need information in consumption taxes and domestic and imported demand shares for the new agent, we borrow those from investments. This implies that we have to delete all MACC demand shares where no investment demand is found:

```
*
*   --- MACC demand is derived from investment demand, delete shares where not inv demand is observed
*
p_maccsShare(rNat,i,aProcEmis) $ (not xaFlag(rnat,i,"inv")) = 0;
```

We defined from the remaining shares all products which show MACC demand and scale the demand shares to unity to ensure the FOC condition for cost-minimal marginal abatement:

```

set iMaccs(i) "Products with MACCS demand";
iMaccs(i) $ sum((r,aProcEmis),p_maccsShare(r,i,aProcEmis)) = YES;
*
* --- scale sajres to unity
*
p_maccsShare(rNat,i,aProcEmis) $ p_maccsShare(rNat,i,aProcEmis)
= p_maccsShare(rNat,i,aProcEmis) * 1/sum(iMaccs,p_maccsShare(rNat,iMaccs,aProcEmis));

```

Next, we introduce the Armington demand into the model based on the flag *xaFlag* and borrow the domestic and imported demand shares and related taxes from investment demand:

```

*
* --- Set armington flag and tax over shares and taxes from investment
*
xaFlag(rNat,iMaccs,"maccs") $ xaFlag(rNat,iMaccs,"inv") = YES;
xaFlag(rNat,"maccs",iMaccs) $ xaFlag(rNat,"inv",iMaccs) = YES;

p_alphaD(rnat,iMaccs,"maccs",t0) = p_alphaD(rnat,iMaccs,"inv",t0);
p_alpham(rnat,iMaccs,"maccs",t0) = p_alpham(rnat,iMaccs,"inv",t0);

mintx(rnat,iMaccs,"maccs",t0) = mintx(rnat,iMaccs,"inv",t0);
dintx(rnat,iMaccs,"maccs",t0) = dintx(rnat,iMaccs,"inv",t0);

```

We finally assign a quite tiny starting value and from there import and domestic demands:

```

*
* --- almost zero starting values
*
xa(rnat,iMaccs,"maccs",t0) $ xa(rnat,iMaccs,"inv",t0) = 1.E-10;
pa(rnat,iMaccs,"maccs",t0) $ xa(rnat,iMaccs,"inv",t0) = 1;
xm(rnat,iMaccs,"maccs",t0) $ xa(rnat,iMaccs,"inv",t0) = xa(rnat,iMaccs,"maccs",t0) * xm(rnat,iMaccs,"inv",t0)/xa(rnat,iMaccs,"inv",t0);
xd(rnat,iMaccs,"maccs",t0) $ xa(rnat,iMaccs,"inv",t0) = xa(rnat,iMaccs,"maccs",t0) * xd(rnat,iMaccs,"inv",t0)/xa(rnat,iMaccs,"inv",t0);

valMaccs(rnat,t0) = 0;
valAbat(rnat,a,t0) = 0;

```

## Interface

On the interface, in the CO2 emission panel, there is now a checkbox to switch the MACC mechanism on and to select the file the MACC information (stored under *gams/scen/ghg/*; the file name must start with *maccs\_*):

The “Maximum abatement option to consider” allows to restrict the considered points on the MACCs. If emission prices are endogenous under a ceiling, removing more expensive options can imply only that other abatement options, such as movement on the production function towards less emitting inputs, must be used instead which should normally imply that total cost will be higher.

## Reporting

The reductions in emission coefficients are reflected when the process emissions are assigned to the *p\_results* array used by the exploitation tool, both for taxed and individual emissions:

```

p_results(rsNat,"CO2eqTaxed","xp",a,"dom",%arg1%, "%version%") $ xp.1(rsNat,a,%arg2%)
= emiaCo2EqTaxed(rsNat,a,%arg2%)*(1-v_redEmiaCo2Eq.1(rsNat,a,%arg2%))/gblScale * (xp.1(rsNat,a,%arg2%)+xTNest.1(rsNat,"fcost",a,%arg2%));

*
* --- linked to firm output
*
p_results(rsNat,nCO2,"xp",a,"dom",%arg1%, "%version%") $ xp.1(rsNat,a,%arg2%)
= emiIN(rsNat,nCO2,"",a,%arg2%)*(1-v_redEmiaCo2Eq.1(rsNat,a,%arg2%))/gblScale * (xp.1(rsNat,a,%arg2%)+xTNest.1(rsNat,"fcost",a,%arg2%));

```

Finally, the table with the market demands are expanded to comprise the new agent:

Demand, by product [1]

Regions: Romania Items: Quantity Origins: Total Years: 136

nnn\_GDP\_SSP2\_POP\_SSP2\_CC\_None\_CS\_ON

|                             | Output  | Total   | Household demand | Government demand | Intermediate demand | Investment demand | MACCs demand | Exp |
|-----------------------------|---------|---------|------------------|-------------------|---------------------|-------------------|--------------|-----|
| All commodities             | 1369.20 | 1635.43 | 323.40           | 71.64             | 717.43              | 182.08            | 14.53        |     |
| Grains and Crops            | 18.00   | 24.67   | 4.85             | 0.11              | 15.59               | 0.15              |              |     |
| Livestock and Meat Products | 5.84    | 8.15    | 3.95             | 0.01              | 3.91                | 0.07              |              |     |
| Mining and Extraction       | 7.50    | 18.90   | 1.56             | 0.12              | 15.89               | 0.45              |              |     |
| Processed Food              | 34.03   | 43.58   | 25.10            | 0.09              | 11.81               | 0.00              |              |     |
| Textiles and Clothing       | 20.91   | 36.87   | 12.03            | 0.00              | 15.32               | 0.01              |              |     |
| Light Manufacturing         | 135.60  | 186.62  | 21.45            | 0.98              | 94.78               | 15.76             |              |     |
| Heavy Manufacturing         | 303.10  | 443.89  | 26.19            | 3.07              | 176.08              | 47.68             |              |     |
| Utilities and Construction  | 227.46  | 230.41  | 31.00            | 0.14              | 87.31               | 90.10             | 14.53        |     |
| Transport and Communication | 308.30  | 319.34  | 100.03           | 7.71              | 167.56              | 15.37             |              |     |
| Other Services              | 308.46  | 323.00  | 97.25            | 59.40             | 129.38              | 12.49             |              |     |
| All commodities             | 1369.20 | 1635.43 | 323.40           | 71.64             | 717.43              | 182.08            | 14.53        |     |
| Grains and Crops            | 18.00   | 24.67   | 4.85             | 0.11              | 15.59               | 0.15              |              |     |
| Residuals                   | 0.01    | 0.01    |                  |                   | 0.01                |                   |              |     |

## References

Britz, W., Ciaian, P., Gocht, A., Kanellopoulos, A., Kremmydas, D., Müller, M., Petsakos, A., Reidsma, P. (2021): A design for a generic and modular bio-economic farm model, *Agricultural Systems* 191(June 2021)

Kiuiila, O., Markandya, A., & Ščasný, M. (2019). Taxing air pollutants and carbon individually or jointly: results from a CGE model enriched by an emission abatement sector. *Economic Systems Research*, 31(1), 21-43.

Kiuiila, O., & Rutherford, T. F. (2013). The cost of reducing CO2 emissions: Integrating abatement technologies into economic modeling. *Ecological Economics*, 87, 62-71.

### 3.10.3 Air pollution emissions

There is one equation  $e_{airpol}$  and related variable  $airPol$  to determine air pollution emissions<sup>34</sup> in the core model:

```

$iftheni.air_pol "%ModulesAir_Pollution%"=="on"

variable airPol(rNat,emisBst,t) "Air pollution emissions";
equation e_airPol(rNat,emisBst,t) "Air pollution emissions";

set aaAir(aa) "Armington agents with air pollution emission factors";

e_airPol(rsNat,emisBst,ts) $ (airPol.range(rsNat,emisBst,ts) ne 0) ..

    airPol(rsNat,emisBst,ts) = E=
    [
        --- link to final consumption
        sum((aaAir,i) $ ( (xaFlag(rsNat,aaAir,i) or xaFlag(rsNat,i,aaAir)) $ emiiA(rsNat,emisBst,i,aaAir)),
            mm_xa(rsNat,i,aaAir,ts) * emiiA(rsNat,emisBst,i,aaAir))
        --- link to firm output
        + sum(a $ (xpFlag(rsNat,a) $ emiiA(rsNat,emisBst,"",a)),
            (m_xp(rsNat,a,ts) + xtNest(rsNat,"fCost",a,ts) $ xtNest.l(rsNat,"fCost",a,ts))
            * emiiA(rsNat,emisBst,"",a))
        --- link to factor use
        + sum((a,f) $ (xfFlag(rsNat,f,a) $ emiiA(rsNat,emisBst,f,a)),
            xf(rsNat,f,a,ts) * emiiA(rsNat,emisBst,f,a))
    ] $ (airPol,'rsNat,emisBst,ts');

$endif.air_pol

```

<sup>34</sup> Chepeliev, M. (2020). *Development of the Air Pollution Database for the GTAP Data Base Version 10A* (No. 6163). Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University, GTAP research memorandum 33. <https://www.gtap.agecon.purdue.edu/resources/download/10042.pdf>

In opposite to the CO2 and non-CO2 emissions, air pollution emission factors are not differentiated between the domestic and imported origin. Similar to non-CO2 emissions, there are also pollution factors linked to factor use and production output.

The set *aaAir* makes sure that only the individual Armington agents are used in the equations, and not the potentially aggregated positions “armA” and “int”, see *cal\cal.gms*:

```
$$iftheni.air pol %ModulesAir_Polution%==on

aaAir(aaArm) = YES;
aaAir(aa)    = YES;
aaAir("armA") = NO;
aaAir("int")  = NO;
```

Air emissions are in the standard model post-triangular as they don’t show in any other equation. It is however possible to add equations in a shock file where air emissions impact, for instance, efficiency shifters to consider impacts of air emission on labour productivity.

#### 3.10.4 Reporting of energy use in Million tons of oil equivalents (MTOE)

The energy use accounting in Million tons of oil equivalents (*mtoe*) is solely done for reporting purposes and part of the post-model reporting code.

### 3.11 GTAP-E

GTAP-E is entirely based on the flexible nesting approach and is detailed there, see section “Flexible nesting.”

### 3.12 GTAP-AGR

GTAP-AGR is based on several changes in the GTAP in the standard layout and implemented in CGEBox by using the flexible nesting approach for the production function and the factor supply (see “*gams\gtapAgr\gtapAgr-Model.gms*”).

For livestock activities *lstk*, it introduces a sub-nest under the ND nests which allows for substitution between feedstock feed:

```
if ( card(feed),
    tNest("feed") = yes;
    tNest_i_a("feed", feed, lstk) = yes;
    tNest_n_a("ND", "feed", lstk) = YES;
);
```

The two sets are defined in “*gams\gtapAgr\gtapAgr\_def.gms*”:

```

set lstk0(i0) /
    ctl    "Cattle,sheep,goats,horses"
    oap    "Animal products nec"
    rmk    "Raw milk"
    wol    "Wool, silk-worm cocoons"
;/

set lstk(a); lstk(a) $ sum(mapa(lstk0,a),1) = YES;

set feed0(i0) /
    set.icrops
    cmt
    omt
    vol
    mil
    pcr
    sgr
    ofd
;/

set feed(i); feed(i) $ sum(mapi(feed0,i),1) = YES;

```

Intermediate demand substitution is also introduced in the food processing industries:

```

set foodPrc0(i0) / cmt,omt,vol,mil,pcr,sgr,ofd,b_t /;
set foodPrc(a); foodPrc(a) $ sum(mapa(foodPrc0,a),1) = YES;

if ( card(agr_c),
    tNest("foodPrc") = yes;
    tNest_i_a("foodPrc",agr_c,foodPrc) = yes;
    tNest_n_a("ND","foodPrc",foodPrc) = YES;
);

```

Factor supply to mobile factors is differentiated between agricultural and non-agricultural activities:

```

set agrFNest(f);
agrFNest(fm(cap)) $ sum(rnat, (p_omegaf(rnat, cap) ne inf)) = YES;
agrFNest(l) $ sum(rnat, (p_omegaf(rnat, l) ne inf)) = YES;

if ( card(agr),
    if ( card(agr) and card(agrFNest),
        fNest("agr") = yes;
        fNest_a_f("agr",agr,agrFNest) = yes;
        fNest_n_f("xft","agr",agrFNest) = yes;
        p_omegafNest(r,"agr",agrFNest) = inf;

        fNest("nonagr") = yes;
        fNest_a_f("nonagr",nonagr,agrFNest) = yes;
        fNest_n_f("xft","nonagr",agrFNest) = yes;
        p_omegafNest(r,"nonagr",agrFNest) = inf;
    );
);

```

The transformation elasticities are population weighted from the parameters:

```

*
* --- for the top nest (total labor/capital to agr/nonAgr) use the CET-elasticities provided by Hertel&Keeney
*
p_omegaf(rNat,fm) $ l(fm) = sum(roecd,p_factSupply(roecd,"labor")*roecd_r(roecd,rNat));

$$iftheni.cap not "%modulesCapVintages%"=="on"
p_omegaf(rNat,f) $ cap(f) = sum(roecd,p_factSupply(roecd,"capital")*roecd_r(roecd,rNat));
$$endif.cap

p_omegaf(rNat,fm) $ (lnd(fm) and (not sum(sameas(fNest,"cropLand"),1))) = sum(roecd, p_factSupply(roecd,"land")*roecd_r(roecd,rNat));
p_omegaf(rNat,fm) $ mapf("water",fm) = sum(roecd, p_factSupply(roecd,"land")*roecd_r(roecd,rNat));

```

The parameterization is based on parameters by Keeney and Hertel, details do not matter here, the code in “*gams\gtapAgr\gtapAgr\_def.gms*” should be rather self-explanatory.



### 3.13 Recursive-dynamics

#### 3.13.1 Overview

The model can be run in recursive-dynamic mode. Compared to the comparative-static case, there are four major changes:

- The capital stock in each period is last year's capital stock plus investments minus depreciation.
- If the vintage module is active, the share parameters are updated in each period, see "Dynamic updating of share parameters".
- Share parameters in factor supply are updated as discussed next.
- The G-RDEM model can be added.

#### 3.13.2 Dynamic updates of share parameters in factor supply

Share parameters in factor supply nests are defined at the benchmark based on the observed allocation across production activities and sub-regions. The model might produce implausible differences in wages, land or capital rents if the share parameters are kept unchanged over time. Consider, for instance, a model set-up where labor markets are segmented into an agricultural and non-agricultural block with limited mobility and structural change implying dropping demand for agricultural output. As a consequence, the agricultural segment demands less labor and the wage for labor employed in agriculture drops while it increases in the rest of the economy. With each year, these wage differences will grow. In reality, new entrants in the labor force will observe the shrinking demand in agriculture and less of them will opt for training in agriculture. One might implement this by splitting up the factor supply of labor similar to the capital vintages into new entrants and the already trained labour force and differentiate their transformation elasticities. Similar to the vintage module, a share of the old labour would retire from the labour force in each year while previous year entrants would be added to the old stock. Share parameters would need to be updated as a weighted average of the entrants of last year and the remaining old stock. Similar for capital if no capital vintages are present.

CGEBox implements a reduced form solution working in the dual domain instead, for both capital and labor. Instead of explicitly having the two differentiated in new and old, emerging price differences between wage and capital rents are reduced in each year. To do so, we first define the factors subject to this mechanism:

```
option kill=curf;  
curF(cap) $ sum(rs, p_omegaf(rs, cap) ne inf) = yes;  
curF(l)   $ sum(rs, p_omegaf(rs, l) ne inf) = yes;  
curF(wat) $ sum(rs, p_omegaf(rs, wat) ne inf) = yes;  
curF(cap) $ sameas(cap, "oldCap") = no;
```

As seen, these are capital, labor and water if they are not depicted as fully mobile. The code first defines an update factor, temporarily stored on xft.m. For capital and water, it is based on last period's depreciation rate times the number of years from the last to the current period:

```
xft.m(rs, cap, tSim) = sum(r_r(rs, rsNat), p_depr(rsNat, tSim-1)) * p_timeStep(tSim-1);  
xft.m(rs, wat, tSim) = sum(r_r(rs, rsNat), p_depr(rsNat, tSim-1)) * p_timeStep(tSim-1);
```

For labor, we assume an average of 40 active years and add 2% additional mobility for retraining:

```

--- assume 40 years of active working years, plus 2% retraining additional mobility
xft.m(rs,l,tSim) = (1/40 + 0.02) *p_timeStep(tSim-1);

```

The factors are increased if the actual change in the factor stock between the last two simulated periods is higher, such as in case of high rates of nest investments:

```

xft.m(rs,curF,tSim) $ xft.l(rs,curF,tSim-2)
= max(xft.m(rs,curF,tSim),xft.m(rs,curF,tSim)+(xft.l(rs,curF,tSim-1)-xft.l(rs,curF,tSim-2))/xft.l(rs,curF,tSim-2));

```

The expected factor prices for an activity are then defined as a weighted average of the activity specific price in the last and the economy-wide average price:

```

pf.m(xfFlag(rs,curF,a),tSim) $ (p_omegaf(rs,curF) ne inf)
= pf.l(rs,curF,a,tSim-1) * (1-xft.m(rs,curF,tSim))
+ pft.l(rs,curF,tSim-1) * xft.m(rs,curF,tSim);

```

Based on these updated prices, we define expected prices for factor supply nests:

```

*
* --- define results exhaustion prices for nests and the economy wide price
*
pfNest.m(rs,curF,fNest,tSim) $ xfNest.l(rs,curf,fNest,tSim-1)
= sum(fNest_a_f(fNest,a,curF), pf.m(rs,curF,a,tSim)*xf.l(rs,curF,a,tSim-1))/xfNest.l(rs,curf,fNest,tSim-1);

```

And the new economy wide price:

```

pft.m(rs,curF,tSim) $ xft.l(rs,curF,tSim-1)
= sum(xfFlag(rs,curF,a), xf.l(rs,curF,a,tSim-1)*pf.m(rs,curF,a,tSim))/xft.l(rs,curF,tSim-1);

pft.m(rs,curF,tSim) $ sum(fNest_n_f("xft",fNest,curF),xfNest.l(rs,curF,fNest,tSim-1))
= sum(fNest_n_f("xft",fNest,curF), xfNest.l(rs,curF,fNest,tSim-1)*pfNest.m(rs,curF,fNest,tSim))/xft.l(rs,curF,tSim-1);

```

The assumption is hence that if there would be shock in one period change factor price relation and no changes ever after, flexibility in deciding to which sectors to allocate new investments and new entrants in the labor force would remove wage and capital rent differences in the long run.

Based on the updated prices, we re-define the share parameters:

```

*
* --- update factor supply revenue shares
*
p_gf(rs,curF,a,tSim) $ (xfFlag(rs,curF,a)
and sum(fNest_a_f(fNest,a,curF) $ (p_omegafNest(rs,fNest,curF) ne inf), xfNest.l(rs,curF,fNest,tSim-1)))
= sum(fNest_a_f(fNest,a,curF),
(xf.l(rs,curF,a,tSim-1)/xfNest.l(rs,curF,fNest,tSim-1))*(pfNest.m(rs,curF,fNest,tSim)/pf.m(rs,curF,a,tSim))**p_omegafNest(rs,fNest,curF));

```

For the example above, the new lower price differences between agricultural and economy-wide ages, the share parameters  $p_{gf}$  would be reduce.

### 3.14 G-REDM: Long-term baseline generation and analysis

This section is prepared by Roberto and Wolfgang Britz, drawing on: G-REDM: A GTAP - Based Recursive Dynamic CGE model for Long Term Baseline Generation and Analysis, in: Bocconi Working Paper Series No. 105, April 2018. The working paper was in revised form published as Britz,

W., Roson, R. (2019): G-RDEM: A GTAP-Based Recursive Dynamic CGE Model for Long-Term Baseline Generation and Analysis, *Journal of Global Economic Analysis* 4(1): 50-96. Since then, G-RDEM had been extended by additional elements: expenditure shares of the government and of investors can be rendered a function of income per capita changes, drawing on econometric work, crop yields and crop land expansions can be driven by exogenous projections, and calorie intakes per calories controlled based on an empirical relation to income per capita (Britz, W. (2020): Maintaining plausible calorie intakes, crop yields and crop land expansion in long-run simulations with Computable General Equilibrium Models. Discussion Paper 2020:2, Institute for Food and Resource Economics, University Bonn). Finally, the AIDADS demand system was extended to a MAIDADS system and re-estimated (Britz, W. (2021): Estimating a global MAIDADS demand system considering demography, climate and norms, Bio-based and Applied Economics, pre-Journal print available online).

### **3.14.1 Summary**

The G-RDEM module allows the user to construct long-run baselines drawing on the so-called Socio-Economic Pathways (SSPs) developed for the International Panel on Climate Change (IPCC), adding model mechanisms important for long-term structural change. These baselines can be subsequently used for counterfactual analysis. The necessary data for the SSPs are distributed with the model code and need not to be supplied by the user. The different mechanisms to consider in the G-RDEM model extension as well as the period to cover and the time resolution can be steered via the interface.

### **3.14.2 Background and introduction**

Due to issues such as climate change and depletion of global resources, there is an increasing demand for long-term quantitative analyses. Computable General Equilibrium models can contribute in that direction as they consistently consider the manifold interrelations occurring in the economy, while providing the often-needed sectoral detail. They therefore complement approaches working at the more aggregate level (e.g. Dellink et al. 2017<sup>35</sup>) or focusing in detail on specific sectors (e.g. Alexandratos and Bruinsma 2012<sup>36</sup> for agriculture). On the other hand, it should be noted that CGE models were not originally designed to this purpose, but rather for short-term policy assessment, like simulating the effects of a fiscal reform, or the implementation of a trade agreement. Accordingly, most parameters in CGE models are usually “calibrated” to a relatively recent Social Accounting Matrix or Input Output Table, such that the observed structure of an economic system is taken as a benchmark, from which counterfactual experiments are conducted.

Of course, when the economy is analysed at a horizon of 20, 30 years, or even more, the economic structure as emerging from some past national accounts, which may refer to five years back, is no more a valid starting point. One should consider trends in structural adjustment, driven by changing preferences, demographic composition, new technologies, variations in the endowments of primary resources (including human capital), etc. The whole issue is not about forecasting: nobody actually knows which “breakthrough” technologies could emerge, or which unexpected phenomena could shape the economic structure in the future. What we do know from past observation is that a number

---

<sup>35</sup> Dellink, R., Chateau, J., Lanzi, E., & Magné, B. (2017). Long-term economic growth projections in the Shared Socioeconomic Pathways. *Global Environmental Change*, 42, 200-214

<sup>36</sup> Alexandratos, N., & Bruinsma, J. (2012). World agriculture towards 2030/2050: the 2012 revision (Vol. 12, No. 3). FAO, Rome: ESA Working paper

of “slow” adjustment processes are active and therefore they should be taken into account in the generation of a credible and internally consistent future baseline.

The study of time evolution of the economic structure (“structural change”) is a rather active research field in theoretical and applied economics (Matsuyama, 2008<sup>37</sup>). Most of the studies in the literature, however, look at the past. Typical research questions are: the contribution of the changing industrial mix to aggregate productivity (e.g., Duarte and Restuccia 2010<sup>38</sup>); the declining share of the agricultural sector in developing economies (e.g., Üngör 2013<sup>39</sup>), etc., where some specific transition processes are identified. With the G-RDEM module, rather than studying the past, we aim at drawing from some empirical findings and methodologies in this literature to infer, inside a CGE modelling framework, a possible future evolution of the economy.

To this end, a number of “unconventional” features have to be introduced into the standard CGE formulation, to create a model specifically designed for the generation and assessment of long-term economic scenarios. We present therefore in here a newly develop CGE model of this kind, termed G-RDEM (GTAP based Recursive Dynamic Economic Model). This model considers drivers of long run structural change, which we regard as especially relevant, namely: (1) non-linear Engel curves in household consumption, (2) productivity growth differentiated by sector, (3) debt accumulation from foreign savings and trade imbalances, (4) aggregate saving rates linked to population and income dynamics, and (5) time-varying and income dependent industrial cost shares.

G-RDEM extends the flexible and modular CGE modelling platform CGEBox, from which it inherits some important features. Firstly, the code is open source, to ensure transparency and invite the community of modellers to use the tool and contribute to its further development. Secondly, it maintains full flexibility in sectoral and regional aggregation. Thirdly, G-RDEM as a seamless integrated module in CGEBox offers the possibility to combine it with other modules such as CO2 and Non-CO2 emissions, GTAP-Water, GTAP-AEZ etc.. All new features are based on econometrically estimated parameters, thereby making the implementation fully documented and transparent. G-RDEM is encoded in the GAMS modelling language and, as a module of CGEBox, shares its graphical user interface

### 3.14.3 Overview

The construction of a long-term baseline in CGE models typically draws on population and GDP projections from other studies. Indeed, a standard recursive dynamic CGE only considers capital accumulation as an endogenous mechanism driving growth, while productivity changes and other drivers of structural change are usually kept exogenous. In order to let a CGE model replicate a given growth path, a total factor productivity shifter is endogenously determined during the construction of the reference baseline, by fixing GDP at each time period. In subsequent model runs and counterfactual simulations, productivity parameters are then maintained at those levels determined during baseline construction, while national income is endogenously computed.

---

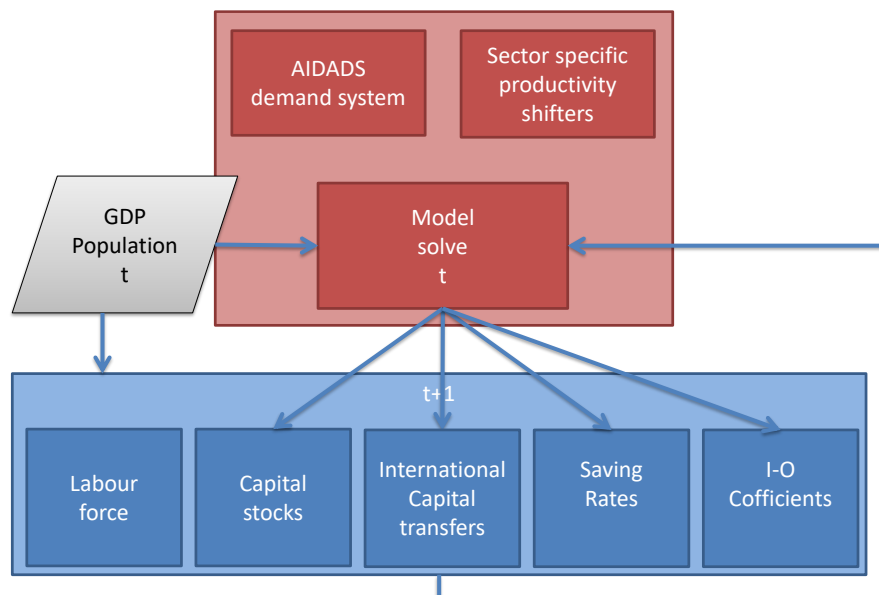
<sup>37</sup> Matsuyama, K. (2008). Structural change. In S. Durlauf, & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics* (2 ed.). Palgrave-Macmillan

<sup>38</sup> Duarte, M., Restuccia, D., (2010). The role of the structural transformation in aggregate productivity. *The Quarterly Journal of Economics* 125, 129-173

<sup>39</sup> Üngör, M., (2013). De-agriculturalization as a result of productivity growth in agriculture. *Economics Letters* 119, 141-145

This simple methodology aligns the output of the CGE model to a pre-determined aggregate growth path, but of course does not capture some fundamental structural changes which may take place in the economy, i.e. in the composition of output and demand. Instead, we aim in here to address the key elements driving such compositional change (Figure 7).

To this end, we introduce an MAIDADS demand system to consider how budget shares in household consumption adjust to the changing levels of per capita income, to capture “non-linear Engel curves”, which are a salient feature of economic development. Secondly, the economy wide total factor productivity (TFP) shifter, aligning the model to the target GDP in any period, is here differentiated by sector. These two features are introduced through specific equations directly into the CGE framework itself (red boxes in Figure 1). Other elements are activated in between the solution points (blue boxes). Therefore, the intra-period equilibrium computed by the model, in combination with exogenous projections from the current period  $t$ , updates some parameters for the following period  $t+1$ . The labour force (by skill category) is adjusted to population and work force projections. Next year’s capital stocks reflect last year’s ending stocks and gross investments. International capital transfers reflect past foreign savings. Saving rates adjust to population and GDP growth, and I-O coefficients (factor shares in production processes) are updated on the basis of national income.



**Figure 7: Overview of the recursive-dynamic modelling framework G-RDEM**

The process thus requires some exogenous projections for GDP and population. G-RDEM offers the possibility to draw on a set of projections for the so-called SSPs (Shared Socio Economic Pathways) (Riahi et al. 2016<sup>40</sup>), available online from the IIASA Shared Socio Economic Pathways Database.<sup>41</sup> These SSPs were developed in the context of the IPCC scientific assessment on Climate Change. For each of these five SSPs, a single population and urbanization scenario, jointly developed by IIASA and NCAR (National Center for Atmospheric Research), can be combined with GDP projections from either the OECD or IIASA. These GDP and population projections are available in 5-year steps up to 2100, at a single country basis. They are aggregated in G-RDEM to the desired regional aggregation

<sup>40</sup> Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O’neill, B. C., Fujimori, S., & Lutz, W. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Global Environmental Change*, 42, 153-168

<sup>41</sup> <https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=about>.

level and interpolated to yield yearly time series. They can also be complemented by Climate Change impacts on yields for a set of RCPs (Representative Concentration Pathways) and various combinations of GCMs (Global Circulation Models) and global gridded growth models provided by the AgCLIM50 project (van Meijl et al. 2017<sup>42</sup>).

A user might also add its own scenario assumptions during the construction of the baseline, such as about trade policies or alternative GDP projections. After the definition of the baseline, the software saves the resulting productivity shifters and other variables, which can subsequently be loaded as exogenous parameters for counterfactual analysis. The set of results from the baseline can also be directly employed, to get a much-disaggregated definition of the economic scenario.

### **3.14.4 An AIDADS demand system with detail for food consumption**

#### **Background**

It is universally acknowledged that the relationship between consumption level and income (also known as Engel curve) can be complex and non-linear. Yet, many CGE models still adopt demand systems such as Cobb-Douglas (CD) or Linear Expenditure System (LES), having linear Engel curves. Those simplifying assumptions make the model easier to handle, but are defensible only if the model is used for simulations involving limited changes in income levels. Of course, this is not the case for long-term analyses. Keeping constant marginal budget shares would lead to an overestimation of the demand for necessities, such as food, while demand in other sectors will hence be underestimated. The consequences are implausible long-run structural changes in production, demand, and trade patterns. Some models employed for long-term analysis therefore use different demand systems and/or re-parameterize along the dynamic path. For instance, MAGNET (Woltjer et al. 2014, p. 84<sup>43</sup>) incorporates a module for re-calibrating the parameters of a CDE (Constant Differences in Elasticity) demand system to given income elasticities. Nonetheless, the authors admit: “All of these parameters and functional forms are very much ad hoc, and should be improved.”.

Following Roson and van der Mensbrugghe 2018<sup>44</sup>, we rather implement an empirically estimated AIDADS demand system into the G-RDEM model, for broad product groups. The AIDADS is An Implicit, Directly Additive Demand System (Rimmer and Powell 1996<sup>45</sup>). It can be understood as a generalization of a LES demand system, where marginal budget shares are not fixed, but are a linear combination of two vectors, depicting the marginal budget structure at very low and very high utility (income) levels. Given that the marginal budget shares in the two vectors fulfil the adding up condition to unity, any linear combination of the two vectors also leads to regular budget shares. In order to improve the detail inside the agri-food sector, we also took econometric estimates of income

---

<sup>42</sup> van Meijl, H., Havlik, P., Lotze-Campen, H., Stehfest, E., Witzke, P., Dominguez, I.P., Bodirsky, B., van Dijk, M., Doelman, J., Fellmann, T. and Humpenoeder, F., 2017. Challenges of Global Agriculture in a Climate Change Context by 2050 (AgCLIM50). JRC Science for Policy Report, EUR 28649 EN, doi:10.2760/7724

<sup>43</sup> Woltjer, G. B., Kuiper, M., Kavallari, A., van Meijl, H., Powell, J. P., Rutten, M. M., ... & Tabeau, A. A. (2014). The MAGNET model: Module description (No. 14-57). LEI Wageningen UR

<sup>44</sup> Roson R, & van der Mensbrugghe D. (2018), Demand-Driven Structural Change in Applied General Equilibrium Models, in F. Perali and P. L. Scandizzo (eds.), The New Generation of Computable General Equilibrium Models, Springer International, doi: /10.1007/978-3-319-58533-8\_2

<sup>45</sup> Rimmer, M. T., & Powell, A. A. (1996). An implicitly additive demand system. *Applied Economics*, 28(12), 1613-1622

dependent marginal expenditure shares for food categories from Muhammad et al. 2011<sup>46</sup>, and incorporate them in the AIDADS framework.

Cranfield et al. 2000<sup>47</sup> improve on the original Rimmer and Powell 1996 approach, by developing an estimation method that does not rely on an approximation of utility. We follow their notation in the following. The demand system is defined below. Equation (1) determines the Marshallian demand, which is similar to that of a LES. Here, however, the marginal budget shares  $\delta_i$  are endogenous variables, defined by (2), expressed as a linear combination of two vectors  $\alpha$  and  $\beta$ , function of the utility level  $u$ , implicitly defined by (3).

$$x_i = \bar{\gamma}_i + \frac{\delta_i}{p_i} \left[ Y - \sum_j \bar{\gamma}_j p_j \right] \quad (1)$$

$$\delta_i = \frac{\bar{\alpha}_i + \bar{\beta}_i u}{1 + u} \quad (2)$$

$$\ln(u) = \sum_i \delta_i \ln(x_i - \bar{\gamma}_i) + \bar{\kappa} \quad (3)$$

### Estimation and integration into modelling framework

We first econometrically estimated  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $u$  using data from the International Comparison Program ICP 2015<sup>48</sup>, for ten broader expenditure categories (food, beverages and tobacco, clothing, housing, furniture, transportation, recreation, communication, health, education) plus details for food expenditures. The integration in the CGE model requires mapping the parameter estimates to the commodity resolution of the model.

The demand system is calibrated against the benchmark data of regional household consumption, from the GTAP v.9 data set. To this purpose, we regressed the utility levels  $u$  from our findings to total per capita consumption expenditure  $Y$  in each region. That allows us to estimate (from (2)) the marginal budget shares  $\delta_i$  in the calibration point. We then discarded the previously estimated  $\gamma$  and instead solve (1) for  $\gamma$  at given  $x$ ,  $Y$ ,  $p$  and the calibrated marginal budget shares. In the case that this implies a negative  $\gamma$ , we use a penalty minimization approach, which minimizes the difference between the estimated  $\alpha$ ,  $\beta$  and the “corrected” ones, such that all  $\gamma$  turn out to be positive.

---

<sup>46</sup> Muhammad, A., Seale, J. L., Meade, B., & Regmi, A. (2011). International evidence on food consumption patterns: an update using 2005 international comparison program data. (March 1, 2011). USDA-ERS Technical Bulletin No. 1929. Available at SSRN: <https://ssrn.com/abstract=2114337> or <http://dx.doi.org/10.2139/ssrn.2114337>

<sup>47</sup> Cranfield, J. A., Preckel, P. V., Eales, J. S., & Hertel, T. W. (2000). On the estimation of an implicitly additive demand system'. *Applied Economics*, 32(15), 1907-1915

<sup>48</sup> ICP (2015). Purchasing Power Parities and Real Expenditures of World Economics: A Comprehensive Report of the 2011 International Comparison Program. World Bank, Washington, D.C.

## Technical implementation of the MAIDADS demand system

Note that the equations as discussed in here refer to the updated estimation by Britz 2021<sup>49</sup> who uses an extended data set from IPC with more agro-food detail and introduces additional controls.

### Equations

Due to its resemblance with the LES system, which was already available in CGEBox, the code is a direct extension of the existing code both for the calibration to the benchmark and implementation of the demand system in the model's equation (see model\cal\_les.gms and model\dem\_les.gms).

The demand equations for both systems are indeed identical:

```
*
* --- demand equations for individual products (if not not under sub-nest, this is checked by flag p_hasXA)
*
e _xacles(rs,i,h,ts) $ p_hasXA(rs,i,h) ..
    xa(rs,i,h,ts) /xa.scale(rs,i,h,ts)
    =e= (
        $$iftheni.baseline "%fixCalories%"=="on"
            gammaLES(rs,i,h,ts)
        $$else.baseline
            gammaLESM(rs,i,h,ts)
        $$endif.baseline
        /1000*pop(rs,h,ts) + alphaLES(rs,i,h,ts)/m_pa(rs,i,h,ts) * yCNonCom(rs,h,ts))/xa.scale(rs,i,h,ts);
```

The demand is equal to a constant term gammaLES respectively gammaLESM per head multiplied with population in the household type h (which might be just one) plus the marginal budget share alphaLES times the non-committed income yCNonCom, divided by the Armington price index for the private household aggregate, defined in the macro m\_pa.

The non-committed income yNonCom is the difference between the given expenditures for final consumption yc and the value of the commitments, i.e. the constant terms multiplied with the Armington price:

```
*
* --- non committed income: total expenditure for private consumption minus
* value of commitments. Note that top-level nests (only) of the nested
* demand system are taken into account. Lower nests are all CES-functions
*
e_yCNonCom(rs,h,ts) $ sum(i,xaFlag(rs,i,h)) ..
    ycNonCom(rs,h,ts)/yc.scale(rs,h,ts)
    =E= (yc(rs,h,ts)
*
* --- note: commitment terms are defined in a per capita basis
*
    $$iftheni.baseline "%fixCalories%"=="on"
        - sum(i $ p_hasXA(rs,i,h), gammaLES(rs,i,h,ts) *pop(rs,h,ts)/1000 * m_pa(rs,i,h,ts))
        - sum(dNest_n_fd("top",dNest,h) $ p_hasXA(rs,dNest,h), gammaLES(rs,dNest,h,ts) *pop(rs,h,ts)/1000 * pdNest(rs,dNest,h,ts))/yc.scale(rs,h,ts);
    $$else.baseline
        - sum(i $ p_hasXA(rs,i,h), gammaLESM(rs,i,h,ts) *pop(rs,h,ts)/1000 * m_pa(rs,i,h,ts))
        - sum(dNest_n_fd("top",dNest,h) $ p_hasXA(rs,dNest,h), gammaLESM(rs,dNest,h,ts) *pop(rs,h,ts)/1000 * pdNest(rs,dNest,h,ts))/yc.scale(rs,h,ts);
    $$endif.baseline
```

Note that we allow for demand nests dNest which are CES aggregates of other such nests or individual commodity demands.

In the case of AIDADS, the marginal budget shares alphaLES are not fixed (as in the LES), but a function of utility uh:

<sup>49</sup> Britz, W. (2021): Estimating a global MAIDADS demand system considering demography, climate and norms, Bio-based and Applied Economics, in: online available 19 July 2021



```

*
* --- AIDADS: marginal budget shares are a function of utility (individual products or nests)
*
e_alphaAIDADS(rs,i,h,ts(tRun)) $ (xaFlag(rs,i,h) $ (alphaLES.range(rs,i,h,tRun) ne 0) $ p_hasXA(rs,i,h)
$ (alphaAIDADS.l(rs,i,h) + betaAIDADS.l(rs,i,h))
$ (p_demSystem eq AIDADS) $ hFlag(rs,h) $ hCur(h))..

alphaLES(rs,i,h,tRun)/(alphaAIDADS.l(rs,i,h) + betaAIDADS.l(rs,i,h) + 1.E-6)
=E= { [ alphaAIDADS(rs,i,h) + betaAIDADS(rs,i,h) * sqExp(uhExpD(rs,h,tRun),10) ]
/ [ 1 + sqExp(uhExpD(rs,h,tRun),10) ]

*
* --- shifter terms to meet calorie intake
*
+ [alphaLES.l(rs,i,h,tRun-1) $ (alphaLES.l(rs,i,h,tRun-1) gt 1.E-8)]
* ( alphaShiftF(rs,h,tRun) $ foodCal(i)
+ alphaShiftNF(rs,h,tRun) $ (not foodCal(i))) $ (not sameas(tRun,"%t0%"))
}/(alphaAIDADS.l(rs,i,h) + betaAIDADS.l(rs,i,h) + 1.E-6);

```

The two vectors `alphaAIDADS` and `betaAIDADS` are the limiting cases for the marginal budget shares at very low and very high utility levels. The utility of the (representative) private household is set to unity in the benchmark. In order to line up with the empirical implementation (see next section), the parameter `p_uh0` is introduced.

All `alphaAIDADS` might change endogenously according to a uniform shifter `alphaShiftF` for food product and `alphaShiftNF` for non-food items, which have the opposite sign. These shifters are driven by fixing calorie intakes as discussed below.

Cardinal utility is thus defined as follows:

```

*
* --- indirect (and implicate) utility definition in MAIDADS
*
e_uhAIDADS(rs,h,ts) $ (sum(i,xaFlag(rs,i,h)) $ (p_demSystem eq AIDADS)) ..

uh(rs,h,ts) =E= sqexp(uh1(rs,h,ts),10)/p_uh0(rs,h);

e_uh1AIDADS(rs,h,ts) $ (sum(i,xaFlag(rs,i,h)) $ (p_demSystem eq AIDADS)) ..

uh1(rs,h,ts) =E= log(ycNonCom(rs,h,ts)/pop(rs,h,ts)*1000)
+ sum[i $ (p_hasXA(rs,i,h) $ (alphaAIDADS.l(rs,i,h) + betaAIDADS.l(rs,i,h))),
alphaLES(rs,i,h,ts) * log(alphaLES(rs,i,h,ts)/m_pa(rs,i,h,ts))]
+ sum[dNest $ (p_hasXA(rs,dNest,h) $ (alphaAIDADS.l(rs,dNest,h) + betaAIDADS.l(rs,dNest,h))),
alphaLES(rs,dNest,h,ts) * log(alphaLES(rs,dNest,h,ts)/pdNest(rs,dNest,h,ts))]
- v_kappa.l(rs,h,"U");

```

where `p_uh0` is an additional parameter in the demand system (corresponding to the parameter `A` in Cranfield et al. 2000, see their equation (10)). Note the equivalence of the utility definition to the one in the LES system: only consumption quantities exceeding the commitments `gammaLES` generate utility.

The cardinal utility is mapped into a driver of the logistic functions of the marginal budget shares and the commitment terms:

```

*
* --- utility level is transformed of a driver of the logistic functions for the marginal
* budget shares
*
e_uhExpG(rs,h,ts) $ (sum(i,xaFlag(rs,i,h)) $ (p_demSystem eq AIDADS) $ hFlag(rs,h) $ hCur(h)) ..

(uhExpG(rs,h,ts)) =E= (p_parAIDADS("UG","cnst") + v_kappa(rs,h,"UGC") + (v_kappa(rs,h,"UGL") + p_parAIDADS("UG","lin")) * uh1(rs,h,ts));

*
* --- utility level is transformed of a driver of the logistic functions of the commitment terms
*
e_uhExpD(rs,h,ts) $ (sum(i,xaFlag(rs,i,h)) $ (p_demSystem eq AIDADS) $ hFlag(rs,h) $ hCur(h)) ..

(uhExpD(rs,h,ts)) =E= (p_parAIDADS("UD","cnst") + v_kappa(rs,h,"UDC") + (v_kappa(rs,h,"UDL") + p_parAIDADS("UD","lin")) * uh1(rs,h,ts));

```

We also estimate the elasticity of private expenditure with regard to private utility:

```

e_phiPLES(rsNat,ts) $ (pkip.range(rsNat,ts) ne 0) ..

phiP(rsNat,ts) =e=

sum( r_r(rs,rsNat) $ sum(h $ hFlag(rs,h),1),
*
*   --- population weighted 1/[ du/dm]  from indirect utility function
*
sum(h $ hFlag(rs,h), pop(rs,h,ts) * prod(i $ (p_hasXa(rs,i,h) $ alphaLES.1(rs,i,h,"%t0%")),
(m_pa(rs,i,h,ts)/alphaLES(rs,i,h,ts))**alphaLES(rs,i,h,ts))
* prod(dNest_n fd("top",dNest,h) $ (p_hasXa(rs,dNest,h) $ alphaLES.1(rs,dNest,h,"%t0%")),
(pdNest(rs,dNest,h,ts)/alphaLES(rs,dNest,h,ts))**alphaLES(rs,dNest,h,ts))
/sum(h $ hFlag(rs,h), pop.1(rs,h,"%t0%") * prod(i $ (p_hasXa(rs,i,h) $ alphaLES.1(rs,i,h,"%t0%")),
(pa.1(rs,i,h,"%t0%)/alphaLES.1(rs,i,h,"%t0%"))**alphaLES.1(rs,i,h,"%t0%")),
* prod(dNest_n fd("top",dNest,h) $ (p_hasXa(rs,dNest,h) $ alphaLES.1(rs,dNest,h,"%t0%")),
(pdNest.1(rs,dNest,h,"%t0%)/alphaLES.1(rs,dNest,h,"%t0%"))**alphaLES.1(rs,dNest,h,"%t0%"))))
*
*   --- conversion to elasticity
*
*   * sum(h $ hFlag(rs,h), pop(rs,h,ts) * uh(rs,h,ts))/sum(h $ hFlag(rs,h), pop(rs,h,ts) * yc(rs,h,ts))
*
*   --- and correct for benchmark relation and weight with population
*
*   * sum(h $ hFlag(rs,h), pop.1(rs,h,"%t0%") * yc.1(rs,h,"%t0%"))
*   /sum(h $ hFlag(rs,h), pop.1(rs,h,"%t0%") * uh.1(rs,h,"%t0%"))
*
)
/sum((r_r(rs,rsNat)) $ sum(h,hFlag(rs,h),1);

```

### Controlling for calorie intakes as a function of per capita income

In order to arrive at a plausible development of calorie intakes, an empirical estimated relation between calories and per-capita income is used,  $Y$  is per capita income in USD and  $T$  is mean temperature:

$$calTot^* = -947,275 - 16,257T + 738,910 \ln(Y) - 30,083 \ln(Y)^2 \quad (60)$$

The integration into the modelling framework needs to consider the error at the benchmark. Therefore, only changes in calories equation (60) from changes in income are considered in the second part of equation. There are two additional elements in the equations: (1) The first block defines a weighted average between the equation (60) and its first differences in the second block, such that after two hundredth years, the simulation would be identical to the empirically estimated relations. (2) the actual intake and the estimate at the benchmark is multiplied with the first differences, such that countries which an above average intake will also tend to expand their calorie stronger.

```

e_caloriesShift(rs,h,ts(tRun)) $ (hFlag(rs,h) $ (not sameas(ts,"%t0%")) $ (alphaShiftF.range(rs,h,tRun) ne 0) ) ..

v_nutrients(rs,h,"cal",ts)/v_nutrients.scale(rs,h,"cal",ts) =E=

[
  $$iftheni.dyn not "%dynMode%"=="Comparative static"
*
*   --- draw the calorie intakes in year t0+200 to the estimates
*
  [1-( 201-sum(sameas(t,tRun),t.pos))/200]*
  {p_calCnst(rs) + 738.910 * log( yc(rs,h,tRun)/(pop(rs,h,tRun)*pCons(rs,h,tRun))*1000)
    - 30.083 * sqr(log( yc(rs,h,tRun)/(pop(rs,h,tRun)*pCons(rs,h,tRun))*1000))}
*
*   --- second term updates estimates depending on difference between benchmark results
*   and estimates
*
  + [(201-sum(sameas(t,tRun),t.pos))/200]*
  $$endif.dyn

```

```

{
*
*   --- last year's intake
*
*   v_nutrients.l(rs,h,"cal",tRun-1)
*
*   --- add to last year the change implied by the income increase
*   (income dependent betas from regression)
*
*   + ( 738.910 * [ log( yc(rs,h,tRun) / (pop(rs,h,tRun) * pCons(rs,h,tRun) * 1000)
*     - log( yc.l(rs,h,tRun-1) / (pop.l(rs,h,tRun-1) * pCons.l(rs,h,tRun-1) * 1000) ]
*     - 30.083 * [ sqrt(log( yc(rs,h,tRun) / (pop(rs,h,tRun) * pCons(rs,h,tRun) * 1000) )
*     - sqrt(log( yc.l(rs,h,tRun-1) / (pop.l(rs,h,tRun-1) * pCons.l(rs,h,tRun-1) * 1000) ) )
*
*   --- in case of benchmark intake below estimate, dampend by the relation between the actual intake
*   and estimates
*
*   * min(10.0, (v_nutrients.l(rs,h,"cal", "%t0%") / {p_calCnst(rs) + 738.910 * log( yc.l(rs,h, "%t0%") / pop.l(rs,h, "%t0%") * 1000)
*     - 30.083 * sqrt(log( yc.l(rs,h, "%t0%") / pop.l(rs,h, "%t0%") * 1000) ) } )
*
* }
] / v_nutrients.scale(rs,h,"cal",ts);

```

Which can be depicted in mathematical notation as:

$$\begin{aligned}
 calTot_t = & \left(1 - \frac{201-t}{200}\right) calTot_t^* \\
 & + \frac{201-t}{200} \left( calTot_{t-1}^* + \frac{calTot_0}{calTot_0^*} [calTot_t^* - calTot_{t-1}^*] \right)
 \end{aligned} \tag{61}$$

As the RHS of this equation is exogenous if income is fixed during baseline construction, an endogenous variable need to adjust. This is the shifter *alphaShiftF* for food products in the *e\_alphaAidAds* equation. It will be added to the marginal budget share as driven by the cardinal utility. As the budget shares need to add up to unity, a positive/negative shifter for food products (which is equivalent of increasing/decreasing the calorie intake) requires a shift of the marginal budget shares for non-food products in the opposite direction.

The correction factors for the marginal budget shares are integrated as follows where the endogenous correction factor *cor* differs for food *f* and non-food items *nf* and are defined relative to last year's marginal budget share, i.e. the corrections are defined in relative terms:

$$\delta_{i,t}^* = \delta_{i,t} + \delta_{i,t-1}^* \left( cor_f \wedge i \in f + cor_{nf} \wedge i \in nf \right) \tag{62}$$

For food items, *cor<sub>f</sub>* it is implicitly determined by the calorie constraints (61). For non-food items, *cor<sub>nf</sub>* is driven by the adding up conditions for marginal budget shares:

$$\sum_i \delta_{i,t-1}^* \left( cor_f \wedge i \in f + cor_{nf} \wedge i \in nf \right) = 0 \tag{63}$$

The correction for the commitment terms uses last year's per capita demands as weights. Using large year's marginal budget shares in (62) and last year's per capita demand in (63) should ensure that the provoked changes do not fundamentally change the composition inside of the food and inside of the non-food commodity bundle from what the benchmark step suggests.

The correction is also applied to the commitment terms:

```

*
* --- MAIDADS: commitments are a function of utility (individual products or nests)
*
e_gammaAIDADS(rs,i,h,ts(tRun)) $ ( (p_demSystem eq AIDADS) $ (gammaLESM.range(rs,i,h,tRun) ne 0)
$ p_hasXa(rs,i,h) $ hFlag(rs,h) $ hCur(h) )..

gammaLESM(rs,i,h,tRun)/(abs(gammaLo.scale(rs,i,h) + gammaHi.scale(rs,i,h)) + 1)
=E= { ([ gammaLo.scale(rs,i,h) + gammaHi.scale(rs,i,h) * sqExp(uhExpG(rs,h,tRun),10)]
/ [ 1 + sqExp(uhExpG(rs,h,tRun),10)] ) $ abs(gammaLo.scale(rs,i,h) + gammaHi.scale(rs,i,h))

* --- G-REDM might introduce shifters to the commitment terms to let total calorie
* demand follow the calorie supply as a function of the household s per capita income
+ [ (alphaShiftF(rs,h,tRun) $ foodCal(i) + alphaShiftNF(rs,h,tRun) $ (not foodCal(i)))
* max(xa.1(rs,i,h,tRun-1)/pop.1(rs,h,ts)*1000,gammaLESM.1(rs,i,h,tRun-1))
] $ ((not sameas(tRun,"%t0%")) $ ( (alphaShiftF.range(rs,h,tRun) ne 0) or alphaShiftF.1(rs,h,tRun) ne 0) )
)/(abs(gammaLo.scale(rs,i,h) + gammaHi.scale(rs,i,h)) + 1);

e_gammaDAIDADS(rs,dNest,h,ts(tRun)) $ ( (p_demSystem eq AIDADS) $ (gammaLESM.range(rs,dNest,h,tRun) ne 0)
$ p_hasXa(rs,dNest,h) )..

gammaLESM(rs,dNest,h,tRun) /(abs(gammaLo.scale(rs,dNest,h) + gammaHi.scale(rs,dNest,h)) + 1)
=E= { ([ gammaLo.scale(rs,dNest,h) + gammaHi.scale(rs,dNest,h) * sqExp(uhExpG(rs,h,tRun),10)]
/ [ 1 + sqExp(uhExpG(rs,h,tRun),10)] ) $ abs(gammaLo.scale(rs,dNest,h) + gammaHi.scale(rs,dNest,h))

* --- G-REDM might introduce shifters to the commitment terms to let total calorie
* demand follow the calorie supply as a function of the household s per capita income
+ [
(alphaShiftF(rs,h,tRun) $ foodCal(dNest) + alphaShiftNF(rs,h,tRun) $ (not foodCal(dNest)))
* max(xdNest.1(rs,dNest,h,tRun-1)/pop.1(rs,h,ts)*1000,gammaLESM.1(rs,dNest,h,tRun-1))
] $ ((not sameas(tRun,"%t0%")) $ ( (alphaShiftF.range(rs,h,tRun) ne 0) or (alphaShiftF.1(rs,h,tRun) ne 0) ))
)/(abs(gammaLo.scale(rs,dNest,h) + gammaHi.scale(rs,dNest,h)) + 1)
;

```

The correction for the commitment terms uses last year's per capita demands  $xa$  respectively  $xdNest$  as weights. Using large year's marginal budget shares in (62) and last year's per capita demand in (63) should ensure that the provoked changes do not fundamentally change the composition inside of the food and inside of the non-food commodity bundle from what the benchmark step suggests.

Large commitment terms relative to the demand lead to quite low price elasticities and can provoke problems with solving the model. We fudge therefore the updated commitments with a Veelken-Ulbrich smooth min operator at 90% of last year's simulated per capita demands, see (64). This operator is an inbuilt function by GAMS such that it provides derivative information to the solver which might be better compared to, for instance, programming directly the otherwise somewhat simpler Huber max approximation. While we don't allow for negative commitment terms during benchmarking, the correction process could actually introduce negative ones which is another argument to control for calorie intakes at higher income level already during benchmarking to keep the changes from the correction small.

$$\gamma_{it}^* = -VU \left[ \begin{aligned} & -\gamma_{i,t} - \frac{xa_{i,t-1}}{pop_{t-1}} \left( cor_f \wedge i \in f + cor_{nf} \wedge i \in nf \right), \\ & -\frac{9}{10} \frac{xa_{i,t-1}}{pop_{t-1}} \end{aligned} \right] \quad (64)$$

The code implementation looks as follows:

```

*
* --- Veelken-Ulbrich smoothed min to prevent that the gamma shifts provoke negative demands
*
e_gammaAIDADS(rs,dNest,h,ts(tRun)) $ ( (p_demSystem eq AIDADS) $ (gammaLES.range(rs,dNest,h,tRun) ne 0)
$ p_hasXa(rs,dNest,h) ) ..

gammaLES(rs,dNest,h,tRun) / (abs(gammaLo.scale(rs,dNest,h) + gammaHi.scale(rs,dNest,h)) + 1)
=E= [ [-ncpVUSin(-gammaLES(rs,dNest,h,tRun),
-(-0.90*max(0,xdNest.1(rs,dNest,h,tRun-1))/pop.1(rs,h,tRun-1)*1000),abs(xdNest.1(rs,dNest,h,tRun-1))/pop.1(rs,h,tRun-1)))
$ ( (not sameas(tRun,"%t0%")) $ ( (alphaShiftF.range(rs,h,tRun) ne 0) or (alphaShiftF.1(rs,h,tRun) ne 0)))
+ gammaLES(rs,dNest,h,tRun)
$ ( sameas(tRun,"%t0%") or ( (alphaShiftF.range(rs,h,tRun) eq 0) and (alphaShiftF.1(rs,h,tRun) eq 0)))
] / (abs(gammaLo.scale(rs,dNest,h) + gammaHi.scale(rs,dNest,h)) + 1);

e_gammaAIDADS(rs,i,h,ts(tRun)) $ ( (p_demSystem eq AIDADS) $ (gammaLES.range(rs,i,h,tRun) ne 0)
$ p_hasXa(rs,i,h) $ hFlag(rs,h) $ hCur(h) ) ..

gammaLES(rs,i,h,tRun) / (abs(gammaLo.scale(rs,i,h) + gammaHi.scale(rs,i,h)) + 1)
=E= [ [-ncpVUSin(-gammaLES(rs,i,h,tRun),
-(-0.90*max(0,xa.1(rs,i,h,tRun-1))/pop.1(rs,h,tRun-1)*1000),abs(xa.1(rs,i,h,tRun-1))/pop.1(rs,h,tRun-1)))
$ ( (not sameas(tRun,"%t0%")) $ ( (alphaShiftF.range(rs,h,tRun) ne 0) or (alphaShiftF.1(rs,h,tRun) ne 0)))
+ gammaLES(rs,i,h,tRun)
$ ( sameas(tRun,"%t0%") or ( (alphaShiftF.range(rs,h,tRun) eq 0) and (alphaShiftF.1(rs,h,tRun) eq 0)))
] / (abs(gammaLo.scale(rs,i,h) + gammaHi.scale(rs,i,h)) + 1);

```

The calibration code for the AIDADS demand system can be found in “*model/cal\_les.gms*”. The econometrically estimated parameters are read from a GDX file:

```

parameter p_AIDADEst(*,*);
set dummyAidAdsEst / "islam+demo+temp", "gammaUp", "lo", "hi", "kappaUD", "kappaUG", "lin", "cnst", "u", "log", "logSqr", "logSqrt" /;
execute_load "%datdir%/maidads.gdx" p_AIDADEst;

```

Integration into a GTAP model with flexible sectoral aggregation requires a mapping from these categories to the current sectors in the model. That is achieved in two steps:

1. The commodity groups from the AIDADS estimation are mapped to the 57 respectively 65 GTAP sectors (non-aggregated):
2. Next, the mapping *mapi* from the 57 respectively 65 GTAP sectors to the current aggregation is used to assign parameters to the commodities, as defined in the current run:

```

AIDADS_to_GTAP("Miscellaneous goods and services",idat) $ (not sum(AIDADS_to_GTAP(AIDADS_grps,idat),1)) = yes;

set AIDADS_to_i(AIDADS_grps,i) "Link to current sectoral aggregation";
AIDADS_to_i(AIDADS_grps,i) $ sum(AIDADS_to_GTAP(AIDADS_grps,idat),mapi(idat,i)) = YES;

```

3. The estimated parameters are assigned to the products, using budget shares as weights:

```

alphaAIDADS(r,i,h) $ xcshr.1(r,i,h,"%t0%")
= max(1.E-6,sum(AIDADS_to_i(AIDADS_grps,i),
(p_parAIDADS(AIDADS_grps,"alpha")
+sum((r_r(r,rp),deltaDrivers),p_deltaShifters(rp,deltaDrivers)*p_deltaShift(AIDADS_grps,deltaDrivers,"alpha"))
*xcshr.1(r,i,h,"%t0%"))
/sum(AIDADS_to_i(AIDADS_grps,j) $ AIDADS_to_i(AIDADS_grps,i),xcshr.1(r,j,h,"%t0%")));

betaAIDADS(r,i,h) $ xcshr.1(r,i,h,"%t0%")
= max(1.E-6,sum(AIDADS_to_i(AIDADS_grps,i),
(p_parAIDADS(AIDADS_grps,"beta")
+sum((r_r(r,rp),deltaDrivers),p_deltaShifters(rp,deltaDrivers)*p_deltaShift(AIDADS_grps,deltaDrivers,"beta"))
*xcshr.1(r,i,h,"%t0%"))
/sum(AIDADS_to_i(AIDADS_grps,j) $ AIDADS_to_i(AIDADS_grps,i),xcshr.1(r,j,h,"%t0%")));

gammaLo(r,i,h) $ xcshr.1(r,i,h,"%t0%")
= sum(AIDADS_to_i(AIDADS_grps,i), p_parAIDADS(AIDADS_grps,"gammaLo")*xcshr.1(r,i,h,"%t0%"))
/sum(AIDADS_to_i(AIDADS_grps,j) $ AIDADS_to_i(AIDADS_grps,i),xcshr.1(r,j,h,"%t0%")));

gammaHi(r,i,h) $ xcshr.1(r,i,h,"%t0%")
= sum(AIDADS_to_i(AIDADS_grps,i), p_parAIDADS(AIDADS_grps,"gammaHi")*xcshr.1(r,i,h,"%t0%"))
/sum(AIDADS_to_i(AIDADS_grps,j) $ AIDADS_to_i(AIDADS_grps,i),xcshr.1(r,j,h,"%t0%")));

```

In order to ensure that both vectors fulfil the adding up condition exactly, the elements of in the vector *alphaAIDADS* and *betaAIDADS* are scaled such their sum is equal to unity:

```

alphaAIDADS.fx(r,i,h,t0) $ xcshr.1(r,i,h,t0) = alphaAIDADS(r,i,h,t0)/sum(j,alphaAIDADS(r,j,h,t0));
betaAIDADS.fx(r,i,h,t0) $ xcshr.1(r,i,h,t0) = betaAIDADS(r,i,h,t0)/sum(j,betaAIDADS(r,j,h,t0));

```

The cardinal utility level, which drives the marginal budget composition, is an estimated parameter. To this end, Roson and van der Mensbrugghe 2018 regress the estimated utility levels on real GDP per capita. The similar regression results from Britz 2021 are here employed to derive the utility level at the benchmark and thus the marginal budget shares:



```

Uh1.l(r,h,t0) $ p_perCapInc(r,h,"%t0%")
= (p_estU("cnst") + p_estU("log") * log(p_perCapInc(r,h,"%t0%"))
  + p_estU("logSqr") * sqr(log(p_perCapInc(r,h,"%t0%")))
  + p_estU("logSqrt") * sqrt(log(p_perCapInc(r,h,"%t0%")))
  + sum( (r_r(r,rNat),deltaDrivers), p_estu(deltaDrivers) * p_deltaShifters(rnat,deltaDrivers)
);

```

The obtained marginal budget shares *alphaLES* at benchmark income and prices are then used to define the commitment terms. The code is not shown here as it was already comprised in CGEBox code and it is illustrated in the associated documentation.

Whenever that calibration procedure becomes infeasible, some deviations from the given *alphas* and *betas* are allowed:

```

if ( (fitLes.suminfes ne 0) and ( p_demSystem eq AIDADS) ,
*
*   --- AIDADS failed, GAMMA bounds became binding, release bounds on alpha and beta
*
alphaAIDADS.lo(rNat,i,h,t0) = min(alphaAIDADS.l(rNat,i,h,t0) * 0.00001,alphaAIDADS.l(rNat,i,h,t0) * 1000);
alphaAIDADS.up(rNat,i,h,t0) = max(0.1,alphaAIDADS.l(rNat,i,h,t0) * 0.00001,alphaAIDADS.l(rNat,i,h,t0) * 1000);

betaAIDADS.lo(rNat,i,h,t0) = min(betaAIDADS.l(rNat,i,h,t0) * 0.00001,betaAIDADS.l(rNat,i,h,t0) * 1000);
betaAIDADS.up(rNat,i,h,t0) = max(0.1,betaAIDADS.l(rNat,i,h,t0) * 0.00001,betaAIDADS.l(rNat,i,h,t0) * 1000);
solve fitLES using NLP minimizing fit;
);

```

In that case, squared relative differences between the corrected and a priori marginal budget shares enter the objective function:

```

e_fitLes ..
fit * sum ( (rs(r),i,h,t0) $ xaFlag(r,i,h),1)
=E= sum ( (rs(r),i,h,t0) $ xaFlag(r,i,h),
  sqr( pElas(r,i,h,t0)-pElas.l(r,i,h,t0))*0.01
  + sqr( incElas(r,i,h,t0)-incElas.l(r,i,h,t0)) )
*
*   --- squared relative deviations from given parameters in case of AIDADS
*   if gamma bounds become binding
*
+ sum ( (rs(r),i,h,t0) $ xaFlag(r,i,h), sqr(v_scaleAlpha(r,i,h)-1)) $ (p_demSystem eq AIDADS);

```

Where *v\_scaleAlpha* enters the definition of *alphaLES* and *betaLES*:

```

e_alphaAIDADS1(rs(r),i,h,t0) $ (alphaAIDADS.l(r,i,h,t0) + betaAIDADS.l(r,i,h,t0) $ (p_demSystem eq AIDADS))..
alphaLES(r,i,h,t0) =E= [ alphaAIDADS(r,i,h,t0)*v_scaleAlpha(r,i,h)
  + betaAIDADS(r,i,h,t0)*v_scaleAlpha(r,i,h) * ( uh(r,h,t0)*p_uh0(r,h)) ]
/ [ 1 + uh(r,h,t0)*p_uh0(r,h) ];

```

Such as:

```

--- marginal budget shares must add to unity

e_sumAlpha(rs(r),h,t0) $ ( (p_demSystem eq AIDADS) $ sum(i,xaFlag(r,i,h))) ..
sum(i $ xaFlag(r,i,h), alphaAIDADS(r,i,h,t0)*v_scaleAlpha(r,i,h)) =E= 1;

e_sumBeta(rs(r),h,t0) $ ( (p_demSystem eq AIDADS) $ sum(i,xaFlag(r,i,h)))..
sum(i $ xaFlag(r,i,h), betaAIDADS(r,i,h,t0)*v_scaleAlpha(r,i,h)) =E= 1;

```

Finally, a multiplier *p\_uh1* (termed *A* in Cranfield et al. 2000, eq 10 ) is calculated:

```

$iftheni.AIDADS "%demSystem%" == "AIDADS"

Uh.lo(r,h,t0) = 0.001;
Uh.up(r,h,t0) = inf;

p_uh1(r,h) = { sum{i $ alphaLES(r,i,h,"%t0%"),
                    alphaLES(r,i,h,"%t0%") * log( xa(r,i,h,"%t0%") - gammaLES(r,i,h,"%t0%")*pop.l(r,h,"%t0%"))]}
+sum{dNest $ alphaLES(r,dNest,h,"%t0%"),
      alphaLES(r,dNest,h,"%t0%") * log( xdNest(r,dNest,h,"%t0%") - gammaLES(r,dNest,h,"%t0%")*pop.l(r,h,"%t0%"))]}
}
- (log(uh.l(r,h,"%t0%"))+p_uh0(r,h));

$endif.AIDADS

```

The code allows introducing aggregates of individual commodities under which CES-sub-nests can be defined, which comprise these individual commodities.

**Table 8: Mapping between AIDADS categories and GTAP sectors**

| AIDADS category       | GTAP sector              | Detailed description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Cereals               | pdr<br>wht<br>gro<br>pcr | Paddy Rice: rice, husked and unhusked<br>Wheat: wheat and meslin<br>Other Grains: maize (corn), barley, rye, oats, other cereals<br>Processed Rice: rice, semi- or wholly milled                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Ruminant meat         | Ctl<br><br>cmt           | Cattle: cattle, sheep, goats, horses, asses, mules, and hinnies; and semen thereof<br>Cattle Meat: fresh or chilled meat and edible offal of cattle, sheep, goats, horses, asses, mules, and hinnies. raw fats or grease from any animal or bird                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Other animal products | oap<br><br>omt           | Other Animal Products: swine, poultry and other live animals; eggs, in shell (fresh or cooked), natural honey, snails (fresh or preserved) except sea snails; frogs' legs, edible products of animal origin n.e.c., hides, skins and furskins, raw, insect waxes and spermaceti, whether or not refined or coloured<br><br>Other Meat: pig meat and offal. preserves and preparations of meat, meat offal or blood, flours, meals and pellets of meat or inedible meat offal; greaves                                                                                                                                                                                                                                                                                                                            |
| Fish                  | fsh                      | Fishing: hunting, trapping and game propagation including related service activities, fishing, fish farms; service activities incidental to fishing                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Dairy                 | mil                      | Milk: dairy products                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Milk                  | rmk                      | Raw milk                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Oils_Fats             | Osd<br>vol               | Oil Seeds: oil seeds and oleaginous fruit; soy beans, copra<br>Vegetable Oils: crude and refined oils of soya-bean, maize (corn), olive, sesame, ground-nut, olive, sunflower-seed, safflower, cotton-seed, rape, colza and canola, mustard, coconut palm, palm kernel, castor, tung jojoba, babassu and linseed, perhaps partly or wholly hydrogenated,inter-esterified, re-esterified or elaidinised. Also margarine and similar preparations, animal or vegetable waxes, fats and oils and their fractions, cotton linters, oil-cake and other solid residues resulting from the extraction of vegetable fats or oils; flours and meals of oil seeds or oleaginous fruits, except those of mustard; degreas and other residues resulting from the treatment of fatty substances or animal or vegetable waxes. |
| Fruits_vegs           | v_f<br><br>ocr           | Veg & Fruit: vegetables, fruitvegetables, fruit and nuts, potatoes, cassava, truffles<br>Other Crops: live plants; cut flowers and flower buds; flower seeds and fruit seeds; vegetable seeds, beverage and spice crops, unmanufactured tobacco, cereal straw and husks, unprepared, whether or not chopped, ground, pressed or in the form of pellets; swedes, mangolds, fodder roots, hay, lucerne (alfalfa), clover, sainfoin, forage kale, lupines, vetches and similar forage                                                                                                                                                                                                                                                                                                                               |

| AIDADS category                    | GTAP sector                                                                 | 9 Detailed description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------------------------|-----------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                    | c_b                                                                         | products, whether or not in the form of pellets, plants and parts of plants used primarily in perfumery, in pharmacy, or for insecticidal, fungicidal or similar purposes, sugar beet seed and seeds of forage plants, other raw vegetable materials<br>Cane & Beet: sugar cane and sugar beet                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Food_other                         | ofd                                                                         | Other Food: prepared and preserved fish or vegetables, fruit juices and vegetable juices, prepared and preserved fruit and nuts, all cereal flours, groats, meal and pellets of wheat, cereal groats, meal and pellets n.e.c., other cereal grain products (including corn flakes), other vegetable flours and meals, mixes and doughs for the preparation of bakers' wares, starches and starch products; sugars and sugar syrups n.e.c., preparations used in animal feeding, bakery products, cocoa, chocolate and sugar confectionery, macaroni, noodles, couscous and similar farinaceous products, food products n.e.c.                                                                                                                                                                                 |
| Sugar                              | sgr                                                                         | sugar                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Beverages and tobacco              | b_t                                                                         | Beverages and Tobacco products                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Clothing                           | tex<br>wap<br>lea<br><br>wol<br>pfb                                         | Textiles: textiles and man-made fibres<br>Wearing Apparel: Clothing, dressing and dyeing of fur<br>Leather: tanning and dressing of leather; luggage, handbags, saddlery, harness and footwear<br>Wool: wool, silk, and other raw animal materials used in textile<br>Plant Fibres: cotton, flax, hemp, sisal and other raw vegetable materials used in textiles                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Housing                            | dwe<br><br>obs<br>wtr<br>gdt<br><br>ely<br>coa<br>oil<br><br>gas<br><br>isr | Dwellings: ownership of dwellings (imputed rents of houses occupied by owners)<br>Other Business Services: real estate, renting and business activities<br>Water: collection, purification and distribution<br>Gas Distribution: distribution of gaseous fuels through mains; steam and hot water supply<br>Electricity: production, collection and distribution<br>Coal: mining and agglomeration of hard coal, lignite and peat<br>Oil: extraction of crude petroleum and natural gas (part), service activities incidental to oil and gas extraction excluding surveying (part)<br>Gas: extraction of crude petroleum and natural gas (part), service activities incidental to oil and gas extraction excluding surveying (part)<br>Insurance: includes pension funding, except compulsory social security |
| Furniture, Maintenance, appliances | p_c<br>frs<br>omn<br><br>lum<br><br>crp<br><br>nmm<br>i_s<br>nfm<br><br>fmp | Petroleum & Coke: coke oven products, refined petroleum products, processing of nuclear fuel<br>Forestry: forestry, logging and related service activities<br>Other Mining: mining of metal ores, uranium, gems. other mining and quarrying<br>Lumber: wood and products of wood and cork, except furniture; articles of straw and plaiting materials<br>Chemical Rubber Products: basic chemicals, other chemical products, rubber and plastics products<br>Non-Metallic Minerals: cement, plaster, lime, gravel, concrete<br>Iron & Steel: basic production and casting<br>Non-Ferrous Metals: production and casting of copper, aluminium, zinc, lead, gold, and silver<br>Fabricated Metal Products: Sheet metal products, but not machinery and                                                          |



| AIDADS category        | GTAP sector                                                   | Detailed description                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|------------------------|---------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                        | cns<br>ele<br><br>ome<br><br>omf                              | equipment<br>Construction: building houses factories offices and roads<br>Electronic Equipment: office, accounting and computing machinery, radio, television and communication equipment and apparatus<br><br>Other Machinery & Equipment: electrical machinery and apparatus n.e.c., medical, precision and optical instruments, watches and clocks<br>Other Manufacturing: includes recycling                                                              |
| Transport              | Otp<br><br>wtp<br>atp<br>mvh<br><br>otn<br><br>Isr<br><br>p_c | Other Transport: road, rail ; pipelines, auxiliary transport activities; travel agencies<br>Water transport<br>Air transport<br><br>Motor Motor vehicles and parts: cars, lorries, trailers and semi-trailers<br>Other Transport Equipment: Manufacture of other transport equipment<br>nsurance: includes pension funding, except compulsory social security<br>Petroleum & Coke: coke oven products, refined petroleum products, processing of nuclear fuel |
| Communication          | cmn                                                           | Communications: post and telecommunications                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Recreation             | ros<br><br>ppp                                                | Recreation & Other Services: recreational, cultural and sporting activities, other service activities; private households with employed persons (servants)<br>Paper & Paper Products: includes publishing, printing and reproduction of recorded media                                                                                                                                                                                                        |
| Education              | osg                                                           | Other Services (Government): public administration and defense; compulsory social security, education, health and social work, sewage and refuse disposal, sanitation and similar activities, activities of membership organizations n.e.c., extra-territorial organizations and bodies                                                                                                                                                                       |
| Health                 | Osg                                                           | See above                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Restaurants and hotels | Trd                                                           | Trade: all retail sales; wholesale trade and commission trade; hotels and restaurants; repairs of motor vehicles and personal and household goods; retail sale of automotive fuel                                                                                                                                                                                                                                                                             |
| Rest                   |                                                               | Other Financial Intermediation: includes auxiliary activities but not insurance and pension funding (see next)                                                                                                                                                                                                                                                                                                                                                |

### 3.14.5 Differentiated productivity growth

#### Background and literature review

Productivity does not vary uniformly among industries and sectors. Harberger 1998<sup>50</sup> points out that the whole dynamics of economic progress actually resembles the growth process of “mushrooms”, rather than the steady rise of “yeast”. Indeed, differential productivity growth is one key factor of structural change in the economic systems, and probably the most important one (Swiecki 2017<sup>51</sup>). Several implications of different growth rates have been investigated in the literature, e.g.: relevance and empirics of the so-called “Baumol's disease” (Baumol 1986<sup>52</sup>, Triplett and Bosworth 2003<sup>53</sup>,

<sup>50</sup> Harberger, A. C., (1998). A Vision of the Growth Process. *The American Economic Review* 88 (1), 1-32

<sup>51</sup> Swiecki, T. (2017) Determinants of structural change. *Review of Economic Dynamics* 24, 95-131

<sup>52</sup> Baumol, W. J. (1986). Productivity Growth, Convergence and Welfare: What the Long-Run Data Show. *The American Economic Review* 76 (5), 1072-1085

Young 2014<sup>54</sup>); specialization and international trade (McMillan and Rodrik 2011<sup>55</sup>, Caron and Markusen 2014<sup>56</sup>); “premature deindustrialization” (Rodrik 2016<sup>57</sup>).

To introduce differentiated productivity growth in the G-RDEM model, we build on Roson 2018, who estimated trends in labour productivity, using the Groening GGDC 10-Sector Database (de Vries et al. 2015<sup>58</sup>). In that study, some trends and country specific dummies for labor productivity (VA/employment) are estimated. Results are subsequently employed in a cluster analysis, where three groups of countries with similar characteristics are identified. Table 9 below shows some of the findings used to obtain parameters for G-RDEM:

**Table 9: Average labour productivity growth rates**

| Cluster | AGR  | MAN   | SER  | TOT  |
|---------|------|-------|------|------|
| Rising  | 6.23 | 11.43 | 5.65 | 8    |
| Steady  | 7    | 7.88  | 5    | 5.93 |
| Lagging | 5.17 | 5.32  | 2.34 | 3.16 |

Source: Roson 2018

The last column in Table 1 (TOT) displays the average (yearly) growth rate in labor productivity in each group. It refers to value added per worker or hour, so it accounts for capital deepening and similar effects. Interestingly, the differences among industries depends on how fast an economy is growing.

### Estimation of differentiated productivity growth

In the development of the G-RDEM model we are not concerned about labour productivity in itself, but rather on the relative differences among the three broad sectors of Agriculture, Manufacturing and Services. To this end, a correspondence between the three clusters and the annual GDP growth rates used in the SSPs was established. The distribution of IIASA SSP data (OECD) on GDP was considered, and it was assumed that the average GDP growth in the Lagging group of countries corresponds to the 20% percentile of the SSP distribution, 50% for Steady, 80% for Rising. This means 1.2%, 2.5%, and 4.9%, respectively.

Second, the ratio of each sector productivity rate, relative the slowest growing sector, which is Services, was computed. Third, for each industry a quadratic interpolation between the three multipliers and the references GDP growth rates was undertaken, thereby getting three parameters of a quadratic polynomial relationship between a sectoral productivity shifter (ratio between industry

<sup>53</sup> Triplett, J. E., Bosworth, B. P., (2003). Productivity Measurement Issues in Services Industries: Baumol's Disease Has Been Cured. *Economic Policy* (September), 23-33

<sup>54</sup> Young, A., (2014). Structural transformation, the mismeasurement of productivity growth, and the cost disease of services. *American Economic Review* 104 (11), 3635-3667

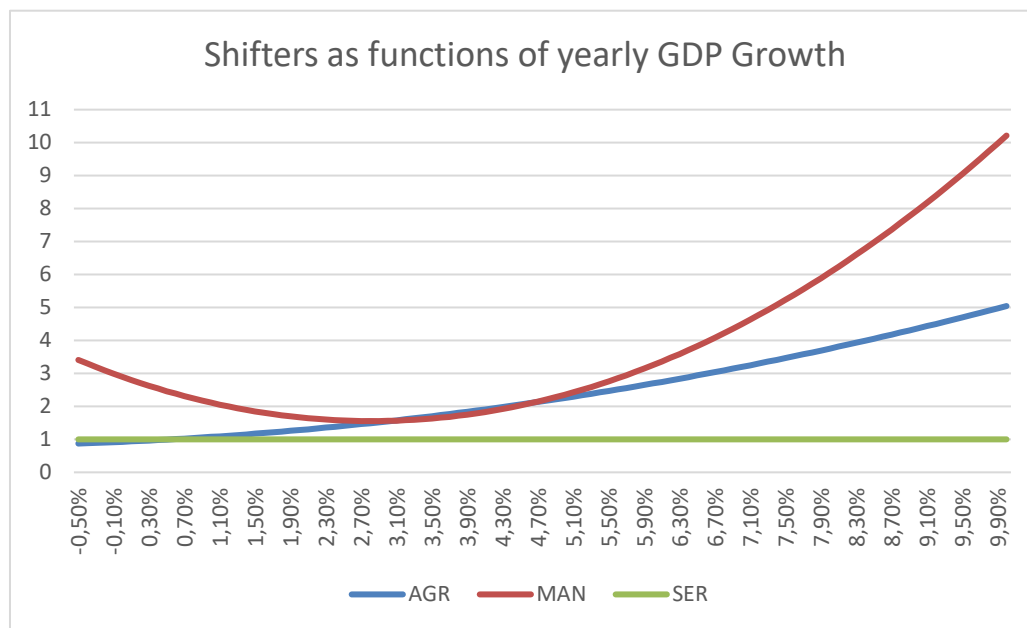
<sup>55</sup> McMillan, M., Rodrik, D., (2011). Globalization, Structural Change and Productivity Growth. NBER Working Paper w17143

<sup>56</sup> Caron, J., Markusen, J. R., (2014). International Trade Puzzles: a Solution Linking Production and Preferences. *The Quarterly Journal of Economics*, 1501-1552

<sup>57</sup> Rodrik, D., (2016). Premature deindustrialization. *Journal of Economic Growth* 21, 1-33

<sup>58</sup> de Vries, G., Timmer, M., de Vries, K., (2015). Patterns of structural change in develop-ing countries. In: Weiss, J., Tribe, M. (Eds.), *Routledge Handbook of Industry and Development*. Routledge, pp. 65-83

growth rate and the corresponding one in the Services) and GDP annual growth. This gives raise to the functions displayed in Figure 8.



**Figure 8 : Productivity growth relative to GDP growth**

The key finding is that productivity differentials are minimized (although still significant) at a moderate GDP growth of around 2%. For higher or lower rates, we can see that differences amplify, with manufacturing becoming the key sector. Notice that the shifter is a multiplier: if aggregate growth is negative, it will likely become negative for the reference slow sector as well. When the shifter for Manufacturing is high and positive, this means that productivity is decreasing there more than in the rest of the economy. In other words, productivity growth in Manufacturing appears as strongly correlated with the aggregate productivity growth, which suggests the existence of inter-sectoral externalities.

Implementation in G-RDEM is straightforward. Total factor productivity in the Services  $tfp(r)$  becomes endogenous during the construction of the baseline and is kept then fixed during counterfactual simulations. Total factor productivity for other sectors (indexed by  $i$ ) in region  $r$  at time  $t+1$  are defined as  $tfp(r)*sh(i,r)$ , where the latter is determined by equations like:

$$Sh(i,r) = a + b \frac{gdp(r,t+1) - gdp(r,t)}{gdp(r,t)} + c \left( \frac{gdp(r,t+1) - gdp(r,t)}{gdp(r,t)} \right)^2 \quad (4)$$

Here are the estimated values for the three parameters  $a$ ,  $b$  and  $c$ :

Table 10: Estimated parameter for sector specific productivity growth

|   | AGR      | MAN      |
|---|----------|----------|
| a | 0.925391 | 2.893917 |
| b | 11.99205 | -94.8599 |
| c | 291.8147 | 1680.554 |

### Implementation of the sector specific productivity shifters

The parameters are derived from estimates in “GTAPRDEM\diffProdGrowth.gms”:

```

: set tfpPars / a,b,c/
:
: Table p_diffProdShifters(tfpPars, globSec)
:
:      AGR      MAN      SER
:  a      0.925391 2.893917      1
:  b      11.99205 -94.8599
:  c      291.8147 1680.554
:
:
: p_prodShft(tfpPars,a) = sum( (globSec_iOri(globSec,iOri),mapa(iOri,a)),p_diffProdShifters(tfpPars,globSec))
:                          /sum( (globSec_iOri(globSec,iOri),mapa(iOri,a)),1);

```

Which requires a mapping from the original GTAP sectors iOri to the currently active sector aggregation defined in the same program, of which here only an extracted part is shown:

```

:
: set globSec / AGR,MAN,SER /;
:
: set globSec_iOri(globSec,iOri) /
:
:      AGR.      pdr      "Paddy rice"
: ----- 13 line(s) not displayed
:      MAN.      coa      "Coal"
: ----- 31 line(s) not displayed
:      SER.      trd      "Trade"
: ----- 10 line(s) not displayed
: /;
:

```

The production shifter equation derives from the economy wide, regional specific tfp shifter *axpNat* sector specific shifters *lambdaVa* based on the quadratic function as detailed above (see *gams/model/model.gms*). The equation updates the last year's shifter *lamdbaf.l(rs,f,a,tRun-1)* for primary factors based on a uniform economy-wide productivity shifter *axpNat* – implicitly defined by targeting real GDP. It can be modulated by changes in regional productivity *axpReg*. The if condition *\$ l(f)* allows to only target changes in labor productivity. The last line restricts changes in the productivity of natural resources such as oil field and fish stocks to 10% of the general change. The middle parts modulate the economy-wide shifters depending on GDP growth, applying the regression coefficients are defined above.

```

--- production shifter equation
set enerNatRes(a) "Oil, gas, coal extraction";

e_prodShift(rs,f,aIn(a),ts(tRun)) $ (xFlag(rs,f,a) $ (lambdaf.range(rs,f,a,tRun) ne 0)
    $ sum(r_r(rs,rNat),rgdmp.l(rNat,tRun-1)) $ (not disr(rs))) ..

lambdaf(rs,f,a,tRun)s(lambdaf,'rs,f,a,tRun')

=E= {lambdaf.l(rs,f,a,tRun-1)
    + [
        --- can be endogenous to target e_gvaShare
        expReg(rs,tRun) $ subr(rs)

    + sum(r_r(rs,rNat), expNat(rsNat,tRun)
        $iftheni.tfpLab "%TFPLab%"=="on"
        $ l(f)
        $endif.tfpLab
        * (
            p_prodShift("a",a)
            + p_prodShift("b",a) * [ (rgdmp(rsNat,tRun)/rgdmp.l(rsNat,tRun-1))**(1/p_timeStep(tRun))-1 ]
            + p_prodShift("c",a) * sqrt( (rgdmp(rsNat,tRun)/rgdmp.l(rsNat,tRun-1))**(1/p_timeStep(tRun))-1 ) ]
        ) * (1 - 0.90 $ nr(f) $ enerNatRes(a)) } s(lambdaf,'rs,f,a,tRun')

```

### 3.14.6 Endogenous saving rates

#### Background and literature review

We aim at developing a simple but robust mechanism to render aggregate saving rates in G-RDEM endogenous. One strand of literature, relying on cross-country differences of saving rates (e.g. Kirsanova and Sefton 2007<sup>59</sup>), works with micro-economic survey data. It explicitly accounts for factors such as demography, welfare state, retirement behaviour, borrowing constraints, income distribution over a lifetime and its uncertainty, as well as capital gains. The focus here is on the life-cycle hypothesis, which considers the change in available income over a lifetime. While these papers give robust evidence that the factors indeed explain the saving behaviour of individuals or households, they typically offer results only for one or a smaller group of countries.

Rather, we draw here on studies which employ cross-sectional analyses over countries to evaluate the factors affecting the economy-wide aggregate saving rates. Most of these works also take the lifecycle hypotheses into account (although indirectly) and find that even in cross-country analyses larger proportions of the young and the elderly compared to persons in working age (dependency ratios) generally decrease the saving rate (Doshi 1994<sup>60</sup>, Masson et al. 1998<sup>61</sup>, Laoayza et al. 2000<sup>62</sup>).

#### Estimation

Instead of directly using parameter values from the literature, we carry out our own cross-section estimation, using GTAP 9 and other data used in our modelling framework, to overcome any potential divergence in definitions, measurement units etc.. The reader might note that we face a potential endogeneity issue: higher rates of GDP growth require increased capital accumulation, thus larger net investments and consequently higher saving rates. The saving rate and GDP growth are hence structurally dependent. However, this is not an issue of major concern in this context, since we are not integrating the estimated equation into the model, but only updating saving rates, given GDP projections. Hence, our aim is solely to ensure that correlation, not causation, is properly accounted

<sup>59</sup> Kirsanova, T., & Sefton, J. (2007). A comparison of national saving rates in the UK , US and Italy. *European Economic Review*, 51(8), 1998-2028

<sup>60</sup> Doshi, K. (1994). Determinants of the saving rate: an international comparison. *Contemporary Economic Policy*, 12(1), 37-45

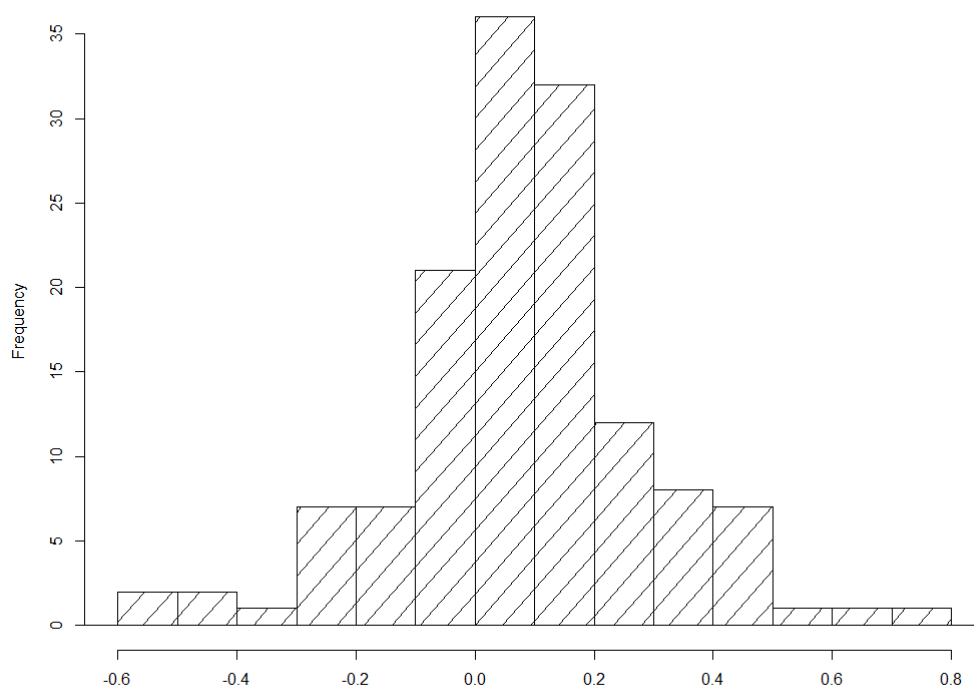
<sup>61</sup> Masson, P. R., Bayoumi, T., & Samiei, H. (1998). International evidence on the determinants of private saving. *The World Bank Economic Review*, 12(3), 483-501

<sup>62</sup> Loayza, N., Schmidt-Hebbel, K., & Servén, L. (2000). What drives private saving across the world?. *Review of Economics and Statistics*, 82(2), 165-181

for. Notice also that we obtained our estimates from a sectional data base, which would make it impossible the introduction of lagged variables as instruments.

The distribution of the national aggregate saving shares in the GTAP 9 data set reveals a large spread, as shown in the Figure 9. We regressed those saving rates with OLS against the following explanatory variables:

1. Population composition by age group from the IIASA repository for 2010 (Lutz et al. 2017<sup>63</sup>)
2. GDP growth per capita from 2010 to 2011, in PPPs, from the OECD Env. Growth Model data base as found in the IIASA repository
3. Foreign savings (trade balance) relative to regional income, from the GTAP 9 data base
4. We also tested, as a potential explanatory variable, the share of government consumption on regional income, but did not find a statistically significant relation.



**Figure 9 : Distribution of aggregated savings rates in GTAP 9**

We found a very good fit for our sectional analysis, with a  $R^2$  at 92% and all variables (with the exemption of the young dependency rate) statistically significant at  $<0.1\%$ . The young DR is nonetheless significant at the 5% level. All variables have the expected sign: dependency ratios decrease the saving rates, as postulated by the life cycle hypothesis, while a higher income per capita and a higher growth rate increase the saving rate. A positive trade surplus (i.e. negative foreign savings) also tends to increase the saving rates.

<sup>63</sup> Lutz, W., Butz, W. P., & Samir, K. E. (Eds.). (2017). World Population & Human Capital in the Twenty-First Century: An Overview. Oxford University Press

**Table 11: Regression output for saving rate estimation**

| Coefficient                                         | Estimate | Std. error | t-value | Pr(>  t )    |
|-----------------------------------------------------|----------|------------|---------|--------------|
| Intercept                                           | 0.258    | 0.058      | 4.401   | 2.22e-05 *** |
| Yearly GDP growth %                                 | 2.208    | 0.390      | 5.659   | 9.18e-08 *** |
| GDP per capita                                      | 0.001    | 0.0003     | 2.899   | 0.00439 **   |
| Dependency ratio elderly, square root               | -0.364   | 0.079      | -4.590  | 1.02e-05 *** |
| Dependency ratio young persons, square root         | -0.145   | 0.056      | -2.583  | 0.01089 *    |
| Trade balance, relative to regional income          | -0.967   | 0.047      | -20.433 | < 2e-16 ***  |
| Trade balance, relative to regional income, squared | 0.334    | 0.067      | 4.944   | 2.30e-06 *** |

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05

Residual standard error: 0.05893 on 131 degrees of freedom

Multiple R-squared: 0.9213, Adjusted R-squared: 0.9177

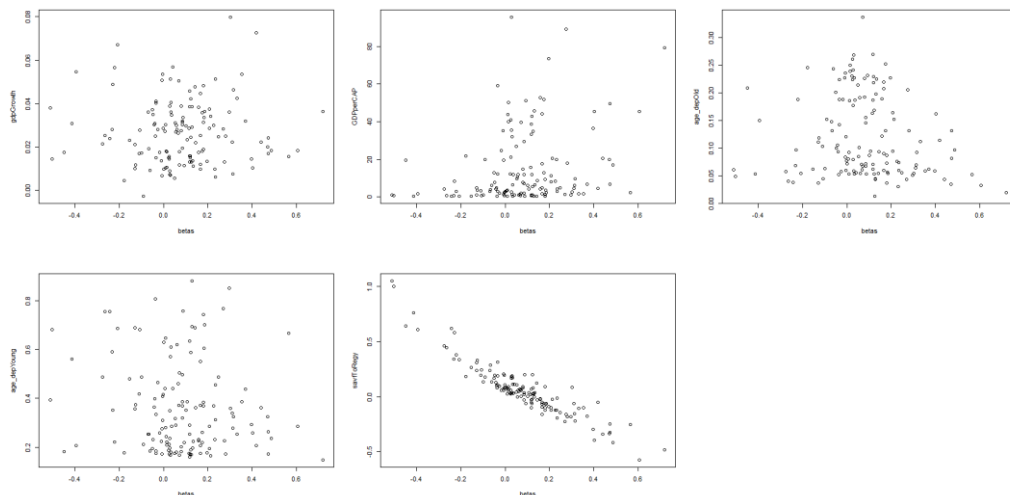
F-statistic: 255.5 on 6 and 131 DF, p-value: < 2.2e-16

The good fit of the regression stems to a large degree from the inclusion of foreign savings relative to regional income, i.e. a trade surplus indicator, (see Table 4 below), while the contributions of the dependency ratios and GDPperCap are in a similar range, with GDP growth trailing.

**Table 12: Analysis of Variance for saving rate estimation**

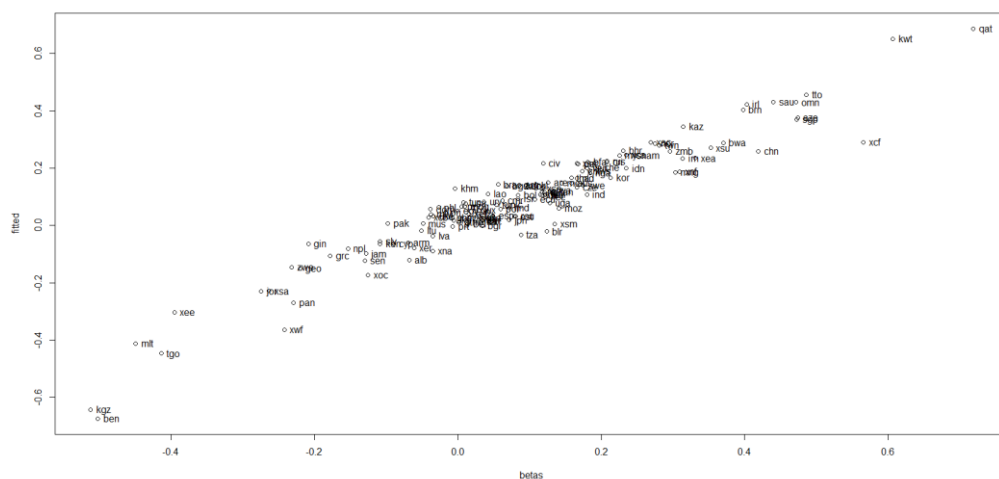
| Coefficient                                         | DF  | Sum Sq | F value | Pr(>  t )       |
|-----------------------------------------------------|-----|--------|---------|-----------------|
| Yearly GDP growth %                                 | 1   | 0.0001 | 0.0287  | 0.8658          |
| GDP per capita                                      | 1   | 0.5271 | 151     | < 2.2e-16 ***   |
| Dependency ratio elderly, square root               | 1   | 0.5092 | 146     | < 2.2e-16 ***   |
| Dependency ratio young persons, square root         | 1   | 0.3712 | 106     | < 2.2e-16 ****  |
| Trade balance, relative to regional income          | 1   | 3.8306 | 1103    | < 2.2e-16 ***** |
| Trade balance, relative to regional income, squared | 1   | 0.0849 | 24      | 2.297e-06 ***   |
| Error term                                          | 131 | 0.4549 |         |                 |

Some scatter plots (visualizing the ANOVA results above) between the explanatory variables and the saving rate are shown below in Figure 10.



**Figure 10 : Scatter plots of explanatory variables against the saving rate**

The high contribution of the relation between foreign savings and regional income is not astonishing, because it controls for cases such as oil exporting countries (high saving rates) as well as some other countries, often developing ones, with very low saving rates. The foreign saving indicators can be hence rather understood as a control variable for country specific unobserved features (large receiver of development aid in a group a country with otherwise similar macro-economic indicators, rich oil and gas reserves, tax havens ...). Accordingly, we do not use changes in the foreign savings during the process of baseline construction, to update the saving rates.



**Figure 11 : Fitted against observed saving rates, GTAP region codes as labels**

Note that the fitted values cannot be used as such, since we would then neglect any unexplained additional factors, which could imply large changes in the saving rates from the benchmark to subsequent simulation periods in some countries. Thus, we use relative changes in the estimates – neglecting foreign saving – to update the saving rates used in the model.

### Implementation of the endogenous saving rates

The GAMS code does not use the estimated intercept, but rather updates over the simulation horizon the “betas” for the individual countries in terms of the difference between two consecutive time



periods. Note that we leave the impact of the saving rate out – the estimated coefficient is multiplied in the formula below. The names of the RHS terms should be hopefully self-explaining.

```
p_betaTrack(rNat,tSim,"full") $ (tSim.pos gt 1)
=
( 0.2584670
+ sqrt(p_driver(rNat,"age0_14",tSim)/p_driver(rNat,"age15_64",tSim)) * (-0.1454039)
+ sqrt(p_driver(rNat,"age_gt65",tSim)/p_driver(rNat,"age15_64",tSim)) * (-0.3640043)
+ p_perCapIncome(rNat,tSim) * (1+p_growthRate(rNat,"GDPPerCAP",tSim)/(100*p_timeStep(tSim))) * 0.0010213
+ (p_growthRate(rNat,"GDPPerCAP",tSim)/(100*p_timeStep(tSim))) * 2.2080586
+ (savf.l(rnat,tSim-1)/regy.l(rnat,tSim-1)) * (-0.9674676)*0
+ sign(savf.l(rnat,tSim-1)/regy.l(rnat,tSim-1)) * sqrt(savf.l(rnat,tSim-1)/regy.l(rnat,tSim-1)) * 0.3347001*0
);
```

Note that using dependency ratios restrict the application to simulation exercises where a projection of the population by age group is available. Given that the IIASA repository features these projections for SSP1-SSP5, we suggest to use one of these SSPs population scenarios directly, or at least the projected age composition.

The parameter `p_betaTrack` gives the estimated saving rates at a zero trade balance which is by definition is average in any simulation. It is calculated for  $t-1$  and  $t-2$ , and its relative change measured against the current saving rate or the simulated is used to update the saving rates in the model:

```
betas.l(rNat,tSim) $ (tSim.pos gt 2)
= betas.l(rNat,tSim) * (1 + sign(betas.l(rNat,tSim))
* (p_betaTrack(rNat,tSim,"full") - p_betaTrack(rNat,tSim-1,"full"))
/ max(abs(betas.l(rNat,tSim)),abs(p_betaTrack(rNat,tSim-1,"full"))));
```

In order to maintain consistency, in the top-level expenditure distribution system of the regional household, the expenditure shares of savings, private and government consumption are scaled to unity:

```
p_betaTrack(rNat,tSim,"sum") = ( betas.l(rNat,tSim) + betag.l(rNat,tSim) + betaphi.l(rNat,tSim));

betas.fx(rNat,tSim) = betas.l(rNat,tSim) / p_betaTrack(rNat,tSim,"sum");
betag.fx(rNat,tSim) = betag.l(rNat,tSim) / p_betaTrack(rNat,tSim,"sum");
betap.fx(rNat,tSim) = betap.l(rNat,tSim) / p_betaTrack(rNat,tSim,"sum");
betaphi.l(rNat,tSim) = betaphi.l(rNat,tSim) / p_betaTrack(rNat,tSim,"sum");
```

### 3.14.7 Debt accumulation from foreign savings

Accounting identities in the model ensure (for each time period) that the sum of regional and foreign savings in each region equals gross investments, while foreign savings are equal to the foreign trade deficit. The latter is determined, in the GTAP model (Hertel and Tsigas 1997<sup>64</sup>, Corong et al. 2017<sup>65</sup>), which defines the intra-periodal equilibrium in G-RDEM) by the mechanism of regional allocation of investments. It is based on a distribution of global savings, driven by relative expected returns on capital, as it is briefly illustrated in the following.

Let denote the price of a homogeneous capital factor (services) as  $p_c$  and  $p_i$  as the price of investments (the cost of producing one unit of new capital good),  $\kappa$  the tax rate on capital earnings,  $fdepr$  the depreciation rate. The net rate of return in a region  $r$  ( $rorc$ ) is defined in the GTAP model as:

<sup>64</sup> Hertel, T.W. & Tsigas, M.E. (1997) "Structure of GTAP", in: T.W. Hertel (ed.), Global Trade Analysis: Modeling and Applications, Cambridge University Press

<sup>65</sup> Corong, E., Hertel, T., McDougall, R., Tsigas, M., & van der Mensbrugghe, D. (2017). The Standard GTAP Model, Version 7. Journal of Global Economic Analysis, 2(1), 1-119. doi:http://dx.doi.org/10.21642/JGEA.020101AF

$$rorc_r = \frac{p_{c,r} [1 - \overline{\kappa_r}]}{p_{i,r}} - \overline{fdepr} \quad (5)$$

The expected rate of return  $rorc$  considers the difference between start and end of period capital stocks,  $k_s$  and  $k_e$ . The logic is that investors should become more cautious when aggregate investments lead to large changes in capital stocks:

$$rore_r = rorc_r \left( \frac{k_{s,r}}{k_{e,r}} \right)^{\overline{rorFlex}} \quad (6)$$

The parameter  $rorFlex$  (whose default value is 10 for all for regions) dampens the relative differences in expected returns, thereby avoiding the generation of unrealistically large flows of (real) capital in international markets. In addition, a regional risk factor is introduced, to ensure that an arbitrage condition for the international investor holds in the calibration data set, meaning that a single global, risk-adjusted return  $rorg$  is identified:

$$rore_r \overline{risk_r} = rorg \quad (7)$$

The condition (7) holds in all periods in G-RDEM, where  $rorg$  and  $rore$  are endogenous variables. Therefore, the relationships above drive the distribution of foreign savings  $fsav$  or, equivalently, the amount of investments in each region (which do not generally match with regional savings).

The global investor hence expects equal returns of  $rorg$  on his savings in any region. Accordingly, the returns in year  $t$  from foreign savings add up to zero as, by construction, the global economy is closed, and total investments equal total savings (equivalently, the global trade balance is zero):

$$\sum_r fsav_r rorg \equiv 0 \quad (8)$$

The model considers four elements to define the BOP: (1) remittances, (2) development aid, (3) foreign savings and (4) payments to or from the global bank. The first two elements are only found if the myGTAP module is used. The debt accumulation mechanism determines the payments to or from the global bank and is based on the following core elements. Firstly, foreign savings are loans paid back in equal instalments over ten years, with a fixed interest rate equal to the expected rate of return in the year  $t$  and country  $rorc$  where the loan is granted. This expected rate  $rorc$  reflects returns to capital after depreciation and direct taxation and a country specific risk parameter  $risk$ . Note that a lower risk parameter implies a higher risk assumed by the global bank, according to the relation between  $rorc$  and the global average expected return  $rorg$ :

$$rorc_{r,t} risk_{r,t} = rorg_{r,t} \quad (65)$$

The total amount of outstanding debt  $debt$  for a country in a year is the sum of past loans granted by the global bank, i.e. in years with a positive foreign saving, minus repayments on past loans. We consider further that countries might default on part of their debt, expressed by the parameter  $accum$  which is the share of the debt honoured:

$$debt_{r,t} = \sum_{tt < t} \max(0, sav_{r,tt}) accum_{r,t} \max(0 - 1/10 [t - tt]) \quad (66)$$

The payments related to old debt comprise the repayments, i.e. 10% on the stock, and the interest rate  $rorc$ :

$$paym_{r,t} = \sum_{tt < t} \max(0, savf_{r,tt}) \text{ accum}_{r,t} \max(0 - 1/10 [t - tt]) (1/10 + rorc_{r,tt}) \quad (67)$$

The payments which the global bank receives in the current year  $t$  from each country  $r$  and for each past year  $tt$  are:

$$returns_{tt,t} = \sum_{r, tt < t} \max(0, savf_{r,tt}) \text{ accum}_{r,t} \max(0 - 1/10 [t - tt]) (1/10 + rorc_{r,tt}) \quad (68)$$

The returns for a country  $r$  in each year  $t$  are a share on the returns on the total global loans granted in previous years  $tt$ . Note that returns from past years are only occurring when the country acted as a lender, i.e. when its foreign savings were negative:

$$returns_{r,t} = \sum_{tt < t} \frac{\max(0, -savf_{r,tt})}{\sum_{rr} \max(0, -savf_{rr,tt})} returns_{tt,t} \quad (69)$$

The regional income equation of the country in each year reflects the different between the returns and the payments, i.e.

$$capTrans_{r,t} = paym_{r,t} - returns_{r,t} \quad (70)$$

That means that payments to the bank decrease the share of the GDP which can spent on final consumption and otherwise.

The share of the debt accum which is assumed to be honoured, i.e. not defaulted, is assumed to depend on the debt stock relative to GDP. Note that we take only foreign debt into account, i.e. the net position with the global bank. Specifically, we assume, that defaulting on part of the debt starts if the foreign debt to gdp level exceeds 30%. For each percentage point above that thresholds, one percentage of defaulted past loans are assumed. That mechanism avoids that country can accumulate very large debt stocks:

$$\text{accum}_{r,t} = \min(\max(1 - \left[ \frac{\text{debt}_{r,t}}{\text{gdp}_{r,t}} - 1 \right] - 0.3)) \quad (71)$$

That correction alone would however assume that the global bank does not respond to the fact that countries are not fully paying back past loans and serving interest. We therefore let the global bank adjust its country specific risk parameter depending on the current year's return received from payments to the bank relative to GDP. The global bank reacts before critical debt levels occur and part of the loans are not honoured, i.e. the risk parameter is already updated earlier. Similar, receiving payments from the bank decreases the risk. As the global bank mechanism is quite sensitive to very low or high risk parameters, we need some security bounds:

$$\text{risk}_{r,t} = \min(0.2, \text{risk}_{r,t0} \min(1.25, \max\left(0.4, \left(1 - \frac{\text{debt}_{r,t}}{\text{gdp}_{r,t}}\right)^8\right))) \quad (72)$$

Note that lenders to the bank also receive an update of their risk parameter, i.e. they become more attractive for foreign investments. It is obvious that the thresholds and the elasticity of eight are not the outcome of econometric work but of a trial-and-error approach which various aggregations.

A practical issue emerged when the mechanism above was applied to some special cases, where foreign savings account for a large share of investments or total final consumption. Examples are some developing countries, receiving large amounts of development aid or remittances, but also “tax havens” such as Malta. In such cases, we noticed that the mechanism above can lead, after some periods, to a situation where regional income gets unrealistically small. To avoid such extreme cases, while allowing for the existence of capital inflows or outflows determined by factors other than expected returns, we introduced a regional share parameter, such that only part of the debt may actually be served (see the Technical Annex for more details).

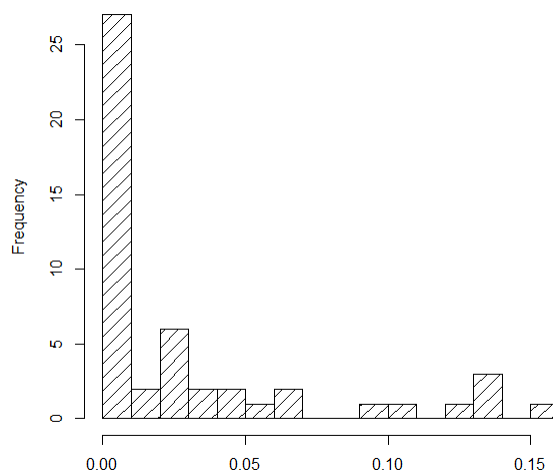
On the position “debt” the cumulated debt is defined: previous year debt level (which can be positive or negative) plus foreign saving in the current year (again positive or negative). The debt servings in the current year are defined in the “paym” column: the foreign savings on all previous years times expected returns – assuming hence a return fixed when the investment is made – times the size of the time step. Note the scaling of  $p\_capTrans$  to ensure that the global expenditures serving past debt and related revenues add up to zero.

### 3.14.8 Cost-share adjustment

The question if and to what degree the cost share of total intermediate demand factors changes in the course of economic development is an empirical question, whose answer depends on many factors, such as the industrial composition, the price structure, etc.. Analogously, cost shares might be subject to change for individual sectors. We therefore use again the GTAP 9/10 Data base for a statistical analysis, taking from there data on cost shares for the intermediate composite in the 57/65 GTAP industries and regional GDP per capita.

We exclude values which were 2.5 times the inter-quartile range away from the median. The estimations were performed on the log of the intermediate costs share relative to the log of income per capita and the square of log income per capita to capture potential saturation effects or turning points. We used a model selection procedure to remove insignificant coefficients. Population acted as regression weights to avoid that smaller countries with similar income levels, such as e.g. in case of the EU27 or partly for Africa, carry too much weight.

We find that the link between GDP per capita and the intermediate cost share can be generally confirmed (see Figure 12). However, for nine industries, the model selection process dropped the relation altogether while the remaining coefficients are mostly highly significant (see Table 13).



**Figure 12 : Significance level of regression coefficients of per capita income for cost share of intermediate composite**

There is however no linear trend connecting economic growth to the share of intermediate costs. Our regression reveals that up to around 5.000 USD per capita, the share of intermediate costs in the economy normally increases (first line, “tot”, in Table 6). A possible explanation could be the growing share of manufacturing in the economy, where the share of intermediates is higher. At higher income levels, the share drops, possibly reflecting the generally lower cost shares of intermediates in service sectors.

**Table 13: Intermediate costs shares in total and for the 57 GTAP sectors, mean, max and min estimates at increasing per capita GDP level in cst USD**

|     | mean  | max   | min   | 500          | 1000  | 5000         | 10000 | 50000 | 100000       |    |
|-----|-------|-------|-------|--------------|-------|--------------|-------|-------|--------------|----|
| tot | 0,515 | 0,669 | 0,291 | 0,495        | 0,524 | <b>0,568</b> | 0,558 | 0,475 | 0,421        |    |
| pdr | 0,460 | 0,992 | 0,043 | 0,264        | 0,269 | 0,296        | 0,328 | 0,478 | <b>0,596</b> | ++ |
| wht | 0,571 | 1,000 | 0,001 | <b>0,637</b> | 0,592 | 0,499        | 0,463 | 0,390 | 0,363        | -- |
| gro | 0,425 | 0,957 | 0,001 | 0,275        | 0,303 | 0,377        | 0,414 | 0,515 | <b>0,566</b> | ++ |
| v.f | 0,361 | 0,847 | 0,002 | 0,177        | 0,200 | 0,263        | 0,296 | 0,391 | <b>0,440</b> | ++ |
| osd | 0,451 | 0,994 | 0,011 | 0,249        | 0,255 | 0,288        | 0,328 | 0,529 | <b>0,701</b> | ++ |
| c.b | 0,442 | 0,989 | 0,029 | 0,300        | 0,303 | 0,317        | 0,333 | 0,398 | <b>0,441</b> | ++ |
| pfb | 0,535 | 0,998 | 0,044 | 0,286        | 0,341 | 0,472        | 0,497 | 0,458 | <b>0,405</b> | ** |
| ocr | 0,325 | 0,869 | 0,008 | 0,227        | 0,230 | 0,245        | 0,263 | 0,337 | <b>0,391</b> | ** |
| ctl | 0,520 | 0,989 | 0,046 | 0,378        | 0,386 | 0,430        | 0,483 | 0,737 | <b>0,946</b> | ++ |
| oap | 0,546 | 0,984 | 0,066 | 0,362        | 0,402 | 0,513        | 0,570 | 0,727 | <b>0,807</b> | ++ |
| rmk | 0,546 | 0,999 | 0,106 | 0,276        | 0,315 | 0,427        | 0,488 | 0,662 | <b>0,756</b> | ++ |
| wol | 0,765 | 0,996 | 0,173 | 0,816        | 0,719 | 0,589        | 0,598 | 0,788 | <b>0,982</b> | ++ |
| frs | 0,399 | 0,986 | 0,002 | 0,142        | 0,171 | 0,259        | 0,311 | 0,472 | <b>0,566</b> | ++ |
| fsh | 0,481 | 0,952 | 0,020 | 0,248        | 0,277 | 0,359        | 0,402 | 0,522 | <b>0,584</b> | ++ |
| coa | 0,561 | 1,000 | 0,037 | 0,480        | 0,419 | 0,348        | 0,372 | 0,608 | <b>0,869</b> | +  |
| oil | 0,433 | 1,179 | 0,027 | <b>0,514</b> | 0,399 | 0,255        | 0,246 | 0,325 | 0,429        | -  |
| gas | 0,471 | 1,265 | 0,018 | 0,320        | 0,320 | 0,320        | 0,320 | 0,320 | 0,320        |    |
| omn | 0,528 | 1,000 | 0,071 | 0,411        | 0,427 | 0,466        | 0,484 | 0,528 | <b>0,548</b> | ++ |
| cmt | 0,729 | 0,995 | 0,307 | 0,736        | 0,736 | 0,736        | 0,736 | 0,736 | 0,736        |    |
| omt | 0,726 | 1,008 | 0,223 | 0,763        | 0,763 | 0,763        | 0,763 | 0,763 | 0,763        |    |
| vol | 0,784 | 0,999 | 0,358 | 0,813        | 0,813 | 0,813        | 0,813 | 0,813 | 0,813        |    |
| mil | 0,752 | 0,947 | 0,440 | 0,726        | 0,747 | 0,784        | 0,787 | 0,761 | 0,737        | +  |
| pcr | 0,771 | 1,000 | 0,224 | 0,728        | 0,754 | <b>0,780</b> | 0,753 | 0,615 | 0,536        |    |
| sgr | 0,662 | 1,006 | 0,141 | 0,684        | 0,722 | <b>0,760</b> | 0,720 | 0,530 | 0,430        | -  |
| ofd | 0,690 | 0,968 | 0,417 | 0,738        | 0,751 | <b>0,756</b> | 0,732 | 0,625 | 0,563        | -  |
| b.t | 0,603 | 0,897 | 0,146 | 0,565        | 0,599 | <b>0,656</b> | 0,649 | 0,565 | 0,506        | -  |
| tex | 0,664 | 0,952 | 0,272 | 0,693        | 0,710 | <b>0,724</b> | 0,704 | 0,605 | 0,546        | -  |
| wap | 0,632 | 0,974 | 0,237 | <b>0,726</b> | 0,710 | 0,673        | 0,658 | 0,624 | 0,610        | -- |
| lea | 0,669 | 1,000 | 0,275 | <b>0,713</b> | 0,710 | 0,693        | 0,677 | 0,619 | 0,587        | -- |
| lum | 0,680 | 0,979 | 0,295 | 0,625        | 0,646 | <b>0,683</b> | 0,681 | 0,637 | 0,603        | -  |
| ppp | 0,650 | 0,950 | 0,339 | 0,692        | 0,708 | <b>0,714</b> | 0,685 | 0,557 | 0,486        | -  |
| p.c | 0,860 | 1,374 | 0,418 | <b>0,939</b> | 0,935 | 0,915        | 0,894 | 0,820 | 0,780        | -- |
| crp | 0,732 | 1,000 | 0,395 | <b>0,772</b> | 0,769 | 0,750        | 0,730 | 0,663 | 0,626        | -  |
| nmm | 0,657 | 1,000 | 0,235 | <b>0,729</b> | 0,705 | 0,653        | 0,631 | 0,584 | 0,565        | -- |
| i.s | 0,778 | 0,998 | 0,411 | <b>0,795</b> | 0,792 | 0,774        | 0,757 | 0,694 | 0,660        |    |
| nfm | 0,764 | 0,999 | 0,457 | <b>0,792</b> | 0,785 | 0,769        | 0,762 | 0,746 | 0,740        | -- |
| fmp | 0,680 | 0,998 | 0,294 | <b>0,739</b> | 0,733 | 0,703        | 0,672 | 0,570 | 0,518        | -- |
| mvh | 0,723 | 1,000 | 0,270 | 0,750        | 0,750 | 0,750        | 0,750 | 0,750 | 0,750        |    |
| otn | 0,679 | 1,070 | 0,236 | <b>0,734</b> | 0,716 | 0,677        | 0,661 | 0,625 | 0,610        | -- |

|     | mean  | max   | min   | 500          | 1000  | 5000         | 10000        | 50000 | 100000       |    |
|-----|-------|-------|-------|--------------|-------|--------------|--------------|-------|--------------|----|
| ele | 0,694 | 0,999 | 0,197 | <b>0,857</b> | 0,809 | 0,724        | 0,709        | 0,716 | 0,739        | -  |
| ome | 0,686 | 0,999 | 0,240 | <b>0,801</b> | 0,764 | 0,686        | 0,655        | 0,588 | 0,561        | -- |
| omf | 0,679 | 0,999 | 0,167 | <b>0,752</b> | 0,711 | 0,635        | 0,615        | 0,596 | 0,599        | -- |
| ely | 0,672 | 1,004 | 0,119 | 0,638        | 0,674 | <b>0,711</b> | 0,672        | 0,491 | 0,396        | -  |
| gdt | 0,529 | 1,194 | 0,101 | 0,343        | 0,404 | 0,547        | <b>0,572</b> | 0,522 | 0,461        | +  |
| wtr | 0,463 | 0,951 | 0,002 | 0,422        | 0,422 | 0,422        | 0,422        | 0,422 | 0,422        |    |
| cns | 0,601 | 0,858 | 0,243 | 0,594        | 0,609 | <b>0,621</b> | 0,605        | 0,521 | 0,471        | -  |
| trd | 0,429 | 0,842 | 0,083 | 0,296        | 0,316 | 0,367        | 0,391        | 0,455 | <b>0,485</b> | ++ |
| otp | 0,603 | 0,986 | 0,255 | 0,520        | 0,533 | 0,564        | 0,579        | 0,613 | <b>0,628</b> | ++ |
| wtp | 0,710 | 0,988 | 0,291 | 0,509        | 0,532 | 0,591        | 0,618        | 0,686 | <b>0,718</b> | ++ |
| atp | 0,733 | 1,051 | 0,324 | 0,583        | 0,610 | 0,678        | 0,710        | 0,789 | <b>0,826</b> | ++ |
| cmn | 0,393 | 0,702 | 0,113 | 0,271        | 0,295 | 0,361        | 0,393        | 0,481 | <b>0,524</b> | ++ |
| ofi | 0,383 | 0,912 | 0,000 | 0,232        | 0,236 | 0,260        | 0,287        | 0,417 | <b>0,519</b> | ++ |
| isr | 0,461 | 0,911 | 0,000 | 0,253        | 0,308 | 0,450        | <b>0,487</b> | 0,482 | 0,441        | +  |
| obs | 0,381 | 0,740 | 0,079 | 0,383        | 0,383 | 0,383        | 0,383        | 0,383 | 0,383        |    |
| ros | 0,447 | 0,964 | 0,025 | 0,405        | 0,407 | 0,419        | 0,432        | 0,483 | <b>0,516</b> | ++ |
| osg | 0,332 | 0,659 | 0,025 | 0,232        | 0,245 | 0,280        | 0,296        | 0,338 | <b>0,358</b> | ++ |
| dwe | 0,142 | 0,483 | 0,000 | 0,090        | 0,090 | 0,090        | 0,090        | 0,090 | 0,090        |    |

Note: +/- depicts the cases the estimate cost share at the highest income is higher/lowest than at the lowest  
 ++/-- depicts cases where the share is estimated to increase/decrease over the whole income range

On the other hand, our estimates provide rather clear results for agricultural activities: with the exception of wheat (wht), all agro-industries show increasing cost shares of intermediates at higher levels of per capita income. For animal processes, the differences in the shares are more pronounced compared to crops. Similarly, there are strong increases in intermediate shares for forestry (frs), fisheries (fsh) and coal mining (coa), and moderate ones for “other minerals” (omn), whereas the shares are dropping for oil extraction and no significant relation was found for the gas extraction.

Contrary to primary sectors, almost no manufacturing industry displays increasing shares. Food and meat processing (cmt, omt) as well as oil seed crushing (vol) show no clear relation to income. For the remaining food processing sectors, the intermediate cost shares peak at around 5.000 USD. For the rest of manufacturing, the strongest are typically found at the low income levels, although differences in the shares are often not very large.

There is also no uniform tendency for electricity, gas and water distribution activities (ely, gas, wtr) and construction (cns), where intermediate shares peak at medium income levels. Interestingly, results for all service industries highlight a clear direction: intermediate shares increase at higher per capita income levels. Especially high shares are found in the transportation services (otp, wtp, atp), which are around 80%, while the remaining services typically reach values around 50%.

If preferences are a function of income per capita, reflected in non-linear Engel curves, then the portfolio of products offered by the economy clearly changes. As Chenery et al. 1986 put it “On the demand side, a rise in income can only be sustained if the goods and services made available correspond to the proportions in which consumers wish to spend their income”. We already addressed this issue for the final demand through the introduction of an AIDADS demand system, but further adjustments are in order on the production side, to account for income-dependent variations in

intermediate demand. Indeed, an often-neglected aspect in CGE and input-output models is that industries internally include many diverse production processes, characterized by different technologies. Variations in demand patterns therefore occur not only between the macro-industries, but also inside them: aggregate industrial cost structures should be better interpreted as reflecting the internal composition of a sector, rather than describing the production function of a representative firm. Consequently, input-output coefficients can well evolve over time, following changes in income, prices, foreign trade, demography, etc., in a way not too different from the one affecting household final consumption. In parallel, processes such as capital accumulation and the related shift to more capital-based production technologies can systematically affect the cost shares in certain industries.

Already Arrow and Hoeffenberg, 1959<sup>66</sup> decomposed changes in input-output coefficients into variations due to real disposable income and variations due to technology and tastes. Skolka, 1989 further provided a structural decomposition analysis for Austria along these lines, thereby explicitly considering that I-O coefficients are not static, but actually change along the process of economic development. This contrasts with the approach followed in most dynamic CGE models, where changes in the industrial cost shares are only attributed to two causes: non-Hicksian technological progress and changes in relative prices.

As an example, consider the “Other food” sector in GTAP, comprising a wide range of processing activities of crop and fish-based products. Here the ICP data set, which was used to estimate the AIDADS demand system, includes expenditure data of subcategories such as “bread”, “Other bakery products”, “Pasta products”, “Jams, marmalades and honey”, “Confectionery, chocolate and ice cream” etc.. A simple regression on the shares of more disaggregated data relative to totals reveals that they typically are income dependent. As some of these categories require different inputs in production and are based on different technological processes, also the I-O composition of the “Other food” industry should hence be income dependent.

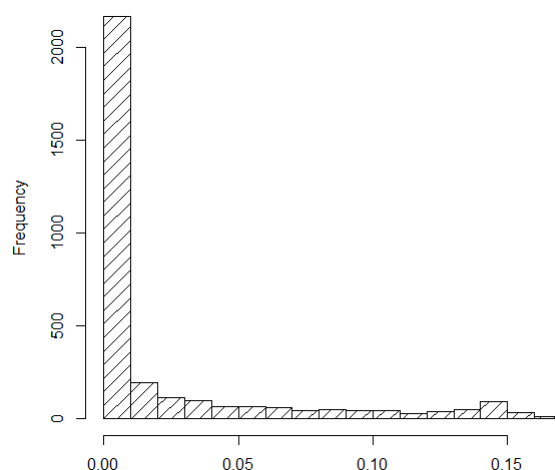
Therefore, I-O coefficients should be not considered as constant in the long-term, where income varies significantly. Since the model already accounts for price-induced compositional changes in intermediate demand, and possibly Hicksian non-neutral technical progress, we include in G-RDEM the modelling of income-related variations.

Our basic hypothesis is that I-O coefficients are income dependent, likewise final consumption shares. Since time series consistent with the GTAP industrial classification are not available, we test our hypothesis using a sectional approach, using once more the GTAP 9 data base, as we did for the intermediate composite. From the 3.249 I-O coefficients (57x57) in the data base, the model selection process filtered a significant relation to GDP per capita in as many as 3.213 of them. The histogram in Figure 9, showing the significance levels, highlights that for more than two thirds of elements in the sample they are significant at 1% level or below.

---

<sup>66</sup> Arrow, K. J. (1959). A time series analysis of interindustry demands (Vol. 17). North-Holland





**Figure 13 : Significance level of regression coefficients of per capita income for individual input coefficients**

Consider the agricultural “coarse grains” industry as an example, and its six largest cost shares (see Table 14). For two of them, corresponding to seeds and transport, there is no clear income dependence. Services tend to grow with higher per capita income, whereas shares of chemical and petroleum products decrease. Coarse grains could hence be seen as an example where economic development, through changes in technology – intensification in crop production by using more fertilizers, machinery etc. – and outsourcing of activities (contract work) drive differences in costs shares.

**Table 14: Main cost shares in coarse grains as a function of per capita income**

|             | mean  | max   | min   | 500   | 1000  | 5000  | 10000 | 50000 | 100000 |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| gro.a_crp.c | 0,173 | 0,666 | 0,000 | 0,113 | 0,138 | 0,199 | 0,210 | 0,187 | 0,161  |
| gro.a_trd.c | 0,048 | 0,611 | 0,000 | 0,041 | 0,041 | 0,047 | 0,053 | 0,083 | 0,109  |
| gro.a_gro.c | 0,043 | 0,315 | 0,000 | 0,043 | 0,043 | 0,043 | 0,043 | 0,043 | 0,043  |
| gro.a_p.c.c | 0,030 | 0,400 | 0,000 | 0,008 | 0,014 | 0,044 | 0,056 | 0,051 | 0,037  |
| gro.a_otp.c | 0,027 | 0,118 | 0,000 | 0,027 | 0,027 | 0,027 | 0,027 | 0,027 | 0,027  |
| gro.a_obs.c | 0,012 | 0,217 | 0,000 | 0,002 | 0,005 | 0,018 | 0,028 | 0,049 | 0,052  |
| gro.a_ofi.c | 0,008 | 0,071 | 0,000 | 0,001 | 0,003 | 0,014 | 0,018 | 0,012 | 0,007  |
| Sum         |       |       |       | 0,233 | 0,267 | 0,378 | 0,416 | 0,439 | 0,428  |

Integration of the regression results in the model is not a trivial task, as shares found at the base year may deviate considerably from the estimates. One option is to realign estimates and data by adding an error term to the intercept in the regressions, while considering also the effect of changes.

This is necessary because, for example, absolute changes may lead to negative shares when the shares are small. On the other hand, relative changes would keep small shares forever small, much like as it happens for trade shares in the Armington formulation.

We therefore opt here for an approach where the error terms  $\text{err}$  are adjusted on the basis of the following equation, where  $\text{inc}$  stands for GDP per capita in the benchmark 0 and in the current simulation year  $t$ , whereas  $\text{share}$  refers to the share parameter for the intermediate composite:

$$and_{corr} = and_{est} + err \left( \max \left( 0, 1 - \frac{1}{2} abs \left[ \frac{inc_t - inc_0}{inc_0} \right] \right) \right) \quad (11)$$

The idea behind the correction is simple: the farther away the given GDP per capita from the benchmark, the less structural information is provided by the initial SAM. Specifically, equation (11) defines a relation where at an income growth of 50%, 75% of the error term will be added. If real income doubles, 50% is added and if income triples, the estimates are directly used.

A further complication is due to the fact that the cost share is not a given parameter but reflects price changes for the output and the intermediate input composite. The price variations for the intermediate composite (pnd) and the unit cost of output (px) need to be considered in the model. The cost share parameter and is therefore defined as:

$$and_{fin} = and_{corr} \frac{px}{pnd} \quad (12)$$

Table 8 illustrates the approach. The estimated cost share at benchmark mark income is about 27% (row Econometric estimate), while the share found in the SAM is about 33.8% (row AND parameter). That results in an error term of around 6%. In the first iteration period (t04), the estimate increases only slightly, as also GDP per capita is not changing much. Accordingly, around 97% of the error term is kept, and the cost share only drops slightly towards the lower econometric estimate. The cost share parameter is further corrected stronger as the per unit price decreases to 0.93, lower than the intermediate composite price index (0.97). In year 28 of the simulation, GDP has tripled and the correction factor drops to zero. Accordingly, from there onwards, the cost share of the intermediate composite (row ND cost share) is identical to the econometric estimate at each iteration round.

**Table 15: Example of cost share parameter adjustment in simulation**

|                      | t00  | t01  | t04  | t08  | t12  | t16  | t20  | t24  | t28  | t32  | t36  | t40  |
|----------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| Econometric estimate |      | 0,27 | 0,27 | 0,27 | 0,27 | 0,28 | 0,28 | 0,28 | 0,29 | 0,30 | 0,31 | 0,31 |
| pnd                  |      | 1,00 | 0,98 | 0,81 | 0,73 | 0,69 | 0,66 | 0,68 | 0,67 | 0,70 | 0,75 | 0,82 |
| px                   |      | 1,00 | 0,93 | 0,80 | 0,76 | 0,77 | 0,78 | 0,88 | 0,93 | 1,03 | 1,14 | 1,29 |
| Error term           |      | 0,06 | 0,06 | 0,06 | 0,06 | 0,06 | 0,06 | 0,06 | 0,06 | 0,06 | 0,06 | 0,06 |
| Correction factor    |      | 1,00 | 0,97 | 0,89 | 0,74 | 0,55 | 0,31 | 0,02 |      |      |      |      |
| GDP                  |      | 1,42 | 1,42 | 1,49 | 1,73 | 2,14 | 2,69 | 3,38 | 4,19 | 5,19 | 6,30 | 7,49 |
| AND parameter        | 0,33 | 0,33 | 0,31 | 0,32 | 0,33 | 0,34 | 0,35 | 0,37 | 0,40 | 0,44 | 0,46 | 0,49 |
| ND cost share        |      | 0,33 | 0,33 | 0,32 | 0,31 | 0,31 | 0,30 | 0,29 | 0,29 | 0,30 | 0,31 | 0,31 |

The estimated coefficients are stored in the GDX file “gams/gtapRdem/io\_regr\_res.gdx” on the parameter p\_estres. As a first step, we remove any non-availables (NA):

```
p_estRes(iDat,jDat,regrCoef) $ (mapVal(p_estRes(iDat,jDat,regrCoef)) eq 5) = 0;
p_estRes(iDat,"nd.share",regrCoef) $ (mapVal(p_estRes(iDat,"nd.share",regrCoef)) eq 5) = 0;
```

Next, we deduct the estimate at benchmark per capita income and store it temporarily on the parameter p\_corr from the given io-coefficient

```

p_corr(r,a,i) $ (io(r,i,a,"%t0%")
    $ sum((mapa(iDat,a),mapi(jDat,i)),p_estRes(iDat,jDat,"log_GDP") or p_estRes(iDat,jDat,"sqr_log_GDP")))
    = io(r,i,a,"%t0%")
    - exp[sum((mapa(iDat,a),mapi(jDat,i)),p_estRes(iDat,jDat,"int")
    + sum(r_r(r,rNat),p_estRes(iDat,jDat,"log_GDP") * log(p_perCapIncome(rNat,"%t0%"))
    + p_estRes(iDat,jDat,"sqr_log_GDP") * sqr(log(p_perCapIncome(rNat,"%t0%")))
    * sign(log(p_perCapIncome(rNat,"%t0%")))))]];

```

Similar statements perform the same operation for io-coefficients in technology nests and for the intermediate composite cost share. During simulation, we first calculate the error correction factor `p_errCorr` as a function of per capita income in the current simulation point `tSim` and the benchmark `t0`, as discussed above:

```

p_errCorr(rNat) = max(0,1 - 0.5*abs((p_perCapIncome(rNat,tSim)-p_perCapIncome(rNat,"%t0%"))
    /p_perCapIncome(rNat,"%t0%")));

```

Next, the calculate the estimate at the current income

```

and(r,a,tSim) $ (and(r,a,tSim-1) $ p_corr(r,a,"nd"))
    = exp[sum((mapa(iDat,a),p_estRes(iDat,"nd.share","int")
    + sum(r_r(r,rNat),p_estRes(iDat,"nd.share","log_GDP") * log(p_perCapIncome(rNat,"%t0%"))
    + p_estRes(iDat,"nd.share","sqr_log_GDP") * sqr(log(p_perCapIncome(rNat,"%t0%")))
    * sign(log(p_perCapIncome(rNat,"%t0%")))))]];

```

Adding subsequently a part of the error term `p_corr` at the benchmark according to the factor `p_errCorr` and introduce security bounds:

```

and(r,a,tSim) $ (and(r,a,tSim-1) $ p_corr(r,a,"nd"))
    = min(and(r,a,"%t0%")*2,1-ava(r,a,"%t0%")*0.5,and(r,a,tSim-1)*(1+0.1*p_timeStep(tSim)),
    max(and(r,a,"%t0%")*0.5,and(r,a,tSim-1)*(1-0.1*p_timeStep(tSim)),
    {and(r,a,tSim) + p_corr(r,a,"nd")*p_errCorr(r)}
    *px(r,a,tSim)/px(r,a,"%t0%")*1/pnd(r,a,tSim)
    ));

```

The value-added share is calculated residually:

```

ava(r,a,tSim) = 1 - and(r,a,tSim);

```

Note here:

1. The application of the estimation is not allowed to more than double or half the benchmark coefficient.
2. The coefficients are not allowed to change by more then 0.1 time the number of simulated years. For individual IO-cost shares, the changes are limited to 0.025 per year.
3. The estimated error terms are introduced in the update process as well such that the for the first year `tSim=1`, per capita income for the benchmark will enter, such that the last term will yield exactly the observed IO coefficient. The process hence only adds the changes from year to year in the estimated to the benchmark coefficient.

Similar statements estimate and apply shifters for the other elements. The IO-cost shares are scaled to maintain the original sum:

```

p_scale(r,a,"") $ sum(j,io(r,j,a,tSim))
    = sum(j,io(r,j,a,"%t0%"))/sum(j,io(r,j,a,tSim));

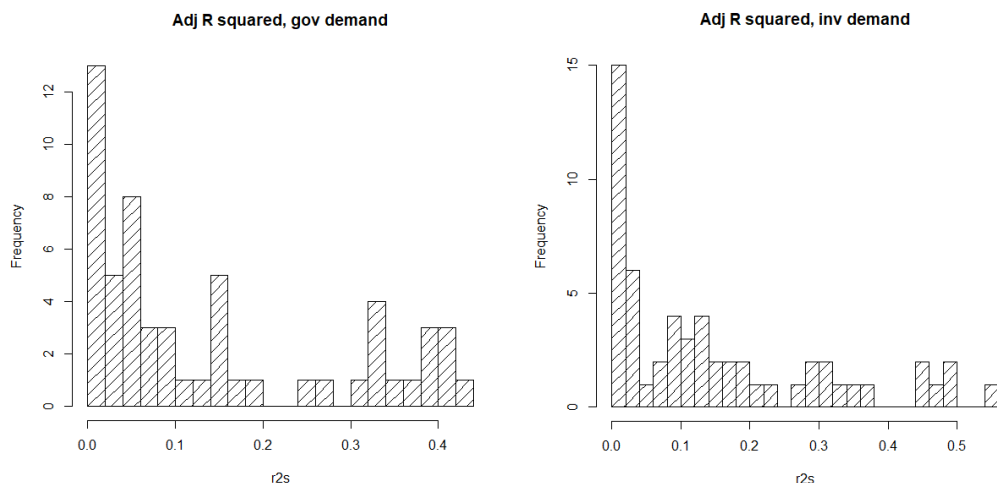
io(r,i,a,tSim) $ io(r,i,a,tSim-1) = io(r,i,a,tSim) * p_scale(r,a,"");

```

### 3.14.9 Endogenous expenditure shares for government and investment demand

Similar to the discussion around industry cost shares depending on economic development, the composition of investment demand might change in the process of economic growth reflecting both a different sectoral composition of the economy and different capital stock composition in each sector. For instance, the share of investments in patents or similar might increase with higher income levels. In the same vein, the composition of government demand might change.

In both cases, we follow an empirical approach by employing regression analysis to the cross-sectional data provided by the GTAP SAM. We use expenditure shares for individual commodities on total investment respectively government demand as dependent variables, and the logarithm of income and its squares as explanatory ones. In case of government expenditure, we also use its total share on regional income, linear and in squares, as independents. A model selection process removes insignificant regressors from the estimation equation to avoid spurious results. We use total population as weights.



The integration into the simulation framework follows the approach used for industry cost shares. We use a weighted average of the expenditure share at the benchmark and the estimates, defining weights based on the relative difference of the projected per capita income from the benchmark. The closer the projection to the benchmark, the closer the expenditure share is to the one found in the SAM while large deviations will drive the shares towards their estimates.

Technically, in order to prepare for simulation, we need therefore the error terms for the current regional and sectoral aggregation. We therefore first define the estimates, temporarily on the parameter `p_corrInv`, using the per capita income `p_perCapIncome` at the benchmark “%t0%” on the RHS together with the parameters found on `p_estResInv`. The estimations are performed on the 57 sector resolution of the data base as depicted by the set `iDat`. In case of an aggregated data base being used for simulation, the cross set `mapi` depicts which of these 57 sectors are aggregated to the sector list in use `i` such that estimated shares are aggregated. The regression results are only used if income dependencies were found, i.e. at least one of the regression coefficients is non-zero:

```
p_corrInv(rNat,i) $ (alphaa.1(rNat,i,"inv","%t0%")
  $ sum((mapi(iDat,i)),p_estResInv(iDat,"log_GDP") or p_estResInv(iDat,"sqr_log_GDP")))
= sum((mapi(iDat,i)),p_estResInv(iDat,"int")
  + p_estResInv(iDat,"log_GDP")*log(p_perCapIncome(rNat,"%t0%"))
  + p_estResInv(iDat,"sqr_log_GDP")*sqr(log(p_perCapIncome(rNat,"%t0%")))
  * sign(log(p_perCapIncome(rNat,"%t0%")))*p_scaleInvToSplit(rNat,i);
```

From there, we define the error term at the benchmark as the difference to the benchmark share `p_alphaa.l`.

```
p_corrInv(rNat,i) $ (alphaa.l(rNat,i,"inv","%t0%") $ p_corrInv(rNat,i))
= alphaa.l(rNat,i,"inv","%t0%") - p_corrInv(rNat,i);
```

The amount of the error applied in year `tSim` is defined as follows (see industry cost shares for a discussion):

```
p_errCorr(rNat) = max(0,1 - 0.5*abs((p_perCapIncome(rNat,tSim)-p_perCapIncome(rNat,"%t0%"))
/p_perCapIncome(rNat,"%t0%")));
```

During simulation, we first define the estimated shared at the projected per capita income, the expression is identical to the one discussed above with the exemption that the current simulation year `tSim` is used instead of the benchmark point “%t0%”:

```
alphaa.l(rNat,i,"inv",tSim) $ (alphaa.l(rNat,i,"inv","%t0%")
$ sum((mapi(iDat,i)),p_estResInv(iDat,"log_GDP") or p_estResInv(iDat,"sqr_log_GDP")))
= sum((mapi(iDat,i)),p_estResInv(iDat,"int")
+ p_estResInv(iDat,"log_GDP") * log(p_perCapIncome(rNat,tSim))
+ p_estResInv(iDat,"sqr_log_GDP") * sqrt(log(p_perCapIncome(rNat,tSim)))
* sign(log(p_perCapIncome(rNat,tSim)))) * p_scaleInvToSplit(rNat,i);
```

Shares comprised in final demand nests `dNest` are aggregated to yield the share of the nest on expenditures:

```
alphaDn(rNat,dNest,"inv",tSim) = sum(dNest_i_fd(dNest,i,"inv"), alphaa.l(rNat,i,"inv",tSim));
```

Next, we add the share part `p_errCorr` of the error term `p_corrInv` at the benchmark to the estimate, that is the expression in the last line in curly brackets. If `p_errCorr` is equal to one, the full error term is added the the last line is equal to the benchmark share. As `p_errCor` decreases to zero, the expression in curly brackets approaches the regression the results. The surrounding min and max operators define security bounds preventing that expenditure shares more than double or half or increase by more than 5% from year to year:

```
alphaa.l(rNat,i,"inv",tSim) $ p_corrInv(rNat,i)
= min(alphaa.l(rNat,i,"inv","%t0%")*2,alphaa.l(rNat,i,"inv",tSim-1)*(1+0.05*p_timeStep(tSim)),
max(alphaa.l(rNat,i,"inv","%t0%")*0.5,alphaa.l(rNat,i,"inv",tSim-1)*(1-0.05*p_timeStep(tSim)),
{alphaa.l(rNat,i,"inv",tSim) + p_corrInv(rNat,i)*p_errCorr(rNat)}
));
```

Similar, expenditure share of demand nest `dNest` are defined:

```
alphaDn(rNat,dNest,"inv",tSim) $ p_corrInv(rNat,dNest)
= min(alphaDn(rNat,dNest,"inv","%t0%")*2,alphaDn(rNat,dNest,"inv",tSim-1)*(1+0.05*p_timeStep(tSim)),
max(alphaDn(rNat,dNest,"inv","%t0%")*0.5,alphaDn(rNat,dNest,"inv",tSim-1)*(1-0.05*p_timeStep(tSim)),
{alphaDn(rNat,dNest,"inv",tSim) + p_corrInv(rNat,dNest)*p_errCorr(rNat)}
));
```

Finally, we scale the shares to maintain the original sum:

```

--- scale to maintain original sum of investment cost shares

p_scaleInv(rNat) = [ sum(notInvInNest(i),alphaa.1(rNat,i,"inv","%t0%"))
+ sum(dNest, alphaDn(rNat,dNest,"inv","%t0%")) ]
/[ sum(notInvInNest(i),alphaa.1(rNat,i,"inv",tSim))
+ sum(dNest, alphaDn(rNat,dNest,"inv",tSim)) ];

alphaa.fx(rNat,i,"inv",tSim) $ (notInvInNest(i)) = alphaa.1(rNat,i,"inv",tSim) * p_scaleInv(rNat);
alphaDn(rNat,dNest,"inv",tSim) = alphaDn(rNat,dNest,"inv",tSim) * p_scaleInv(rNat);

```

The code for the government share is structurally identical.

The impact on the simulated results seems of lower importance compared to other elements in G-RDEM. This firstly reflects that investment demand is generally a minor component of overall demand. Furthermore, construction (“cns”) dominates with shares typically around 55% investment expenditure shares and is found to be little dependent on income, while many other products have very low shares. For government, a very large share is for government services, the share of government demand for other products on total demand is generally quite low.

### 3.14.10 Reovering crop yield projections

The sector specific productivity growth in G-RDEM would imply the also land productivity would follow the estimated relations, potentially leading to quite unrealistic partial productivities, i.e. crop yields.

During baseline construction, the partial land productivity for GTAP crop activity  $a$  in model region  $r$  and simulation year  $t$  is fixed to the yield development of the FAO projections time the benchmark productivity:

$$\frac{xp_{r,a,t}}{xf_{r,lnd,a,t}} = \frac{faoProj_{r,a,t}^{yld}}{faoProj_{r,a,t0}^{yld}} \frac{xp_{r,a,t}}{xf_{r,lnd,a,t}} \quad (73)$$

Were  $xp$  is output and  $xf$  factor use,  $lnd$  refers to land, and  $faoYldProj$  to the yield projection by the FAO in  $t \text{ ha}^{-1}$ . Fixing the yield is made possible by endogenously adjusting the Hoicks-non neutral land productivity shifter for  $r,a,t$  in each simulated point.

The model equations in *gams/GTAPRdem/fao\_baseline.gms* use the projections *p\_faoProjections*, considering the partial land productivity *p\_agrYield* derived from the benchmark SAM, with an correction factor depicted by *v\_yieldDriver*:

```

!
*
* --- target yield: FAO projections plus price driven corrections. If elasticities are zeros or price don't
* change, the FAO projections will be exactly recovered
*
* p_agrYield(rs,lnd,a,ts(t)) $ (p_agrYield(rs,a) $ (not disr(rs)) $ xfFlag(rs,lnd,a) $ (not sameas(ts,"%t0%")))..
*
* --- define output as exogenous yields (from FAO projection) times land used in the activity,
* lambdaf for land will endogenously adjust
*
* m_xp(rs,a,ts)s(xp,'rs,a,ts')
* =E= [sum((r_r(rs,rsNat),faoYears) $ (faoYears.val eq t.val),
* --- relative fao projections times xf/xp at benchmark
* p_faoProjection(rsNat,a,"yield",faoYears)/p_faoProjection(rsNat,a,"yield","%baseYear%") * p_agrYield(rs,a))
*
* ---- price effects
*
* + v_yieldDriver(rs,a,ts)-1
* ] * xf(rs,lnd,a,ts)s(xp,'rs,a,ts');

```

Note that in the equation, the definition of the partial productivity  $xp/xf$  is changed into a linear one by moving  $xf$  on the LHS.

The yield driver variable  $v\_yieldDriver$  uses up to three elasticities:  $p\_yieldElasR$  looks that the price relation between a crop's unit crop relative to the average unit cost of all crops  $v\_cropPrice$ , while  $p\_yieldElasL$  reacts to the land return of the crop of, normalized with price index of GDP  $pgdpmp$  and finally,  $p\_yieldElasO$ , reacts to changes in the overall crop price index.

```
*
*   --- Define price depending shifts in yield that correct the FAO yield projections
*
e_yieldDriver(rs,a,ts(t)) $ (p_agrYield(rs,a) $ (not disr(rs)) $ xfFlag(rs,lnd,a) $ (not sameas(ts,"%t0%"))) ..
    v_yieldDriver(rs,a,ts)
    =E= sum(r_r(rs,rsNat),
*
*           --- add a dampening term which prevents larger differences across crop prices
*
*       + (px(rs,a,ts)/(v_cropPrice(rsNat,ts)*px.l(rs,a,"%t0%")))**p_yieldElasR(rs,a)-1
*
*       --- add (on demand) an elasticity to land returns
*
*       + (pf(rs,lnd,a,ts)/pgdpmp(rsNat,ts))**p_yieldElasL(rs,a)
*
*       --- effect of overall crop price level
*
*       + (v_cropPrice(rsNat,ts)/(v_cropPrice.l(rsNat,"%t0%")*pgdpmp(rsNat,ts)))**p_yieldElasO(rs,a)-1);
```

The average crop price index is a output weighted average of the unit cost:

```
--- crop price index
e_cropPrice(rsNat,ts(tRun)) $ ( v_cropPrice.l(rsNat,"%t0%") $ (not sameas(ts,"%t0%"))) ..
    v_cropPrice(rsNat,ts)
    =E= sum((rs,a) $ (p_agrYield(rs,a) $ r_r(rs,rsNat)), [xp.l(rs,a,tRun-1)+xp.l(rs,a,"%t0%") $ t0(ts)]*px(rs,a,ts) )
    /sum((rs,a) $ (p_agrYield(rs,a) $ r_r(rs,rsNat)), [xp.l(rs,a,tRun-1)+xp.l(rs,a,"%t0%") $ t0(ts)]*px.l(rs,a,"%t0%"));
```

The crop yield projections from FAO 2018<sup>67</sup> are mapped to GTAP based on a mapping between the GTAP sectors and the ones in the FAO:

---

<sup>67</sup> The future of food and agriculture – Alternative pathways to 2050. Rome. 224 pp. Licence: CC BY-NC-SA 3.0 IGO



```

set iOri_ac(%iOri%,faoCrops) "Link between GTAP sectors/products and crops covered by FAO projection" /

pdr.A_PARI
wht.A_WHEA
gro.A_MAIZ
gro.A_BARL
gro.A_MILL
gro.A_SORG
gro.A_XCER
v_f.A_POTA
v_f.A_SWPY
v_f.A_CASS
v_f.A_XRNT
v_f.A_PULS
v_f.A_XVEG
v_f.A_PLAN
v_f.A_BANA
v_f.A_CITR
v_f.A_XFRU
c_b.A_SUGB
c_b.A_SUGC
osd.A_RAPS
osd.A_SOYB
osd.A_GRND
osd.A_SUNF
osd.A_SESA
osd.A_CCNT
osd.A_XOLT
osd.A_PALM
osd.A_OLIV
ocr.A_COCO
ocr.A_COFF
ocr.A_TEAS
ocr.A_TOBA
pfb.A_COTT
pfb.A_XFIB
ocr.A_RUBB
ocr.A_XCRO
/;

```

In case, a link to the FABIO data base is used, a second mapping defines the link between FABIO and the FAO crops, not shown in here. Jointly, the define the mapping between the current products in the data base i and the FAO crops:

```

set i_faoCrops(i,faoCrops) "Link between current commodities and FAO crops";
i_faoCrops(i,faoCrops) $ sum(mapi(%iOri%,i) $ iOri_ac(%iOri%,faoCrops),1) = YES;
i_faoCrops(i,faoCrops) $ sum((fabc_i(fabc,i),fabc_ac(fabc,faoCrops1)),1) = no;
i_faoCrops(i,faoCrops) $ sum(fabc_i(fabc,i) $ fabc_ac(fabc,faoCrops),1) = yes;

```

From this, projections at the level of the GTAP products are defined, aggregation of the system  $fs=\{\text{rainfed, irrigated}\}$  found in the FAO data base:

```

p_faoProjection(rNat,i,"1000ha",faoProj)
= sum( (r_faoReg(rNat,faoReg),i_faoCrops(i,faoCrops),fs), arable_land("%faoScen%",faoReg,faoCrops,fs,faoProj)*0.001);
*
* --- attention: output is harvested-area (not arable_land) times yield.
*
p_faoProjection(rNat,i,"1000t",faoProj)
= sum( (r_faoReg(rNat,faoReg),i_faoCrops(i,faoCrops),fs) $ yield_shift_crop_clim("%faoScen%",faoReg,faoCrops,fs,faoProj),
area_harvested("%faoScen%",faoReg,faoCrops,fs,faoProj)*(yield_crop("%faoScen%",faoReg,faoCrops,fs,faoProj)
$if(.$CCImpact "%faoCCImpact%"="off" /yield_shift_crop_clim("%faoScen%",faoReg,faoCrops,fs,faoProj)
));

```

The FAO projections are not available as a full time series, we therefore first interpolate linearly between neighboring years land use and crop output projections where projections are available:



And next smooth the results to avoid kinks:

```
--- define crop land change multiplier (1 means no change), aggregation over (potentially) sub-regions and AEZ
e_croplandChg(rsNat,ts) $ (sum( aezFlag(rs,aez) $ r_r(rs,rsNat), xaezNest.l(rs,aez,"cropland","%t0%")) $ (not sameas(ts,"%t0%"))) ..
v_croplandChg(rsNat,ts)
=E= sum( aezFlag(rs,aez) $ (xaezNest.l(rs,aez,"cropland","%t0%") $ r_r(rs,rsNat)),xaezNest(rs,aez,"cropland",ts))
/ sum( aezFlag(rs,aez) $ (xaezNest.l(rs,aez,"cropland","%t0%") $ r_r(rs,rsNat)),xaezNest.l(rs,aez,"cropland","%t0%"));
```

No further land use and management information besides the harvested and sown areas for the explicitly covered crops is reported by FAO. We hence miss information on the development of pasture lands or areas devoted to fodder crops such as alfalfa and their productivity developments. This requires further assumptions. The AEZ model assumes some transformation between arable and pasture lands which suggest that returns-to-land to these two categories cannot deviate too much from each other. When letting the model endogenously determine how much pastureland versus forestry is used at given total managed land expansion, we found in some case substantially different developments in cropland and pastureland returns. This impacts cost relations between concentrate feed and feed from pastureland, and between livestock and crop production in a hard to control manner. We therefore opted to render the land productivity in livestock production endogenous to ensure that the prices of crop and pastureland develop in parallel.

In order to project changes in total managed land driven by cropland projections, we processed the land cover statistics from FAO (FAO 2020) and run regressions per country which relate the yearly change in 1000 ha of “Primary forest”, “Managed forest” and “Meadows and pastures” on a constant term (= linear trend), a trend variable (= quadratic trend) and the yearly change in 1000 ha of “Crop land”. The sample for a country comprises all years, potentially starting in 1961, for which all variables are observed. The table below reports as an example the results for Brazil for which data are available from 1990 to 2015. The regression coefficients indicate, that on average of this period, a 1 ha change in cropland provokes a reduction of 0.13 ha of primary forest, of 0.29 ha of managed forest, of 0.54 ha of meadows and pastures and of 0.02 ha for other lands. The coefficients for the quadratic trend suggest a speeding up of the overall land conversion processes. In the last year 2017, the quadratic trend for crop land would imply  $(2017 - 1990 = 27)$  a yearly expansion of  $(-263,60 + 27 * 37 \sim 740.000)$  ha in cropland. Considering these changes in crop lands and the effects of the trends, primary forest would be reduced by  $\sim 140.000$  ha, managed ones by  $-608.000$  ha, while meadows and pastures would expand by  $335.000$  ha and other land shrink by  $321.000$  ha. This suggests a strong relation between crop land change and deforestation in this example.

**Table 16: Regression results for land use changes from FAO land cover data set for Brazil**

|                      |          | par        | pVal       | tVal       | R2         | df |
|----------------------|----------|------------|------------|------------|------------|----|
| Cropland             | const    | -263,59989 | 0,3941049  | -0,8671538 |            |    |
| Cropland             | t        | 37,1091453 | 0,06175627 | 1,95576466 |            |    |
| Cropland             |          |            |            |            | 0,13269777 | 25 |
| Other land           | Cropland | -0,0265623 | 0,83592004 | -0,20938   |            |    |
| Other land           | const    | 4350,69979 | 1,60E-17   | 22,2318278 |            |    |
| Other land           | t        | -172,32039 | 1,37E-12   | -13,333918 |            |    |
| Other land           |          |            |            |            | 0,89627975 | 24 |
| Primary Forest       | Cropland | -0,1377859 | 0,01822203 | -2,5338963 |            |    |
| Primary Forest       | const    | -1077,3289 | 3,03E-12   | -12,843341 |            |    |
| Primary Forest       | t        | 38,3359513 | 3,71E-07   | 6,92055666 |            |    |
| Primary Forest       |          |            |            |            | 0,66617658 | 24 |
| Managed forest       | Cropland | -0,2939631 | 0,00345819 | -3,2432559 |            |    |
| Managed forest       | const    | -2316,0513 | 1,23E-14   | -16,564657 |            |    |
| Managed forest       | t        | 71,2875334 | 5,89E-08   | 7,72063332 |            |    |
| Managed forest       |          |            |            |            | 0,71367956 | 24 |
| Meadows and pastures | Cropland | -0,5416888 | 6,07E-10   | -9,8927345 |            |    |
| Meadows and pastures | const    | -957,31958 | 4,04E-11   | -11,333636 |            |    |
| Meadows and pastures | t        | 62,6969057 | 4,78E-11   | 11,2399276 |            |    |
| Meadows and pastures |          |            |            |            | 0,87308303 | 24 |

If, for instance, crop land expansion was in the past accompanied by reductions of meadows and pastures, as in the example, above, a 1 ha crop land expansion will lead to a less than 1 ha expansion of natural vegetation. But the opposite can be clearly observed as well if crop and pasture land were both expanding. In order to not overly restrain the model, we only exogenize the overall change in natural cover driven by total FAO's crop land expansion estimate and the regression coefficients, and let the CET and CES nests in the GTAP-AEZ module determine changes in managed forest and pasture lands.

In order to estimate the effect of exogenous crop land change on the change in natural vegetation (primary forest and other land) for a model region and a year during baseline construction, we first aggregate the regression coefficients LUCCRegrCoeff from single country to model region, using the current cropland cover (latest data reported by FAO, at least 2015) as weights.

Another issue is the distribution of the nation-wide projected changes in crop land expansion to the different AEZ. We use as a distribution key distKey which considers the minimum of current cropland, pastureland and managed forest, the latter two divided by minus their regression coefficient to reflect expansion possibilities. If, for instance, 1 ha of cropland expansion is accompanied by 0.5 ha reduction of pastureland and only 10% pastureland per ha cropland exist, existing pastureland limits the expansion and reduces the distribution key. We add to this our estimates of the crop land buffer, divided again by the regression coefficient. Here, we add 25% of pastureland to croplands to define a minimum. In most cases, the distribution key is equal to current crop land cover.

In order to correct for cases where the interaction of the regression coefficients, the distribution keys and the exogenous projection of total crop land and crop yields leads to implausible price developments across AEZs and in total, we add two terms. The first term considers the relation between the changes in the returns to land at AEZ level relative to returns to land at the level of the model region, with an elasticity of 0.5. That means that land expansion will be stronger in AEZ with above average land rent increases. Second, we add an elasticity to changes in land rents of 0.15. These two elasticities have two important impacts. First, they imply an endogenous correction of the distribution keys, and second, they correct potentially unwanted consequences of the empirically estimated relation between crop and total land expansion on the development of land prices:

$$x_{AezNest_{r,aez,t}}^{top} = x_{AezNest_{r,aez,t0}}^{top} \left( \text{distkey}_{r,aez} \text{LUCCRegrCoeff}_{r,natVeg} \frac{faoProj_{r,tot,t}^{ha}}{faoProj_{r,tot,t0}^{ha}} + \left[ \frac{p_{AezNest_{r,aez,t}}^{top}}{p_{ft_{r,lnd,t}}} \right]^{0.5} - 1 + p_{ft_{r,lnd,t}} ** 0.25 - 1 \right) \quad (75)$$

The equation in the code looks as follows. The regression coefficients are stored on *p\_topToCrop*. Without corrections, the term above could provoke changes in total managed land that exceed the available land buffer. This is prevented by the fudged min function *ncpVuSin* that considers the 90% of the maximal buffer, corrected by changes in built-up land.

```

#- changed in total managed land (uncorrected), derived from exogenous change in crop land and regression results
e_topLandl(aezFlag(rs,aez),ts(tRun)) $ ((x_{AezNest.l(rs,aez,"top","%t0%")} gt eps) $ (not sameas(ts,"%t0%"))) ..
--- absolute change in managed land
topLandChangel(rs,aez,ts)s(topLandChangel,'rs,aez,ts')
=p=

--- ncpVuSin is a fudged min function and ensures that we don't expand beyond the available buffer
ncpVuSin(
sum(r_r(rs,rsNat),

--- distribution key of changes in total managed land cover from model regions to AEZ and sub-regions
p_topToCrop(rs,aez)

--- statistically estimated relation between change in total managed land cover and in crop land
times the crop land cover (strictly positive),
v_lambdaFCrop.range is zero if the exogenous crop land projections are switched off
* p_topLandChangeShare(rsNat) * sum(aezl,x_{AezNest.l(rsNat,aezl,"cropLand","%t0%")}
* [v_cropLandChg(rsNat,ts)-1] $ (v_lambdaFCrop.range(rsNat,ts) ne 0)

--- price dependent changes (change the estimated relation if active)
+ p_x_{AezNestBase}(rs,aez,"top","%t0%")

* [
--- stronger expansion in AEZ which above average price increases, elasticity is 0.25
[ ((p_{AezNest}(rs,aez,"top",ts)
/p_{AezNest.l(rs,aez,"top","%t0%"))/p_{ft}(rs,lnd,ts))^{**0.25}-1
] $ [sum(aezl $ (aezFlag(rs,aezl) $ x_{AezNest.l(rs,aezl,"top","%t0%"))),1) Gt 1]

--- stronger expansion in regions with above average price increases if p_etaf is zero
+ [ (p_{ft}(rs,lnd,ts)/p_{ft}(rsNat,lnd,ts))^{**0.25}-1
] $ (not p_etaf(rsNat,lnd))

--- all AEZ expand according to change in land price (elasticity 50% p_etaf, the remaining 50% are used in e_cropLand)
if driven by FAO projection: correction term
+ [(p_{ft}(rs,lnd,ts)/pgdpmp(rsNat,ts))^{**((p_etaf(rs,lnd) $ p_etaf(rs,lnd) + p_etaf(rsNat,lnd) $ (not p_etaf(rs,lnd))) * 0.5) - 1}
]
)
,

--- correct strongly the expansion of total managed land if one of the unmanaged land categories
drops below 10% of its benchmark area
p_maxBuffer(rs,aez)
- min(p_maxBuffer(rs,aez)*0.9,
p_landUse(rs,aez,"builtUpLand","ha") * p_builtupLandChange(rs,ts) * 1.E-6)
,max(1.E-4,p_x_{AezNestBase}(rs,aez,"top","%t0%")*1.E-6))
s(topLandChangel,'rs,aez,ts')
;

```

If the second part of the minimum condition becomes binding, less managed land is converted into crop land as suggested by the regression results. The code corrects the resulting absolute error by multiplying the crop land area at the benchmark with the regression coefficient and subtracts the simulated changes in total managed land as defined above:

```

--- Define error between targeted changes in managed land (FAO crop land projection (= v_croplandChg) times the
share of the additional crop land stemming from natural vegetation p_topLandChangeShare)) and the simulated change
in e_topLand1 equation

e_TopLandError(rsNat,ts(tRun)) $ (v_cropPrice.1(rsNat,"%t0%") $ (v_lambdaFCrop.range(rsNat,ts) ne 0)) ..

v_TopLandError(rsNat,ts)s(v_topLandError,'rsNat,ts')
==
{
    --- targeted total change of land in economic use ("top"): a share of crop land change

    [v_croplandChg(rsNat,ts)-1] * sum(aez,xAezNest.1(rsNat,aez,"cropLand","%t0%")) * p_topLandChangeShare(rsNat)

    --- simulated change

    - sum((r_r(rs,rsNat),aez) $ (aezFlag(rs,aez) $ (xAezNest.1(rs,aez,"top","%t0%") gt eps)),
        topLandChange1(rs,aez,ts))
    )s(v_topLandError,'rsNat,ts')
;

```

The error is distributed to the sub-regions and AEZ, taken the maximal remaining buffer into account:

```

--- update the results from e_topLand1 by distribution the error v_topLandError to the regions/aez based on
allocation key p_topToCrop.

e_topLand(aezFlag(rs,aez),ts(tRun)) $ ((xAezNest.1(rs,aez,"top","%t0%") gt eps) $ (not sameas(ts,"%t0%"))) ..

--- absolute change in managed land

topLandChange(rs,aez,ts)s(topLandChange,'rs,aez,ts')
==
--- as above: prevent expansion beyond the available buffer

ncpVuSin(
    --- estimated expansion from above times the nation wide error times the distribution keys

    topLandChange1(rs,aez,ts)+sum(r_r(rs,rsNat), v_TopLandError(rsNat,ts)*p_topToCrop(rs,aez)),

    --- correct strongly the expansion of total managed land if one of the unmanaged land categories
    drops below 10% of its benchmark area

    p_maxBuffer(rs,aez)
    - min(p_maxBuffer(rs,aez)*0.90,p_landUse(rs,aez,"builtUpLand","ha") * p_builtupLandChange(rs,ts) * 1.E-6)
    --- the last term is the fudging factor of the
    ,max(1.E-4,p_xAezNestBAse(rs,aez,"top","%t0%")*1.E-6)

    )s(topLandChange,'rs,aez,ts')
;

```

And enters as an additional term the equations *e\_topLand1* above. Jointly, these equations ensure that both the cropland projections and the changes in managed land according to the regression results are recovered.

Setting the crop land expansion exogenous requires an additional endogenous variable, here chosen is a shift in the productivity of all primary factors with the exemption of land and water. The latter two are shifters are driven by targeting crop yields. The related equation is an extension of the equations targeted real GDP growth, on demand combined with the productivity growth differences by three broad sectors. The first two terms (*lambdaFshift* and *p\_lamdbafShift2*) allow to introduce additional productivity shifters, either endogenous or exogenous. The *axpNat* driven changes are not applied land and water using for crops with targeted yields.

```

*
* --- map uniform shift in non-land productivity for all crops to individual ones
*
e_prodShiftM(rs,f,aIn(a),ts(tRun)) $ (xfFlag(rs,f,a) $ (lambdaf.range(rs,f,a,tRun) ne 0)
$ sum(r_r(rs,rNat),rgdpmp.l(rNat,tRun-1)) $ (not disr(rs))) ..

lambdaf(rs,f,a,tRun)s(lambdaf,'rs,f,a,tRun')

=E= [ lambdaf.l(rs,f,a,tRun-1)

*
*
* --- additional exogenous or endogenous shifters (not part of standard G-RDEM)
*
+ lambdafShift(rs,f,a,tRun) * p_timeStep(tRun) $ (lambdafShift.range(rs,f,a,tRun) eq 0)
+ [sum(r_r(rs,rNat),p_lambdafShift2(rNat,f,a,tRun)) * p_timeStep(tRun)]

+ [

*
* --- sub-national shifter (driven by estimated gva share)
*
axpReg(rs,tRun) $ subr(rs)

*
* --- 3-broad sectors differentiated productivity differences
*
+ sum(r_r(rs,rNat), axpNat(rsNat,tRun)
$ $iftheni.tfpLab "%TFPLab%"=="on"
$ 1(f)
$ $endif.tfpLab

* {
p_prodShift("a",a)
+ p_prodShift("b",a) * [ (rgdpmp(rsNat,tRun)/rgdpmp.l(rsNat,tRun-1))^(1/p_timeStep(tRun))-1 ]
+ p_prodShift("c",a) * sqrt[ (rgdpmp(rsNat,tRun)/rgdpmp.l(rsNat,tRun-1))^(1/p_timeStep(tRun))-1 ] )
]

*
* --- both shifters work differently for natural resources and fossil extraction activities
* We assume that natural resources for coal, oil and gas extractions only participate with 10%
* at the prod shifters for primary factors.
*
* The factor is also not applied for land and water for crops where e_agrYield equation defines
* these shifters.
*

* (1 - 0.90 $ (nr(f) $ enerNatRes(a))) $ (not ((lnd(f) or wat(f)) $ p_agrYield(rs,a)))

+ [

*
* --- shifters to recover crop land expansion affect land for non-crop activities, and all factors besides land and water for crop
* Assume we need less crop land to target the FAO projection and set v_lambdaFCrop > 0.
* This implies a higher capital productivity for crops that should reduce crop land (substitution)
* and a higher land productivity for animals/forest which should let pasture and forest land increase
*
sum(r_r(rs,rNat),v_lambdaFCrop(rsNat,tRun) ) $ ( [cap(f) $ p_agrYield(rs,a)] or [(not p_agrYield(rs,a)) $ lnd(f)])

*
* --- crop and water productivity driven by yield targeting (e_agrYield equation)
*
+ v_lambdafShift(rs,lnd,a,tRun) $ ((lnd(f) or wat(f)) $ p_agrYield(rs,a))

] * p_timeStep(tRun)
]s(lambdaf,'rs,f,a,tRun')
;

```

The yield targeting is based on the endogenous productivity shifter  $v\_lambdafShift$  for land. It is also applied to irrigation water. In order to recover the share of pastureland that is converted to cropland according to the regression results, the endogenous shifter  $v\_lambdafCrop$  changes the capital productivity of crops and the land crop productivity of other land use activity. The reasoning is that increases the capital productivity should provoke a substitution of land by capital, reducing crop land use. Increasing the land productivity of other land use activities should let their land use increase, pressuring on crop land.

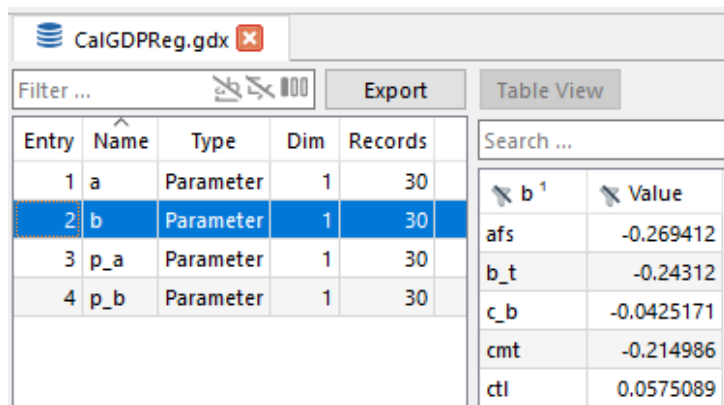
We tested the framework with differently detailed data base with regard to the country aggregation and found that it solved smoothly in most cases. We cannot exclude that especially for projections up to 2050, the overall framework including the G-RDEM mechanism might provoke infeasibilities. In such cases, one might try to increase the price elasticity effect in (3) as implausible high or low land prices can lead to frictions in other parts of model.

In this context it is worth noting that understanding the causes of infeasibilities in baseline construction can be challenging. A (too) low supply elasticity of natural resources, to give an example, can shift global income to a few countries with substantial production of related primary commodities and at the same time provoke quite unbalanced price developments across sectors. These reactions interact with the implicit correction via TFP changes to recover the given real GDP developments.

### 3.14.12 GDP-dependent correction of calorie content

Food consumption changes as prosperity increases. More expensive and highly processed products are consumed resulting in a lower calorie content per dollar spent for food. This is particularly important in long-term simulations in which significant economic development can take place. G-RDEM therefore offers the possibility to correct the calorie content per dollar spent as a function of GDP per capita development.

For this purpose, G-RDEM uses the results of an outlier-corrected exponential regression conducted in R. This determines the correlation between region-specific per capita GDP (real, at market prices) from the GTAP data and region- and sector-specific calorie content from the FAO FBS. The coefficients and corresponding p-values are made available to CGEBox via a GDX file named CalGDPRReg.gdx.



| Entry | Name | Type      | Dim | Records |
|-------|------|-----------|-----|---------|
| 1     | a    | Parameter | 1   | 30      |
| 2     | b    | Parameter | 1   | 30      |
| 3     | p_a  | Parameter | 1   | 30      |
| 4     | p_b  | Parameter | 1   | 30      |

| b <sup>1</sup> | Value      |
|----------------|------------|
| afs            | -0.269412  |
| b_t            | -0.24312   |
| c_b            | -0.0425171 |
| cmt            | -0.214986  |
| cti            | 0.0575089  |

The GAMS code CalGDPCorr.gms located in the gams/GTAPRdem directory is used if the corresponding check mark is set in the G-RDEM tab of the GUI.

GDP depend correction of calorie content



For calorie content correction, the slope coefficients and their p-values are first read in from the GDX file and stored on the parameters p\_coefBOri and p\_pValueBOri.

```
*
* --- load regression coefficient + p-values
*
* execute_load "GtapRDEM/CalGDPRReg.gdx" p_coefBOri=b, p_pValueBOri=p_b ;
```

Using the total nutrient demand for weighting (p\_wgtI, p\_wgtIOri), the regression parameters are adjusted to the sector aggregation in the current model run and stored on p\_coefB and p\_pValueB.



```

*
* --- define basis for nutrient demand (all final demand)
*
p_wgtI(rNat,i) = (xa.l(rnat,i,"hhsld","%t0%")*pa.l(rnat,i,"hhsld","%t0%")
+sum(gov,xa.l(rnat,i,gov,"%t0%")*pa.l(rnat,i,gov,"%t0%"))
+sum(inv,xa.l(rnat,i,inv,"%t0%")*pa.l(rnat,i,inv,"%t0%")));
*
* --- aggregate from single product to product list in GTAP data base
*
p_wgtIOri(rNat,iDat) = sum(mapi(iDat,i),p_wgtI(rNat,i));
*
* --- define elasticity and prob as weighted average
*
p_coefB(rNat,i) $ sum(mapi(iDat,i), p_wgtI(rNat,i))
= sum(mapi(iDat,i), p_coefBOri(iDat) * p_wgtIOri(rNat,iDat))
/sum(mapi(iDat,i), p_wgtIOri(rNat,iDat));
*
* --- define elasticity and prob as weighted average
*
p_pValueB(rNat,i) $ sum(mapi(iDat,i), p_wgtI(rNat,i))
= sum(mapi(iDat,i), p_pValueBOri(iDat) * p_wgtIOri(rNat,iDat))
/sum(mapi(iDat,i), p_wgtIOri(rNat,iDat));

```

Subsequently, the original calorie content  $p\_calT0$  is updated in each time step  $t$  using the slope coefficient  $p\_coefB$  and the regional GDP development if the p-value is significant ( $< 0.05$ ). The new calorie content is stored on the parameter  $p\_nutcontFabio$ .

```

*
* -- shift nutrient content with estimated elasticity coeff_b if p-value < 5%
*
p_nutcontFabio(rNat,i,"cal") $ [p_coefB(rNat,i) and (p_pValueB(rNat,i) lt 0.05)]
= p_calT0(rNat,i) * [ (rgdpmp(rNat,tSim)/pop(rNat,"hhsld",tSim))
/ (rgdpmp(rNat,"%t0%")/pop(rNat,"hhsld","%t0%")) ] ** p_coefB(rNat,i) ;

```

### 3.14.13 Steering of G-RDEM model runs

The screenshot presented in Figure 14 below shows how the user can select which data base aggregation to use, decide the time horizon of the baseline as well as the report frequency. For instance, retaining the flexible aggregation approach from GTAP in G-RDEM allows the user to develop a long-term baseline with a focus on only one country, while aggregating the rest of the world into a single aggregate. So far, the G-RDEM module has been tested with resolution with up to 80 countries.

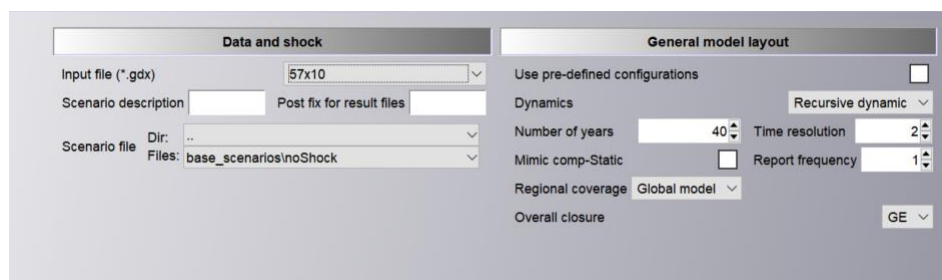


Figure 14 : Main steering panel of CGEBox

The set of years to run will always include the benchmark ( $t0$ ), the last year (eq card( $t$ )) and the first year of the simulation period, which can be used e.g. for estimations from the initialized variables.



```

$iftheni "%dynMode%"=="Recursive dynamic"
    trun(t0) = YES;
    trun(t) $ (t.pos eq card(t)) = YES;
    trun(t) $ (t.pos eq 2) = YES;
    trun(t) $ (mod(t.pos-1,%timeStep%) eq 0) = YES;

```

For the above input, the resulting years to run are:

```

---- 1406 SET trun
t00,    t01,    t05,    t10,    t15,    t20,    t25,    t30,    t35,    t40

```

In order to construct the baseline, the user selects (Figure 15) the desired SSP scenario for population and GDP, as well some procedural modules to use.

Figure 15 : Main steering panel of the G-RDEM module in CGEBox

#### 3.14.14 Population and GDP growth from the IIASA SSPs

IIASA host a web application<sup>68</sup> which gives access to model results relating to the IPCC model activities comprising GDP growth rate projection by country from a range of models and for different Shared Socio-Economic Pathways (SSPs). Some if GDP projections were downloaded in XLSX format and converted into a GDX container. That container is read by the driver program:

```

*
* --- load GDP projections from 2005 to 2100, in 10 year steps
*
execute_load "%datDir%\worldPopProj.gdx" wpp15_Popt;

set modP /"OECD Env-Growth","IIASA GDP"/;
set scen / SSP1,SSP2 /;
set var /"GDP|PPP" /;
set unit /"billion US$2005/yr" /;
set gYears / 2010,2020,2030,2040,2050,2060,2070,2080,2090,2100/;
parameter p_gdp(modP,scen,popreg,var,unit,gYears);
execute_load "%datDir%\GDP_SSP-IIASA.GDX" p_gdp;
option p_gdp:2:5:1;
parameter p_gdpC(popreg,*);
p_gdpC(popreg,gYears) = p_gdp("OECD Env-Growth","SSP2",popreg,"GDP|PPP","billion US$2005/yr",gYears);

```

The labels for the country follow the same codes as for the UN population projections. As seen from above, the projections are not yearly, but by decade. In order to use them flexible with different time steps, the code first interpolates linearly between given time points:

<sup>68</sup> [http://www.iiasa.ac.at/web/home/research/researchPrograms/Energy/SSP\\_Scenario\\_Database.html](http://www.iiasa.ac.at/web/home/research/researchPrograms/Energy/SSP_Scenario_Database.html)

```

*
* --- interpolate linearly between simulation points
*
set pYears_gYears(allYears,gYears,gYears) / (2010*2019).2010.2020
                                              (2020*2029).2020.2030
                                              (2030*2039).2030.2040
                                              (2040*2049).2040.2050
                                              (2050*2059).2050.2060
                                              (2060*2069).2060.2070
                                              (2070*2079).2070.2080
                                              (2080*2089).2080.2090
                                              (2090*2100).2090.2100 /;

alias(gYears,gYears1);

p_gdpC(popReg,allYears) = sum(pYears_gYears(allYears,gYears,gYears1),
                             p_gdpC(popreg,gYears)
                             + (p_gdpC(popreg,gYears1)-p_gdpC(popreg,gYears))/10 * mod(allYears.pos-1,10));

```

The mapping to the regional aggregate used in CGEBox is structural identical to assigning the population projections:

```

*
* --- growth rate between simulation points
*
p_popGrowth(r,tRun) $ (sum(r_popr(r,popreg),1) $ (tRun.pos > 1))
= {[ sum( r_popr(r,popreg),t_yp(tRun,pYears)), wpp15_popt("MEDP",popreg,pYears))
    /sum( r_popr(r,popreg),t_yp(tRun-1,pYears)), wpp15_popt("MEDP",popreg,pYears)) ] -1} * 100;
..

```

For some countries, notably small islands, no projections are available. In these cases, global unweighted averages are used instead.

```

*
* --- use average of counties with data for countries with missing ones
*
p_gdpGrowth(r,tRun) $ (sum(sameas(r,emptyGdp),1) $ (tRun.pos > 1))
= sum(rr, p_gdpGrowth(rr,tRun))/sum(rr $ p_gdpGrowth(rr,tRun),1);
display p_gdpGrowth;

```

For use in the model, per capita growth is used:

```

*
* --- GPD growth per capita is the economy wide growth rate
*      minus population growth rate
*
p_gdpPerCapGrowth(r,tRun) $ (tRun.pos > 1)
= p_gdpGrowth(r,tRun) - p_popGrowth(r,tRun);
..

```

The interface allows to overwrite these projections based on external input. In order to do so, the “Use additional macro projection” checkbox needs to be switched on. The user can then select a scenario file from the folder “*gams/scen/macro*”:

A dummy implementation of overwriting the data from the IIASA portal is shown below (see *gams/scen/macro/macroproj\_dummy.gms*):

```
*
* --- example format to map from data in 1980 ... format to p_driver
*
set myProjItems / GDP /;
set myProjYears(allYears) / 2014*2050 /;
parameter p_myProj(rNat,myProjItems,myProjYears);
option kill=p_myProj;

*
* --- example format to set GDP|PPP
*
p_drvc(rNat,"GDP|PPP",tRun) = p_myProj(rNat,"GDP",myProjYears);
```

### 3.14.15 Using the G-RDEM generated baseline for counterfactual analysis

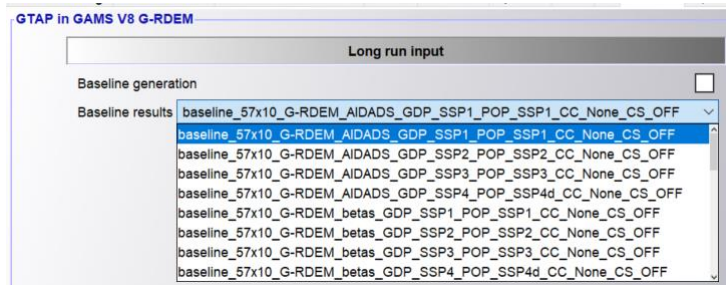
Besides analysing the baseline itself or comparing different baselines against each other, they also might serve as a starting point for counterfactual analysis. Two types of simulations are currently supported in the CGEBox platform. The first one uses the shifters generated during the baseline construction (changes in saving rates, productivity shifters, updated I-O coefficients), as well as population projections, as given parameters into a recursive-dynamic simulation. That makes only sense if the structural set-up of the model is unchanged (production function nesting, bi-lateral trade, etc.). During such a run, GDP would no longer be fixed, but endogenous and, as usual, shocks can be introduced, such as changes in policies or productivity, like climate change impacts on agricultural yields. Parameter estimates for some impacts (like those of climate) are already available in CGEBox and can be readily inserted during the simulation runs, by means of the graphical interface. For an application, see for instance Wilts et al. 2021<sup>69</sup>.

The second type of possible utilization of G-RDEM is in the provision of a benchmark data set data for comparative-static analysis. In this case, the global SAM and parameters for one of the simulated time points are loaded as a replacement for the benchmark data from GTAP. The full modular

<sup>69</sup> Wilts, R., Latka, C., Britz, W. (2021): Who is Most Vulnerable to Climate Change Induced Yield Changes? A Dynamic Long Run Household Analysis in Lower Income Countries, *Climate Risk Management* 33(100330)

flexibility of CGEBox could then be exploited, allowing to modify several characteristics of the model. That approach might be also interesting to produce not a long-run, but a rather a medium-term baseline as an ex-ante benchmark, which is for instance the usual practice in partial equilibrium modelling of agri-food markets.

The first option is to use the G-RDEM module in non-baseline generation mode:



In the case, the user can select the baseline results to employ. That will load the necessary shifters and exogenous parameters into the GAMS code:

```
variable  axpPrev(r,a,t),betasPrev(r,t),betagPrev(r,t),betapPrev(r,t);
parameter ioPrev(r,i,a,t),ioNestPrev(r,i,a,t),avaPrev(r,a,t),andPrev(r,a,t);
$GDXIN "%resdir%/run/%baselineRes%"
$$LOAD axpPrev      = axp
$$LOAD betasPrev    = betas
$$LOAD betagPrev    = betag
$$LOAD betapPrev    = betap
$$LOAD ioPrev       = io
$$LOAD ioNestPrev   = ioNest
$$LOAD andPrev      = and
$$LOAD avaPrev      = ava
$$LOAD p_timeStep
$$LOAD p_growthRate
$GDXIN
```

We have checked that without additional shocks, the existing baseline will be recovered without that GDP is fixed.

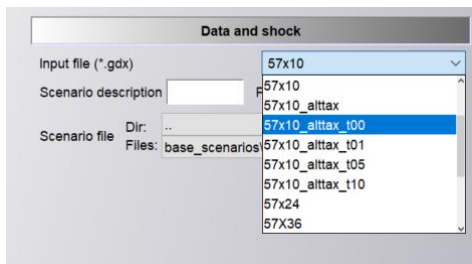
The second option requires that during baseline generation, altertax output is activated:



That will output the SAM and other information required for a benchmark calibration of the model are stored in “build” directory:

```
IDE 57x10_alttax_t01.gdx
IDE 57x10_alttax_t05.gdx
IDE 57x10_alttax_t10.gdx
IDE 57x10_alttax_t00.gdx
```

These benchmark data can be used similarly to the output from the data preparation step. The data can therefore be chosen as a starting point for comparative (or even recursive-dynamic) analysis:



That makes the layout of the model fully adjustable (e.g., by adding myGTAP for several household types, a different production nesting, changes in how bi-lateral trade is modelled).

### 3.14.16 Specific nestings in G-RDEM

We add some specific nestings for the production function (see `gams\gtaprdem\gtaprdem.gms`):

```
set natResA(a); natResA(a) $ sum(r, SAM0(r,"natRes",a)) = YES;

sigmap(r,natResA) = 0.20;
tNest("CAP+NATRES") = YES;
tNest_n_a("UA","CAP+NATRES",natResA) = YES;
tNest_f_a("CAP+NATRES","capital",natResA) = YES;
tNest_f_a("CAP+NATRES","natRes",natResA) = YES;
sigmaNest(r,"CAP+NATRES",natResA) = 0.20;

tNest_n_a(tNest,"CAP+NATRES",a) $ ((not sameas(tNest,"CAP+NATRES")) $ tNest_f_a(tNest,"capital",a)) = YES;
tNest_f_a(tNest,"capital",a) $ (not sameas(tNest,"CAP+NATRES")) = no;
tNest_n_a("UA","CAP+NATRES",a) $ sum(tNest_n_a(tNest,"CAP+NATRES",a),1) = NO;

$$include 'extensions\labor_nest.gms'

set landA(a); landA(a) $ sum(r, SAM0(r,"land",a)) = YES;
tNest("LabLand") = YES;
tNest_n_a("UA","LabLand",landA) = YES;
tNest_f_a("LabLand","Land",landA) = YES;
tNest_n_a("LabLand","Labor",landA) = YES;
tNest_n_a("UA","Labor",landA) = NO;
sigmaNest(r,"LabLand",landA) = 0.5;
```

GTAP-AGR has proven to work badly with some long-term projections, because of limited factor mobility between agricultural and non-agricultural sectors. Therefore, only the technology nests from GTAP-AGR are taken over:

```
$if not defined roecd $include 'gtapAgr\GtapAgr_def.gms'

*
* --- define feed demand sub-nest in livestock production
*
if ( card(feed),
    tNest("feed") = yes;
    tNest_i_a("feed",feed,lstk) = yes;
    tNest_n_a("ND","feed",lstk) = YES;
    sigmaNest(r,"Feed",lstk) = 2;
);

if ( card(agr_c),
    tNest("foodPrc") = yes;
    tNest_i_a("FoodPrc",agr_c,foodPrc) = yes;
    tNest_n_a("ND","FoodPrc",foodPrc) = YES;
    sigmaNest(r,"FoodPrc",foodPrc) = 0.5;
);
```

In order to reflect stronger substitutional relations between animal and crop based products, two final demand nests are constructed:

```

set cropD(i); cropD(i) $ sum(mapi(iCrops,i),1) = YES;
cropD(i) $ mapi("vol",i) = YES;
cropD(i) $ mapi("pcr",i) = YES;
cropD(i) $ mapi("sgr",i) = YES;
cropD(i) $ mapi("ofd",i) = YES;
cropD(i) $ mapi("b_t",i) = YES;

dNest("cropD") = YES;
dNest_i_fd("cropD",cropD,fdn) = YES;
dNest_n_fd("top","cropD",fdn) = YES;
sigmaFDNest(r,"cropD",fdn) = .50;

set animD(i); animD(i) $ sum(mapi(1stk0,i),1) = YES;

animD(i) $ mapi("cmt",i) = YES;
animD(i) $ mapi("omt",i) = YES;
animD(i) $ mapi("mil",i) = YES;

dNest("animD") = YES;
dNest_i_fd("animD",animD,fdn) = YES;
dNest_n_fd("top","animD",fdn) = YES;
sigmaFDNest(r,"animD",fdn) = .50;

```

### 3.14.17 Endogenous non-Hicks neutral technical progress and primary factors supply

In addition to price driven changes in endowments, the marginal productivity can also be rendered price dependent. In the default setting, the labour force is reacting to population changes fully, while land, irrigation water and natural resources follow to a limited extend (positive changes) in population. The other element are by default switched off.

| Elasticities to population growth |      | Hicks non-neutral TP to factor prices |      | Factor supply to factor prices |      |
|-----------------------------------|------|---------------------------------------|------|--------------------------------|------|
| Land                              | 0.05 | Land TP                               | 0.00 | Land XF                        | 0.00 |
| Irrigation water                  | 0.05 | Irrigation water TP                   | 0.00 | Irrigation water XF            | 0.00 |
| Natural resources                 | 0.20 | Natural resources TP                  | 0.00 | Natural resources XF           | 0.00 |
| Labour force                      | 1.00 | Labour force TP                       | 0.00 | Labour force XF                | 0.00 |
|                                   |      | Capital TP                            | 0.00 |                                |      |

The input from the GUI is mapped to the following symbols in the GAMS code:

```

*
* --- link changes in factor endowments to changes in population,
*
p_popShift(r,"land") = %landToPopElas%;

$$ifthene.GTAPWATER %nFactors%>8
  p_popShift(r,"water") = %WaterToPopElas%;
$$endif.GTAPWATER
p_popShift(r,"natres") = %natResToPopElas%;
*
* --- labour shares follow population growth
*
p_popShift(r,l) = %labToPopElas%;

```

```

* --- factor expansion depending on factor price developments:
*   labor participation rates react to wage rate changes

set dummyPS / xf,tp /;
option kill=p_priceShift;

*
* --- endogenous non-Hicks neutral progress to depending on development of factor prices
*
p_priceShift(r,"tp","land") = %landTpToPfeLas%;
$$ifthenelse.GTAPWATER %nFactors%>8
  p_priceShift(r,"tp","water") = %waterTpToPfeLas%;
$$endif.GTAPWATER
p_priceShift(r,"tp","natres") = %natResTpToPfeLas%;
p_priceShift(r,"tp",l) = %labTpToPfeLas%;
p_priceShift(r,"tp",cap) = %capTpToPfeLas%;
*
* --- factor supply to factor price
*
p_priceShift(r,"xf","land") = %landXfToPfeLas%;
$$ifthenelse.GTAPWATER %nFactors%>8
  p_priceShift(r,"xf","water") = %waterXfToPfeLas%;
$$endif.GTAPWATER
p_priceShift(r,"xf","natres") = %natResXfToPfeLas%;
p_priceShift(r,"xf",l) = %labXfToPfeLas%;

```

We also take changes in education levels into account:

```

*
$$ifthenelse.GTAPVersion %GTAPNO%<9
  p_popShift(rnat,l) $ ff("skLab",l)
    = (1+p_growthRate(rnat,"tertiary",tSim)/100) / (1+p_growthRate(rnat,"population",tSim)/100);
  p_popShift(rnat,l) $ ff("unskLab",l)
    = (1+p_growthRate(rnat,"primary",tSim)/100) / (1+p_growthRate(rnat,"population",tSim)/100);
$$else.GTAPVersion
  p_popShift(rnat,l) $ ff("off_mgr_pros",l)
    = (1+p_growthRate(rnat,"tertiary",tSim)/100) / (1+p_growthRate(rnat,"population",tSim)/100);
  p_popShift(rnat,l) $ ff("ag_othlowsk",l)
    = (1+p_growthRate(rnat,"primary",tSim)/100) / (1+p_growthRate(rnat,"population",tSim)/100);
$$endif.GTAPVersion

```

The population and factor price shifters are implemented as follows:

```

*
* ---- shift factor supply depending on population and factor prices
*
aft(r,fnm,tSim) $ aft(r,fnm,tSim)
  = aft(r,fnm,tSim-1)
    * (1 + sum(r_r(r,rNat) $ (p_popt(rNat,tSim) gt p_popt(rNat,tSim-1)),
      p_popShift(rNat,fnm) * (p_popt(rNat,tSim)/p_popt(rNat,tSim-1)-1))
    + sum(r_r(r,rNat),p_priceShift(rNat,"xf",fnm)) * min(10.0,max(-0.90+0.9 $ (not l(fnm)), (pft.1(r,fnm,tSim-1)-1))) );
gf(r,fnm,a,tSim) $ xf.1(r,fnm,a,tSim-1)
  = gf(r,fnm,a,tSim-1)
    * (1 + sum(r_r(r,rNat) $ (p_popt(rNat,tSim) gt p_popt(rNat,tSim-1)),
      p_popShift(rNat,fnm) * (p_popt(rNat,tSim)/p_popt(rNat,tSim-1)-1))
    + sum(r_r(r,rNat),p_priceShift(rNat,"xf",fnm)) * min(10,max(-0.90+0.9 $ (not l(fnm)), (pf.1(r,fnm,a,tSim-1)-1))) );

```

The resulting technical progress changes are calculated as follows:

```

: *
: * --- productivity of factors is updated depending on factor price changes and GDP growth
: *
: lambdaf.fx(r,f,a,tSim) $ xf.1(r,f,a,tSim-1)
:   = m_lambdaf(r,f,a,tSim-1) * (1.0 + sum(r_r(r,rNat),p_priceShift(rNat,"tp",f)
:     * min(10,max(-0.90, (
:       pf.1(r,f,a,tSim) $ (fnm(f) or (omegaf(r,f) ne inf))
:       + pft.1(r,f,tSim) $ fm(f) $ (omegaf(r,f) eq inf)) -1)))
:   + p_growthRate(rNat,"GDPPerCAP",tSim)/100*0.5*0
:   + p_GrowthRate(rNat,"Population",tSim)/100*(1-p_popShift(rNat,f))*0.5*0);

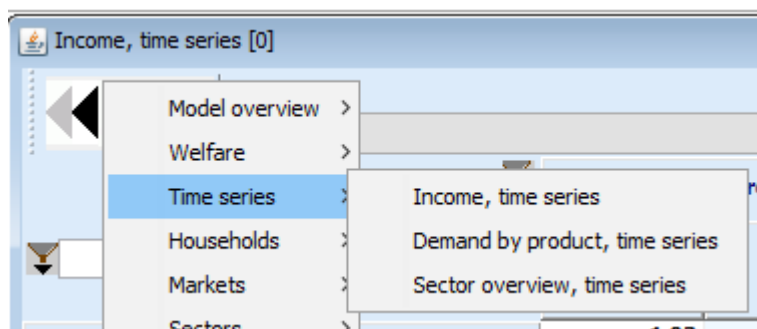
```



### 3.14.18 Analysis of G-REDM simulation output (exploitation)

Loading all results from long time recursive dynamic with detailed global SAMs might require amounts of memory exceeding most personal computer configurations. Therefore, the exploitation selection allows to pick up only a few years for inspecting the results. Therefore, while all output variables are stored in the GDX repository output file, only the selected ones will be loaded into the memory.

Generally, any view existing in the exploitation tool can be changed to a time series using the pivot facility. In order to ease the exploitation, three views have been pre-configured, under the heading “Time series”:



The first one reports changes in GDP, price indices etc:

Income, time series [0]

Regions

Total or household groups

World

Total

10x10\_example\_recDyn

|                                                     | t1       | t2                | t3                | t4                | t5                 | t6                 | t7                 | t8                 | t9 |
|-----------------------------------------------------|----------|-------------------|-------------------|-------------------|--------------------|--------------------|--------------------|--------------------|----|
| Total utility [Index]                               | 1,02     | 1,04<br>2,03%     | 1,06<br>4,13%     | 1,09<br>6,31%     | 1,11<br>8,57%      | 1,13<br>10,92%     | 1,16<br>13,35%     | 1,18<br>15,86%     |    |
| Utility of private household [Index]                | 1,02     | 1,04<br>2,25%     | 1,06<br>4,64%     | 1,09<br>7,09%     | 1,11<br>9,64%      | 1,14<br>12,26%     | 1,17<br>14,98%     | 1,20<br>17,78%     |    |
| Utility of government [Index]                       | 1,03     | 1,06<br>2,50%     | 1,08<br>4,81%     | 1,11<br>7,24%     | 1,13<br>9,77%      | 1,16<br>12,40%     | 1,19<br>15,11%     | 1,22<br>17,91%     |    |
| Utility of saving [Index]                           | 1,04     | 1,05<br>1,08%     | 1,06<br>2,21%     | 1,07<br>3,46%     | 1,09<br>4,80%      | 1,10<br>6,21%      | 1,12<br>7,72%      | 1,13<br>9,31%      |    |
| GDP [Bio \$]                                        | 50969,86 | 52256,51<br>2,52% | 53722,55<br>5,40% | 55362,24<br>8,62% | 57135,19<br>12,10% | 59017,39<br>15,79% | 60990,06<br>19,66% | 63039,17<br>23,68% |    |
| GDP, normalized with price index [Bio \$]           | 50444,38 | 51762,22<br>2,61% | 52973,64<br>5,01% | 54201,20<br>7,45% | 55445,69<br>9,91%  | 56709,14<br>12,42% | 57994,27<br>14,97% | 59302,57<br>17,56% |    |
| GDP, normalized with price index, per cap. [Bio \$] | 7538,05  | 7653,68<br>1,53%  | 7752,19<br>2,84%  | 7852,07<br>4,17%  | 7953,58<br>5,51%   | 8057,09<br>6,89%   | 8163,00<br>8,29%   | 8271,47<br>9,73%   |    |
| Price index [Index]                                 | 1,01     | 1,01<br>-0,09%    | 1,01<br>0,37%     | 1,02<br>1,09%     | 1,03<br>1,98%      | 1,04<br>3,00%      | 1,05<br>4,08%      | 1,06<br>5,20%      |    |
| Consumer price [Index]                              | 1,02     | 1,02<br>-0,00%    | 1,02<br>0,49%     | 1,03<br>1,24%     | 1,04<br>2,17%      | 1,05<br>3,20%      | 1,06<br>4,31%      | 1,07<br>5,44%      |    |
| Government price [Index]                            | 1,00     | 1,00<br>0,45%     | 1,01<br>0,98%     | 1,02<br>1,77%     | 1,03<br>2,69%      | 1,04<br>3,71%      | 1,05<br>4,78%      | 1,06<br>5,89%      |    |
| Savings price [Index]                               | 0,99     | 1,00<br>0,54%     | 1,01<br>1,54%     | 1,02<br>2,82%     | 1,04<br>4,28%      | 1,05<br>5,87%      | 1,07<br>7,54%      | 1,09<br>9,24%      |    |
| Factor income [Bio \$]                              | 40419,89 | 41420,74<br>2,48% | 42554,94<br>5,28% | 43820,08<br>8,41% | 45185,25<br>11,79% | 46632,30<br>15,37% | 48146,78<br>19,12% | 49717,98<br>23,00% |    |
| Factor income, normalized with price index [Bio \$] | 40003,18 | 41028,95<br>2,56% | 41961,71<br>4,90% | 42901,09<br>7,24% | 43849,11<br>9,61%  | 44808,44<br>12,01% | 45781,84<br>14,45% | 46770,98<br>16,92% |    |
| Current account balance [Bio \$]                    | -0,00    |                   | -0,00<br>0,00%    | -0,00<br>0,00%    | -0,00<br>-100,00%  | 0,00<br>300,00%    | -0,00<br>-300,00%  |                    |    |
| Tax income [Bio \$]                                 | 8862,84  | 9096,80<br>2,64%  | 9366,76<br>5,69%  | 9670,67<br>9,11%  | 10000,95<br>12,84% | 10352,84<br>16,81% | 10722,88<br>20,99% | 11108,39<br>25,34% |    |
| Population size [Mil heads]                         | 6691,97  | 6763,05<br>1,06%  | 6833,37<br>2,11%  | 6902,79<br>3,15%  | 6971,16<br>4,17%   | 7038,42<br>5,18%   | 7104,53<br>6,16%   | 7169,54<br>7,14%   |    |

The second one focuses on the demand side:



Demand by product, time series [0]

Regions: World

Items: Quantity

Sectors and institutions: Total

Origins: Total

Percentage diff. to Years t1

|                             | t1        | t2        | t3        | t4        | t5        | t6        | t7        | t8        | t9        | t10       | t11       | t12       | t13       | t14       | t15       | t16       |
|-----------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Total                       | 142485,78 | 146008,50 | 149704,95 | 153565,70 | 157563,80 | 161686,59 | 165926,36 | 170277,59 | 174730,84 | 179282,19 | 183929,30 | 188668,58 | 193499,03 | 198420,80 | 203434,33 | 208447,86 |
| Total                       | 142485,78 | 146008,50 | 149704,95 | 153565,70 | 157563,80 | 161686,59 | 165926,36 | 170277,59 | 174730,84 | 179282,19 | 183929,30 | 188668,58 | 193499,03 | 198420,80 | 203434,33 | 208447,86 |
| Grains and Crops            | 2916,05   | 3018,27   | 3111,50   | 3206,81   | 3303,73   | 3401,67   | 3501,18   | 3602,60   | 3706,14   | 3811,86   | 3919,82   | 4030,00   | 4142,37   | 4256,90   | 4373,56   | 4490,22   |
| Livestock and Meat Products | 2238,93   | 2310,26   | 2379,05   | 2447,52   | 2515,89   | 2584,10   | 2652,63   | 2721,80   | 2791,83   | 2862,86   | 2934,95   | 3008,14   | 3082,46   | 3157,93   | 3234,56   | 3311,22   |
| Mining and Extraction       | 6488,89   | 6680,69   | 6887,33   | 7108,58   | 7345,69   | 7599,84   | 7871,96   | 8162,67   | 8467,89   | 8786,67   | 9118,78   | 9463,23   | 9819,83   | 10188,75  | 10570,14  | 10954,00  |
| Processed Food              | 4741,70   | 4858,49   | 4977,09   | 5095,79   | 5214,73   | 5334,18   | 5454,09   | 5576,63   | 5700,29   | 5825,87   | 5953,49   | 6083,19   | 6215,02   | 6349,05   | 6485,32   | 6622,74   |
| Textiles and Clothing       | 2908,33   | 2982,27   | 3056,45   | 3132,91   | 3210,85   | 3290,30   | 3371,33   | 3454,00   | 3538,37   | 3624,46   | 3712,33   | 3802,02   | 3893,54   | 3986,90   | 4082,13   | 4179,22   |
| Light Manufacturing         | 15828,16  | 16196,59  | 16584,91  | 16988,13  | 17402,62  | 17827,67  | 18262,65  | 18707,15  | 19160,68  | 19623,06  | 20094,20  | 20573,87  | 21062,04  | 21558,85  | 22064,41  | 22578,81  |
| Heavy Manufacturing         | 34507,74  | 35399,53  | 36334,95  | 37309,64  | 38317,01  | 39355,67  | 40424,27  | 41521,92  | 42646,20  | 43796,34  | 44972,09  | 46172,49  | 47397,37  | 48646,94  | 49921,36  | 51211,61  |
| Utilities and Construction  | 11842,57  | 12158,73  | 12489,41  | 12830,10  | 13179,21  | 13536,65  | 13902,15  | 14275,62  | 14656,89  | 15045,89  | 15442,59  | 15846,90  | 16258,82  | 16678,44  | 17105,85  | 17540,05  |
| Transport and Communication | 19758,69  | 20250,46  | 20774,90  | 21322,58  | 21889,05  | 22472,11  | 23070,19  | 23682,21  | 24307,76  | 24946,49  | 25598,09  | 26262,30  | 26939,04  | 27628,33  | 28330,25  | 29042,74  |
| Other Services              | 41254,72  | 42153,21  | 43109,35  | 44123,64  | 45183,03  | 46284,40  | 47415,33  | 48572,99  | 49754,80  | 50958,69  | 52182,95  | 53426,46  | 54688,54  | 55968,72  | 57266,74  | 58581,01  |

The third one reports production quantities and related input demands:

Sector overview, time series [0]

Regions: World

Items: Quantity

Sectors: Total

Origins: Total

Percentage diff. to Years t1

|                                | t1        | t2        | t3        | t4        | t5        | t6        | t7        | t8        | t9        | t10       | t11       | t12       | t13       | t14       | t15       | t16       |
|--------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Total output                   | 111673,89 | 114430,69 | 117323,98 | 120347,07 | 123478,20 | 126711,33 | 130044,46 | 133477,59 | 136910,72 | 140443,85 | 143976,98 | 147510,11 | 151043,24 | 154576,37 | 158109,50 | 161642,63 |
| Output taxes                   | 2089,11   | 2146,66   | 2214,24   | 2290,95   | 2374,69   | 2458,43   | 2542,17   | 2625,91   | 2709,65   | 2793,39   | 2877,13   | 2960,87   | 3044,61   | 3128,35   | 3212,09   | 3295,83   |
| Total intermediate demand      | 56359,58  | 57845,86  | 59441,21  | 61141,64  | 62927,28  | 64787,73  | 66715,77  | 68706,07  | 70752,24  | 72850,90  | 74999,56  | 77195,45  | 79437,23  | 81723,01  | 84054,79  | 86386,57  |
| Total intermediate input taxes | 1688,87   | 1739,42   | 1800,32   | 1870,54   | 1948,00   | 2031,53   | 2120,18   | 2213,23   | 2310,23   | 2410,79   | 2514,61   | 2621,48   | 2731,24   | 2844,00   | 2957,76   | 3071,52   |
| Total factor demand            | 46523,88  | 47852,70  | 49355,00  | 51017,93  | 52805,29  | 54696,36  | 56675,29  | 58730,48  | 60853,46  | 63038,09  | 65279,70  | 67574,77  | 69920,84  | 72311,91  | 74747,98  | 77183,05  |
| Total intermediate demand      | 56359,58  | 57845,86  | 59441,21  | 61141,64  | 62927,28  | 64787,73  | 66715,77  | 68706,07  | 70752,24  | 72850,90  | 74999,56  | 77195,45  | 79437,23  | 81723,01  | 84054,79  | 86386,57  |
| Grains and Crops               | 1444,20   | 1507,18   | 1564,27   | 1623,01   | 1683,04   | 1744,07   | 1806,44   | 1870,38   | 1935,98   | 2003,25   | 2072,17   | 2142,69   | 2214,67   | 2288,11   | 2362,00   | 2436,44   |
| Livestock and Meat Products    | 1029,96   | 1066,93   | 1102,40   | 1138,45   | 1174,97   | 1211,83   | 1249,19   | 1287,16   | 1325,82   | 1365,18   | 1405,26   | 1446,03   | 1487,49   | 1529,74   | 1572,78   | 1616,51   |
| Mining and Extraction          | 3407,47   | 3512,93   | 3624,66   | 3744,48   | 3873,07   | 4011,19   | 4159,28   | 4317,52   | 4484,32   | 4659,34   | 4842,34   | 5032,71   | 5230,26   | 5437,99   | 5655,80   | 5881,79   |
| Processed Food                 | 1351,47   | 1389,92   | 1429,51   | 1470,40   | 1512,28   | 1555,08   | 1598,82   | 1643,53   | 1689,20   | 1735,86   | 1783,50   | 1832,09   | 1881,65   | 1931,19   | 1981,70   | 2033,18   |
| Textiles and Clothing          | 929,78    | 956,88    | 984,46    | 1013,42   | 1043,35   | 1074,17   | 1105,82   | 1138,26   | 1171,46   | 1205,41   | 1240,10   | 1275,54   | 1311,72   | 1348,54   | 1385,91   | 1423,83   |
| Light Manufacturing            | 6248,99   | 6403,78   | 6569,08   | 6744,22   | 6926,68   | 7115,51   | 7309,96   | 7509,48   | 7713,65   | 7922,22   | 8135,00   | 8351,81   | 8572,56   | 8797,26   | 9025,91   | 9258,56   |
| Heavy Manufacturing            | 15394,40  | 15805,70  | 16240,06  | 16697,97  | 17174,96  | 17669,64  | 18180,75  | 18707,43  | 19248,05  | 19802,00  | 20368,90  | 20948,13  | 21539,48  | 22142,81  | 22758,14  | 23384,47  |
| Utilities and Construction     | 3804,34   | 3900,58   | 4003,55   | 4113,45   | 4228,74   | 4348,70   | 4472,76   | 4600,46   | 4731,50   | 4865,63   | 5002,69   | 5142,53   | 5285,05   | 5433,16   | 5585,94   | 5739,41   |
| Transport and Communication    | 8912,65   | 9155,35   | 9417,58   | 9698,51   | 9994,11   | 10302,20  | 10621,16  | 10949,79  | 11287,37  | 11633,36  | 11987,35  | 12349,02  | 12718,19  | 13094,76  | 13478,74  | 13870,11  |
| Other Services                 | 13836,31  | 14146,60  | 14505,64  | 14897,72  | 15316,08  | 15755,32  | 16211,60  | 16682,06  | 17164,88  | 17658,65  | 18162,25  | 18674,90  | 19196,05  | 19725,70  | 20265,94  | 20816,78  |
| Total factor demand            | 46523,88  | 47852,70  | 49355,00  | 51017,93  | 52805,29  | 54696,36  | 56675,29  | 58730,48  | 60853,46  | 63038,09  | 65279,70  | 67574,77  | 69920,84  | 72311,91  | 74747,98  | 77183,05  |
| Land                           | 524,20    | 524,65    | 525,09    | 525,50    | 525,89    | 526,27    | 526,62    | 526,96    | 527,28    | 527,59    | 527,88    | 528,15    | 528,42    | 528,69    | 528,96    | 529,23    |
| UnSkLab                        | 15120,76  | 15406,20  | 15736,19  | 16159,44  | 16646,41  | 17186,67  | 17770,70  | 18391,60  | 19044,12  | 19724,43  | 20429,47  | 21156,87  | 21904,88  | 22673,31  | 23462,24  | 24272,67  |
| SkLab                          | 9495,30   | 9553,90   | 9727,62   | 9953,62   | 10225,02  | 10531,34  | 10865,87  | 11223,51  | 11600,52  | 11994,11  | 12402,14  | 12822,93  | 13255,22  | 13698,91  | 14174,00  | 14671,49  |
| Capital                        | 20739,04  | 21723,07  | 22720,96  | 23733,95  | 24762,29  | 25806,16  | 26865,93  | 27942,01  | 29034,91  | 30145,12  | 31273,17  | 32419,57  | 33584,88  | 34769,99  | 35975,00  | 37199,01  |
| NatRes                         | 644,58    | 644,87    | 645,15    | 645,42    | 645,68    | 645,93    | 646,17    | 646,40    | 646,63    | 646,84    | 647,05    | 647,25    | 647,44    | 647,64    | 647,84    | 648,04    |

### 3.14.19 Computing and data base considerations for G-RDEM runs

Compared to a comparative-static run with the model, constructing a baseline over several decades requires many solves. As indicated above, the time resolution can be chosen by the user. While larger time steps such as five year intervals decrease the number of necessary solves, the resulting shock in each period becomes larger, which implies more time spent by the solver to find a fixed point. For a dataset larger than around 57x24, it is therefore recommended to solve at least in bi-yearly steps, whereas less detailed data sets can still solve quite fast in five year steps. Using the built-in pre-solve mechanism with at least three steps is also recommended for more detailed datasets. Furthermore, we

recommend aggregating over the Armington agents, which removes the agent specific differentiation between domestic and imported shares and reduces the model complexity and rigidities in the demand composition.



Some relatively large data sets have been tested, including 33 from all 57 sectors and 80 countries from the 140 available in GTAP 9. One run over 40 years in 2 year steps with such detailed data base might require about two hours. Note that especially debt accumulation from foreign savings might lead to infeasibilities, such that it might be required to manually reduce the accounted share of related debt servings for specific regions. In the same vein, we typically used an ALTERNAX run to construct a benchmark where very high negative macro-economic saving rates are reduced (see gams\scen\user\_scenarios\corr\_betas.gms). These odd saving rates might point a specific problem in GTAP 9 data, as in the 2011 the global economy was still affected by an economic crisis.

More generally, it should be understood that baseline construction is only partially a process which can be delegated to a ready-to-use code. A close look at some key results, such as the evolution of foreign savings, saving rates and sectoral output per capita is recommended to rule out implausible findings. An ALTERNAX adjustment may be needed to vary the initial benchmark or it might be necessary to add specific shocks on top of the baseline mechanism. We also remind the reader that the SSP1 to 5 storylines comprise elements relating to governance such as policies to abate emissions or control land use change which are not integrated in the model structure discussed here.

### 3.15 GTAP-Melitz: Heterogenous firm module

This section is prepared by Yaghoob Jafari and Wolfgang Britz

#### 3.15.1 Summary

The GTAP-Melitz module replaces the Armington assumption to depict bi-lateral trade by a heterogenous firm module where each sector is assumed to consist of firms of differing productivity which each produce quality differentiated products. Firms face fix costs to enter the industry and to become active on specific trade links. The module does not need additional data, the user can select the parameters for the module on the GUI along which the sector to which the module is applied.

#### 3.15.2 Introduction

Since Armington (1969)<sup>70</sup> proposed to treat imported and domestic varieties of goods in the same classification as imperfect substitutes depicted by a CES-utility function, that approach dominated applied Computable General Equilibrium (CGE) analysis. It provides a powerful, but relatively simple framework for studying international trade policy, not at least as it can accommodate any observed pattern of trade flows and related prices, i.e. the intensive margin of trade. However, preferences for each origin in the Armington model are fixed such that changes in trade cannot impact average imported qualities on a trade link. It hence neglects potential variations at the extensive margin of trade such as trade flows in new products which are found as important in empirical analysis (Hummels and Klenow, 2005<sup>71</sup>; Chaneny, 2008<sup>72</sup>, among others).

The pioneer paper by Melitz (2003<sup>73</sup>) introduced firm productivity heterogeneity drawing from Hopenhayn (1992<sup>74</sup>) into the monopolistic competition framework by Krugman (1980)<sup>75</sup>. The Melitz model combines changes at the intensive and extensive margins of trade by allowing firms to self-select new export markets based on their productivity level. Subsequently, many papers applying the model (Bernard et al., 2003, 2006, 2007<sup>76</sup>; Eaton et al., 2004<sup>77</sup>) supported its empirical evidence by

---

<sup>70</sup> Armington, P.S. 1969. "Theory of demand for products distinguished by place of production." *International Monetary Fund Staff Papers*, 16(1): 159–178. doi:10.2307/3866403

<sup>71</sup> Hummels, D., and P.J. Klenow. 2005. "The variety and quality of a nation's exports." *American Economic Review*, 95(3): 704–723. doi:10.1257/0002828054201396

<sup>72</sup> Chaney, T. 2008. "Distorted gravity: the intensive and extensive margins of international trade." *American Economic Review*, 98(4): 1707–1721. doi:10.1257/aer.98.4.1707

<sup>73</sup> Melitz, M.J. 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica*, 71(6): 1695–1725. doi:10.3386/w8881

<sup>74</sup> Hopenhayn, H. 1992. "Entry, exit, and firm dynamics in long run equilibrium." *Econometrica*, 60(5): 1127–1150. doi:10.2307/2951541

<sup>75</sup> Krugman, P. (1980), "Scale Economies, Product Differentiation, and the Pattern of Trade", *American Economic Review* 70, 950-959

<sup>76</sup> Bernard, A. B., and J. B. Jensen. (1995). Exporters, jobs, and wages in U.S. manufacturing, 1975-1987. In *Brookings Papers on Economic Activity: Microeconomics 1995*, pp. 67-112.; Bernard, A.B., J. Eaton, J.B. Jensen, and S. Kortum. 2003. "Plants and productivity in international trade." *American Economic Review*, 93(4): 1268–1290. doi:10.3386/w7688. ; Bernard, A.B., J.B. Jensen, and P.K. Schott. 2006. "Trade costs, firms and productivity." *Journal of Monetary Economics*, 53(5): 917–937. doi:10.1016/j.jmoneco.2006.05.010.; Bernard, A.B., Redding, S., Schott, P.K., 2007. Comparative advantage and heterogeneous firms. *Review of Economic Studies* 74, 31–66

reproducing salient trade patterns observed in recent micro level studies. As the Melitz model provides a more general framework to depict bi-lateral trade which has proven as empirically superior, there have been a number of efforts to introduce firm heterogeneity into CGE models (Zhai 2008<sup>78</sup>; Balistreri et al., 2011<sup>79</sup>; Oyamada, 2013<sup>80</sup>; Akgul et al., 2016<sup>81</sup>; and Dixon et al., 2016<sup>82</sup>).

The first paper which introduced firm heterogeneity following Melitz (2003) into a CGE framework is Zhai (2008). The Zhai-Melitz approach captures variations in the extensive margin of trade flows in contrast to traditional CGE models based on Armington's (1969) assumption. Zhai (2008) provided the theoretically well-grounded empirical model based on the assumption of no free entry and exit. However, the Zhai implementation allows for adjustments in the extensive margin of trade only as a result of changed export share of firms engaging in a specific trade link, while limiting the variety gains brought by new entrants. Consequently, this assumption results in overestimation of the extensive margin of trade and in turn productivity effect. Balistreri et al. (2011) overcomes that restriction by allowing also for new entrants on each trade link and accounts for a certain exogenous share of firms leaving in each period the industry. Akgul, Vitoria and Hertel (2016) introduced the Melitz framework into the standard GTAP model, abstracting from exogenous firm exit.

Besides, Balistreri and Rutherford (2013)<sup>83</sup> set out stylized versions of the Armington, Krugman, and Melitz under a unified treatment, and then compare the outcome of different approaches. Inspired by Balistreri and Rutherford (2013), Dixon et al. (2016) draws the connections between the three models by developing them sequentially as special cases of an encompassing model. Dixon et al. (2016) interestingly show that the Armington, Krugman, and Melitz models are progressively less restrictive special cases of a more generalized model, derived from a cost minimization problem of a worldwide planner. While Dixon et al. (2016) provide an illustrative numerical general equilibrium model with Melitz sectors, using the earlier version of Dixon et al. (2016), Oyamada (2014) develops a simple AGE framework within which user can switch between Armington, Krugman, and Melitz models, and shows how the Dixon et al. (2016) framework can be parameterized<sup>84</sup>.

---

<sup>77</sup> Eaton, J., S. Kortum, and F. Kramarz. 2004. "Dissecting trade: Firms, industries, and export destinations." *American Economic Review*, 94(2): 150–154. doi:10.3386/w10344

<sup>78</sup> Zhai, F. 2008. "Armington meets Melitz: introducing firm heterogeneity in a global CGE model of trade." *Journal of Economic Integration*, 23(3): 575–604. doi:10.11130/jei.2008.23.3.575

<sup>79</sup> Balistreri, E.J., R.H. Hillberry, and T.F. Rutherford. 2011. "Structural estimation and solution of international trade models with heterogeneous firms." *Journal of International Economics*, 83(2): 95–108. doi:10.2139/ssrn.1153944

<sup>80</sup> Oyamada, K. 2014. "Behavioral characteristics of applied general equilibrium models with an Armington, Krugman, and Melitz encompassing module." 17th Annual Conference of Global Economic Analysis, Dakar, Senegal. [https://www.gtap.agecon.purdue.edu/resources/res\\_display.asp?RecordID=4459](https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=4459)

<sup>81</sup> Akgul, Z., N.B. Vitoria, and T.W. Hertel. 2016. "GTAP-HET: Introducing firm heterogeneity into the GTAP model." *Journal of Global Economic Analysis*, 1(1): 111–180

<sup>82</sup> Dixon, P.B., M. Jerie, and M.T. Rimmer. 2016. "Modern trade theory for CGE modelling: the Armington, Krugman and Melitz models." *Journal of Global Economic Analysis*, 1(1): 111–180

<sup>83</sup> Balistreri, E.J., and T.F. Rutherford. 2013. "Computing general equilibrium theories of monopolistic competition and heterogeneous firms." In *Handbook of Computable General Equilibrium Modeling*, edited by P. B. Dixon and D. W. Jorgenson

<sup>84</sup> See also Itakura, Ken and Kazuhiko Oyamada. 2013. "Incorporating Firm Heterogeneity into the GTAP Model." Paper presented at the 16th Annual conference on Global Economic Analysis, Shanghai, China for technical aspects and Roson and Oyamada (2014) Roson, R., and Kazuhiko Oyamada. 2014. "Introducing

We discuss in here the introduction of the Melitz model in CGEBOX. Besides realized in GAMS (GAMS Development Corporation, 2013<sup>85</sup>) and not in GEMPACK (Harrison and Pearson, 1996<sup>86</sup>), it differs from the standard GTAP model (Hertel, 1997) in several aspects. Inter alia, its equations are all written in levels, whereas the GEMACK realization uses mostly equations depicting relative changes in linearized form. Furthermore, it aims at a modular and easily extendable framework. The implementation of the Melitz model discussed in here provides an example of such a modular extension.

The remainder of this paper is structured as follows: the next section (2) presents a brief review of the trade theories literature; section (3) describe the implementation of firm heterogeneity in the CGEBox model; section (4) presents the technical implementation of the CGEBox model with heterogeneous firms along with results of an example application; and finally section (5) concludes.

### 3.15.3 Brief Literature overview

With regard to international trade, traditional applied general equilibrium models fail in two important aspects (see critique by Keohoe 1994<sup>87</sup>): they do neither allow (trade) policies to impact on aggregate productivity nor on trade along the extensive margin. In order to overcome these shortcomings, a new type of international trade models has emerged drawing on the Melitz (2003) model. It considers heterogeneous firms under monopolistic competition which can self-select into export market.

In traditional trade theories, countries specialize in production and export of those commodities in which they have comparative advantages, in the Ricardian framework based on differences in technology and in the Heckscher-Ohlin one in endowments. These models assume perfect competition and constant returns to scale, often implying that firm size is indeterminate. Due to the resulting specialization, countries either import or export a certain product. Krugman (1980), Helpman (1981)<sup>88</sup> and Ethier (1982)<sup>89</sup> established the so-called “new” trade theories drawing on variety-based models. Here, firms specialize in distinct horizontally-differentiated varieties of a product which provides an explanation why countries simultaneously export and import within the same broader industry. From the demand side, that provides a specific interpretation of the Armington assumption where quality differences reflect these distinct varieties offered by firms. A key simplification in this strand of literature is the assumption of a representative firm within each country. It is also used by Helpman and Krugman (1985)<sup>90</sup> who combine the traditional and new trade theory within an integrated

---

Melitz- Style Firm Heterogeneity in CGE models: Technical aspects and implementations, [University Ca' Foscari of Venice, Dept. of Economics Research Paper Series No. 04/WP/2014](#).

<sup>85</sup> GAMS Development Corporation. General Algebraic Modeling System (GAMS) Release 24.2.1. Washington, DC, USA, 2013

<sup>86</sup> Harrison W.J. and K.R. Pearson, 1996 "Computing Solutions for Large General Equilibrium Models Using GEMPACK", *Computational Economics*, Vol. 9 (1996), pp.83-127

<sup>87</sup> Kehoe, Timothy J., Clemente Polo, and Ferran Sancho, “An Evaluation of the Performance of an Applied General Equilibrium Model of the Spanish Economy,” Federal Reserve Bank of Minneapolis Research Department Working Papers, October 1994, (480)

<sup>88</sup> Helpman, E. (1981), “International Trade in the Presence of Product Differentiation, Economies of Scale, and Monopolistic Competition”, *Journal of International Economics* 11, 305-340

<sup>89</sup> Ethier, W.J. (1982), “National and International returns to Scale in the Modern Theory of International Trade”, *American Economic Review* 72, 225-238

<sup>90</sup> Helpman, E. and P. Krugman (1985), *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*, Cambridge, MA: MIT Press

equilibrium approach to provide an explanation for the pattern of trade where both inter-industry trade and “cross-hauling” can take place.

Increased availability of micro data on plant and firm level since the late 1980s generated empirically evidence challenging these theories of international trade. It became clear that firms are heterogeneous in terms of productivity, export behavior response to trade shocks, and other economic characteristics even within narrowly-defined industries (“one variety”). Specifically, analysis of micro datasets on firms and plants showed that i) only a small minority of firms actually engage in export (Bernard and Jensen 1995, Bernard et al. 2007) while there is considerable variation in export market participation rates across industries (Bernard et al. 2007); ii) exporters are more productive and larger than non-exporters (Bernard and Jensen 1995, Bernard et al. 2003,2007); iii) productivity dispersion exists across coexisting firms within any sector (Bartelsman and Dom 2000<sup>91</sup>) and among exporters in the number of markets (Eaton et.al 2004); iv) there are substantial reallocation effects within an industry following trade liberalization episodes where more productive firm replaces less productive ones (Foster et.al 2001<sup>92</sup>, Aw et al. 2001<sup>93</sup>) increasing aggregate productivity; and v) and endogenous changes in firm productivity via shifting market shares influences within-industry resource allocation (Bernard et al. 2006,2010) and therefore the aggregate productivity.

None of the above-mentioned observations found in micro datasets can be explained with the simplifying assumption of a representative firm within countries or industries made both in the traditional and “new” trade theories. This led to the development of international trade models with firm heterogeneity like that of Melitz (2003). He introduced firm heterogeneity into the model of Krugman (1980) which considers horizontal differentiation into varieties and increasing returns to scale.

Further, the role of the extensive margin, i.e. export of goods not exported before, in observed international trade patterns recently received increased attention. Several studies highlighted the importance of new varieties in export markets and the related welfare implications (see Romer, 1994<sup>94</sup>; Feenstra, 1994<sup>95</sup>; Broda and Weinstein, 2006<sup>96</sup>; Balistreri et al, 2010<sup>97</sup>; Kehoe and Ruhl, 2013<sup>98</sup>; among others). That strand of literature is complemented by more theoretical works which discuss impacts of trade liberalization on the extensive margin (Hummel and Klenow, 2005; Chaney, 2008,

---

<sup>91</sup> Bartelsman, E.J., Doms, M., 2000. Understanding productivity: lessons from longitudinal microdata. *Journal of Economic Literature* 38, 569–594

<sup>92</sup> Foster, L., Haltiwanger, J., Krizan, C., 2001. Aggregate productivity growth: lessons from the microeconomic evidence. In: Dean, E., Harper, M., Hulten, C. (Eds.), *New Directions in Productivity Analysis*. University of Chicago Press, pp. 303–363

<sup>93</sup> Aw, B.Y., Chen, X., Roberts, M.J., 2001. Firm-level evidence on productivity differentials and turnover in Taiwanese manufacturing. *Journal of Development Economics* 66, 51–86

<sup>94</sup> Romer, P.M., 1994. New goods, old theory, and the welfare costs of trade restrictions. *Journal of Development Economics* 43, 5–38

<sup>95</sup> Feenstra, R.C., 1994. New product varieties and the measurement of international prices. *The American Economic Review* 84, 157–177

<sup>96</sup> Broda, C., Weinstein, D.E., 2006. Globalization and the gains from variety. *Quarterly Journal of Economics* 121, 541–585

<sup>97</sup> Balistreri, E.J., Hillberry, R.H., Rutherford, T.F., 2010. Trade and welfare: does industrial organization matter? *Economics Letters* 109, 85–87

<sup>98</sup> Kehoe, Timothy J. and Kim J. Ruhl. 2013 . How Important Is the New Goods Margin in International Trade? (Revised version) Staff Report 324, Federal Reserve Bank of Minneapolis;

among others). Traditional applied general-equilibrium models based on the Armington assumption cannot depict changes at the extensive margin of trade<sup>99</sup> as an important source of new trade (Kehoe, 2005). Heterogeneous firm based models like that of Melitz (2003) overcome that shortcoming by depicting changes simultaneously at the intensive and extensive margins which motivates the implementation into CGEBox discussed in here. The logic of the Melitz model is that actions that facilitate trade will raise both export variety and average productivity.

#### **3.15.4 Implementation of Melitz model into CGEBox**

The standard GTAP model (Hertel 1997), developed by the Centre for Global Trade Analysis is a global, multi-regional, comparative static CGE based on neoclassical assumptions and equilibrium conditions that follow Walras' law. Policies are depicted by fixed relative price wedges. The model and variants thereof are the most widely used tools for the ex-ante analysis of economy-wide trade effects of multilateral or bilateral trade agreements. Based on a modified Cobb-Douglas Utility function, national income in each region is allocated among three types of final aggregate demand agents, namely government, private households, and savings. Each aggregate agent features its own Armington composite of domestic produce and aggregate imports for each product category, while the aggregate import composition for each product category is determined by a shared second Armington nest which also encompasses intermediate demand. Final demand expenditures on the aggregated Armington commodities reflect utility maximization, in the case of the representative private household drawing on a non-homothetic constant difference of elasticity expenditure (CDE) function; in case of the government and investment agents based on constant elasticity of substitution (CES) utility function. Markets are assumed competitive.

Production in each country and all sectors assumes a constant return to scale technology drawing on nested CES functions. In the standard GTAP model, the top level nest is a Leontief aggregate of value added and intermediate input use; the composition of the latter is based on fixed physical input coefficients. The value added nest allows for substitution between primary factors. As for the final demand agents, each sector features its own Armington nest to determine the composition of intermediate input demand for each commodity from domestic product and imports. The import composition is however identical across sectors and final demand, as mentioned above. Primary factors can either be assumed to be perfectly mobile across sectors such that the law of one price holds, or can be treated as “sluggish” based on Constant elasticity of transformation (CET) specification such that return to factors can differ between sectors.

In the standard GTAP framework, saving and capital is determined endogenously through a fictitious Global Bank. The Global Bank allocates investment across regions such that it equates the changes in the expected returns across countries. In the model, ad-valorem wedges can depict policy induces impacts on product price at the level of production, export, import and final consumption. The FOB (free on board) prices are differentiated by exporter and hence reflect bi-lateral export taxes or subsidies, adding international transport margins defines the CIF (cost, insurance and freight price) for each importer to which import taxes or subsidies are added. That allows for a rather detailed analysis of trade policy.

---

<sup>99</sup> Eaton et al. (2004) and Hillberry and Hummels (2008) defines the role of extensive margin in terms of firms serving a market while Hummels and Klenow (2005) and Broda and Weinstein (2006) identify the extensive margin in terms of the role of change in the number of products a firm trades or in the number of its trade partners (countries).

We now turn to the implementation of the Melitz (2003) model into the CGEBox version developed by Van der Mensbrugghe and Britz (2015). The model's equations are written in levels, and not as a mix of equations in levels and in linearized relative differences as found in GEMPACK based CGE implementations. That CGEBox<sup>100</sup> version allows for an exact replica of the standard GEMPACK version as developed by Hertel (1997). However, based on a flexible and modular code structure, it can also accommodate different assumptions of which we mention only some important ones. On the production side, non-diagonal make matrices, potentially combined with a CET approach are possible, while a flexible nesting approach allows more complex CES nests such as e.g. found in GTAP-AGR and GTAP-E. The Armington nests can be combined with a two-stage CET approach which allows for price dependent supply changes in with regard to the shares of domestic sales and total exports, respectively bi-lateral exports. The model can also be used in recursive-dynamic fashion. In the following, we only refer to the standard GTAP replica.

The actual implementation of the Melitz model into CGEBox draws to a large extent on the empirical method by Balistreri and Rutherford (hereafter BR) (2013) to introduce the Melitz (2003) model into an applied equilibrium model. Differences are detailed below. Further, we show how the Melitz structure compares to the Armington one. In our equation structure, sectors either are based on the standard representation, i.e. a two stage Armington structure on the demand side combined with constant return to scale industries, or follow the Melitz model with monopolistic composition and a different demand representation as detailed next.

The monopolistic competition sectors draw strictly on the framework of Melitz (2003): each firm produces one single unique variety over a continuum of varieties under conditions of monopolistic competition arising from imperfect substitution in demand for these varieties. Accordingly, the number of varieties produced in a regional industry is equal to the number of firms operating. New firms can freely enter an industry by paying a fixed entry cost which is thereafter sunk. However, before a firm enters, it is uncertain about its productivity level which becomes known once the sunk entry cost is paid. That productivity level is determined by a draw from a given productivity distribution. Once a firm knows its productivity level, it will only operate on those trade links, i.e. serving the domestic market or a specific export destination, where its profits are positive. The latter are defined as total sales revenues on a trade link minus the bilateral fixed cost of trade plus per unit variable cost times sales quantity, the per unit variable cost are assumed to be independent from the output level and the trade link. All firms face the same bilateral fixed cost on each link; however these costs differ across the trade links. Accordingly, the individual firms' decision to operate on a specific link depends on its productivity level which determines per unit variable cost. At given fixed costs of bi-lateral trade, given potential demand for an additional variety on a trade link and the price received for it, there is hence a cut-off productivity ( $\tilde{\varphi}_{rs}$ ) level beyond which profits become negative.

Those firms which draw a productivity level higher than that zero-profit cutoff productivity will operate on the trade link and those below the cutoff level will not. The firm with the productivity level equal to the cut off level is called the marginal firm and faces zero profits on that link; all other firms operating on that link make positive ones. A reduction in bi-lateral fixed cost of trade or higher demand, for instance from trade liberalization, opens hence a window of opportunity for less

---

<sup>100</sup> Tom Rutherford provides since a long time a GAMS based implementation of a CGE called GTAPinGAMS which however differs somewhat to the standard GTAP model. The M-B implementation is coded in "Dervis KJ, de Melo J, Robinson S (1982) General equilibrium models for development policy. Cambridge University Press, New York tradition" and departs significantly from the nomenclature used by Rutherford and that used by in the GEMPACK version of the model.



productive firms to serve new trade links which benefits consumer by providing more diversity in the product bundle imported from a specific origin.

There is no restriction on the number of markets that each firm can serve. The same firm is typically not active on all trade links as they differ in fixed bilateral cost of trade. Since serving the domestic markets requires lower bilateral (here country to same country) fixed cost, the more productive firms participate in export markets. Considering that intra-industry differentiation allows to depict impacts of trade policy changes on captured in conventional trade models where a trade policy leads only to re-allocation of resources between industries. In firm heterogeneity models however, a trade policy induces additionally re-allocation of resources within each industry. Firms can expand their market shares by absorbing resources of less productive ones forced to exit.

In these models, a policy that reduces the worldwide barriers to trade thus increases profits that existing exporters can earn in foreign markets and reduces the export productivity cutoff above which firms export. Input demand within the industry rises, due both to expansion by existing exporters and to new firms beginning to export. The increase in input demand bids up factor prices and reduces the profits of non-exporters. Reduction in profits in the domestic market induces some low productivity firms who were previously marginal to exit the industry. As low productivity firms exit and output and production factors are reallocated towards higher-productivity firms, average industry productivity rises.

Comparing the constant returns-to-scale sectors in the standard GTAP model to the monopolistic competition sector in Melitz reveals three main differences. First, the standard GTAP model is an aggregate industry level framework capturing the behavior of a representative firm in a perfectly competitive industry. In opposite to that, firms face fixed cost of entry in the monopolistic competition framework which leads to increasing return to scale. Second, the standard GTAP model with one representative firm in each industry cannot reflect productivity differences as depicted in the Melitz model. Third, in the standard GTAP model, consumer's utility is defined by an Armington composite of goods from different origins, while in monopolistic completion models it is defined over the Dixit-Stiglitz (1977)<sup>101</sup> composite of varieties differentiated by origin which allows for monopolistic competition between firms operating on the same trade link. The combination of these differences allows depicting the extensive margin of trade in CGE models with firm heterogeneity.

### **Algebraic representation of the firm heterogeneity into the GTAP model**

This section presents an algebraic representation of the implementation of the Melitz (2003) model as implemented in CGEBox. Note that the equations in GAMS code are documented above in section "Melitz model". The Melitz framework focuses on intra-industry differentiation where each firm produces a single unique variety. However, data at the firm level are limited and applied equilibrium models work at rather aggregate industry levels. Fortunately, Melitz offers a numerical framework build around (marginal changes of) the average firm operating on a trade link. That average firm's productivity comprises all necessary information on the distribution of productivity levels of firms active on that link. That vastly eases the model's implementation by effectively eliminating any data needs at individual firm level as detailed below. Against the background of that definition of an average firm on each trade link, we now focus on the formulation of an empirically computable version of Melitz model and its linkages with the GTAP model.

---

<sup>101</sup> Dixit, A., and J. Stiglitz. 1977. "Monopolistic competition and optimum product diversity." *American Economic Review*, 67(3): 297–308. doi:10.1017/cbo9780511492273.004

Assume that a representative agent  $a$  (private households, government, investors, intermediate inputs by the different firms) in region  $s$  obtains utility  $U_{aais}$  from consumption of the range of differentiated varieties of product  $i$ . Considering the constant elasticity of substitution (CES) utility function as proposed by Dixit and Stiglitz (1997), the aggregate demand by each agent  $a$  for commodity  $i$  in region  $s$  ( $Q_{ais}$ ) which is equivalent to utility ( $Q_{ais} \equiv U_{aais}$ ) can be represented as

$$Q_{ais} = \left( \sum_r \int_{\omega \in \Omega_{irs}} \lambda_{airs} \frac{1}{\sigma_{is}} Q_{airs}(\omega)^{\frac{\sigma_{is}-1}{\sigma_{is}}} d\omega \right)^{\frac{\sigma_{is}}{\sigma_{is}-1}} \quad (5)$$

where  $\Omega_{irs}$  represents the set of products  $i$  sourced from region  $r$  to  $s$  and  $\omega \in \Omega_{irs}$  index the varieties in the set  $\Omega_{irs}$ . In this context,  $Q_{airs}(\omega)$  represents the demand quantity of commodity  $i$  for variety  $\omega$  in region  $s$  by agent  $a$  which is sourced from region  $r$ ,  $\sigma_i$  represents the constant elasticity of substitution for each commodity, and  $\lambda_{airs}$  are preference weights (share parameters)<sup>102</sup> which reflect differences between origins not linked to diversity in varieties. Note that substitution elasticities might be differentiated by destination region  $s$ , but are uniform across agents in each region in our implementation.

The resulting CES unit expenditure function which is the dual price index on Dixit-Stiglitz composite demand in region  $s$  ( $P_{ais}$ ) is given by

$$P_{ais} = \left( \sum_s \int_{\omega \in \Omega_{irs}} \lambda_{airs} P_{airs}(\omega)^{1-\sigma_{is}} d\omega \right)^{\frac{1}{1-\sigma_{is}}} \quad (6)$$

where  $P_{airs}(\omega)$  is agent's  $a$  (purchase) price of product  $i$  for variety  $\omega$  in region  $s$  sourced from region  $r$ . Using the aggregate price index in Melitz (2003) based on his definition of the average firm and considering that varieties do not differ in their marginal utility for the first unit, one can define the price index as equivalent to the dual price defined in (6)

$$P_{ais} = \left( \sum_r \lambda_{airs} N_{irs} \widetilde{P}_{airs}^{1-\sigma_{is}} \right)^{\frac{1}{1-\sigma_{is}}} \quad (7)$$

where  $\widetilde{P}_{airs}$  denotes the agent price inclusive of export, import and consumption taxes for the average firm, and  $N_{irs}$  refers to the number of firms operating on the trade link  $r$ - $s$ . Note that consistent with Melitz (2003), there is a one to one mapping between firms and varieties such that the number of firms is equal to the number of varieties on each trade link. Comparing to (2), which is based on the individual varieties, (3) summarizes the compositional change, i.e. change in the number of varieties which go along with an update of the average price. Note again that we assume the same substitution elasticities across agents

The total ( $Q_{airs}$ ) and average per firm ( $\widetilde{Q}_{airs}$ ) demand for the average variety by an agent to be shipped from  $r$  to  $s$  ( $\widetilde{Q}_{airs}$ ) can be obtained by applying Shephard's Lema on the expenditure function:

<sup>102</sup> The reader should note that the share parameters are absent in the original Melitz paper. We hence allow here a differentiation between products from different origins as in the Armington model in addition to the love of variety effect.

$$Q_{airs} = \tilde{Q}_{airs} N_{irs} = \lambda_{airs} N_{irs} Q_{ais} \left( \frac{P_{ais}}{\bar{P}A_{airs}} \right)^{\sigma_{is}} \quad (8)$$

Which reveals the main differences to a standard Armington composite: the share parameters vary with the number of operating firms, i.e. the number of varieties comprised in the bilateral trade bundles. As the agent demand for the average firm's output in region  $r$  in each industry  $i$  in region  $s$  ( $Q_{airs}$ ) depends on the aggregate regional demand for that industry  $Q_{ais}$ , we need to determine this in equilibrium for each agent. In other word, we need to determine the demand for use of  $i$  as an intermediate input and as final demand for household consumption, government consumption, investment, and for international transport margins. In the standard GTAP model, each agent has a specific preference function which determines the demand for her Armington commodity; the government and saving sector based CES preferences while households used a CDE indirect demand function. The Armington demand for each agent and commodity is then decomposed into a domestic and import component in a first Armington nest. The second one decomposes import demand by each region by origin, independent of the agent.

The implementation of the Melitz model thus simplifies the demand structure present in the standard GTAP model by aggregating the two Armington nests into a single one, however, note that the GTAP data base so far does not differentiate in the SAM bi-lateral flows by agent. We hence used the same shares by origin to split up import demand for the different agents.

Assume a small profit maximizing firm facing the constant elasticity of demand according to (8) for its variety. Based on the assumption of the large group monopolistic competition, firms will not consider its impact on the average price index and therefore follow the usual markup rule to translate their marginal cost of production to the optimal price.

Firms in Melitz (2003) face different types of cost: sunk fixed cost of entry  $f^{ie}$ , fixed cost of operating on a trade link  $f_{irs}$  and marginal cost  $c_{ir}$ . Let  $\varphi_{irs}$  indicate the firm's specific productivity which measures the amount of "variable composite unit" needed per unit of output  $Q_{irs}$ . Accordingly, the marginal cost per unit is the amount of "composite input" required per  $\left(\frac{1}{\varphi_{irs}}\right)$  times the unit cost of the "variable composite input" ( $c_{ir}$ ) in industry  $i$  of region  $r$ . Therefore, a firm wishing to supply  $Q_{irs}$  units from region  $r$  to  $s$  employs  $\left(f_{irs} + \frac{Q_{irs}}{\varphi_{irs}}\right)$  units of "variable composite input". Let,  $\tau_{irs}$  denote the fixed iceberg cost of trade which represent domestic production costs, and not the international trade margins present in GTAP. Focusing on the average firm with a productivity  $\tilde{\varphi}_{irs}$  operating on a trade link, and solving the firm's profit maximization problem, the price charged by the average firm in region  $r$  to supply region  $s$   $\tilde{P}\tilde{F}_{irs}$  (inclusive of domestic transport margin) is

$$\tilde{P}\tilde{F}_{irs} = \frac{\sigma_{is}}{\sigma_{is} - 1} \frac{\tau_{irs} c_{ir}}{\tilde{\varphi}_{irs}} \quad (9)$$

where  $\frac{\sigma_{is}}{\sigma_{is}-1}$  represents the constant markup ratio in industry  $i$  which reflects market power due to product differentiation into varieties. The linkages between the firm and agent prices are provided in Appendix 1.

The average price in (5) depends hence on the price of variable composite input  $c_{ir}$ , which is a function of the price of intermediates and primary factors. Given the assumption of constant return to scale and the way technology is presented in the standard GTAP model, the unit cost function for sector  $i$  in region  $r$   $c_{is}$  in GTAP is given by the Leontief composite of the value added bundle (CES aggregate of factors of production) and the aggregate of intermediate demand (Leontief aggregate of intermediate demands). In the CGEBox,  $m\_px$  is a macro defined as producer price which constitute per unit costs corrected for production taxes. To be consistent with our Melitz formulation, the unit cost inclusive of production tax is directly introduced in the markup equation (5). It should be emphasized that the presence of fixed cost in the Melitz model is the source of increasing returns to scale in a monopolistically competitive industry: if firms expand production, the fixed cost can be distributed over a larger outputs such that per unit cost decrease.

While observed data on quantities traded and related prices allow identifying the necessary attributes on the average firm, additional information is needed to gain information about the marginal firm, i.e. the firm which earns zero profit. Obviously, the distance in productivity between the average and marginal firm reflects properties of the underlying distribution. We rely here on a Pareto Productivity distribution which has analytical tractability (Cheney, 2008) and was shown as empirically relevant (Axtell, 2001<sup>103</sup>; Luttmer, 2002<sup>104</sup>; Gabaix, 2008<sup>105</sup>; Eaton et al., 2001<sup>106</sup>).

Let  $M_r$  indicate the number of firms choosing to incur the fixed entry cost, i.e. total industry size, each individual firm receives its productivity  $\varphi$  draws from a Pareto distribution with Probability Density Function (PDF)

$$g(\varphi) = a \left( \frac{b^a}{\varphi^{a+1}} \right) = \frac{a}{\varphi} \left( \frac{b}{\varphi} \right)^a \quad (10)$$

and Cumulative Distribution Function (CDF)

$$G(\varphi) = 1 - \left( \frac{b}{\varphi} \right)^a \quad (11)$$

where  $b$  is the minimum productivity and  $a$  is a shape parameter. Lower values of the shape parameter imply higher productivity dispersion among firms. As discussed in Melitz (2003),  $a > \sigma_i - 1$  should be applied in order to ensure a finite average productivity level in the industry.

On each bilateral trade link, the given the fixed bilateral trade cost, variable costs and demand define jointly a certain cut off productivity level ( $\varphi_{rs}^*$ ) at which firms will receive zero profit. A firm which has drawn the productivity equal to that threshold level ( $\varphi_{rs} = \varphi_{rs}^*$ ) will hence face zero profits and

---

<sup>103</sup> Axtell, Robert L. 2001. "Zipf Distribution of U. S. Firm Sizes." *Science*, 293(5536): 1818–20. Crozet, Matthieu, and Pamina Koenig. 2007. "Structural Gravity Equation with Extensive and Intensive Margins." Unpublished

<sup>104</sup> Luttmer, Erzo G. J. 2007. "Selection, Growth, and the Size Distribution of Firms." *Quarterly Journal of Economics*, 122(3): 1103–44

<sup>105</sup> Gabaix, Xavier. 2008. "Power Laws." In *The New Palgrave Dictionary of Economics*. 2nd edition, ed. Steven N. Durlauf and Lawrence E. Blume. New York: Palgrave Macmillan

<sup>106</sup> Eaton, J., S. Kortum, and F.Kramarz. 2001. "An anatomy of international trade: evidence from French firms." *Econometrica*, 79(5): 1453–1498. doi:10.3386/w14610

act as the marginal firm from region  $r$  supplying  $s$ . Those firm whose productivity is above the threshold level ( $\varphi_{rs} > \varphi_{rs}^*$ ) will receive a positive profit and will operate on the  $r - s$  link and those firm whose productivity is below the threshold level ( $\varphi_{rs} < \varphi_{rs}^*$ ) will not operate on the  $r - s$  link. Focusing on the fixed operating cost  $f_{irs}$  in composite input units, the marginal firm on  $r-s$  link receives zero profit at

$$c_{ir}f_{irs} = \frac{r(\varphi_{irs}^*)}{\sigma_i} \quad (12)$$

where  $r(\varphi_{irs}^*) = p(\varphi_{irs}^*)q(\varphi_{irs}^*)$  denotes the revenue of marginal firm at the productivity equal to the cut off level ( $\varphi_{irs} = \varphi_{irs}^*$ ).

The zero cut off productivity level in each bilateral market  $\varphi_{irs}^*$  can be obtained by solving (12). However, it is numerically easier to define this condition in terms of the average rather than the marginal firm. To do this, we define the productivity and revenue of the average firm relative to that of the marginal firm.

The average productivity in a trade link is determined by the productivity level of the operating firms on that link which by definition are at or above the cutoff productivity level. Following Melitz (2003), that average productivity is defined as the CES aggregation of productivities of all firms operating on a given trade link

$$\tilde{\varphi}_{irs} = \left[ \frac{1}{1 - G(\varphi_{irs}^*)} \int_{\varphi_{irs}^*}^{\infty} \varphi_{irs}^{\sigma_{is}-1} g(\varphi) d\varphi \right]^{\frac{1}{1-\sigma_{is}}} \quad (13)$$

If these productivities are Pareto distributed, we can write<sup>107</sup>

$$\tilde{\varphi}_{irs} = \left[ \frac{a}{(a + 1 - \sigma_{is})} \right]^{\frac{1}{1-\sigma_{is}}} * \varphi_{irs}^* \quad (14)$$

Eq. (10) provides the relationship between the productivities of the average and marginal firm (for further details see Allen and Arkolakis (2016)<sup>108</sup>.

Using optimal firm pricing according to (9) and given the input technology, the ratio of revenues of the firms with marginal productivity  $r_{irs}(\varphi^*)$  in relation to the revenue of the firm with the average productivity  $r_{irs}(\tilde{\varphi})$  is defined as

---

<sup>107</sup> One could use industry specific shape parameter ( $a_i$ ) given the availability of data at sectoral level. In this study we assume that all firms entering in different industries draw their productivity from the Pareto distribution function with same characteristics (i.e. same scale and shape parameter).

<sup>108</sup> Allen, T., and C. Arkolakis (2016): “Elements of Advanced International Trade,” Unpublished: available at: <http://www.econ.yale.edu/~ka265/teaching/GradTrade/notes/ClassNotes.pdf>.

$$\frac{r_{irs}(\varphi^*)}{r_{irs}(\tilde{\varphi})} = \left(\frac{\varphi^*}{\tilde{\varphi}}\right)^{\sigma_i} \quad (15)$$

Solving (14) for  $\frac{\varphi^*}{\tilde{\varphi}}$ , substituting it into (15), and then solving the resulting equation for  $r_{irs}(\varphi^*)$  and replacing its value in (12), defines a relation between the bilateral fixed cost at current composite input price (the left hand side of (12) below), the average firms revenue ( $\widetilde{PF}_{irs} \tilde{Q}_{irs}$ ), the shape parameter of the Pareto distribution of the productivities and the substitution elasticity of demand:

$$c_{ir}f_{irs} = \frac{(a + 1 - \sigma_{is})}{a\sigma_{is}} \widetilde{PF}_{irs} \tilde{Q}_{irs} \quad (16)$$

Note that average firm's sale in region  $r$  in each industry  $i$  to region  $s$  ( $\tilde{Q}_{irs}$ ) at the equilibrium is composed of the demand for use of  $i$  by different agents<sup>109</sup>.

The optimal pricing in the markup equation (5) requires information on the average productivity on each bilateral trade link. In Melitz (2003), the probability that a firm will operate is  $1 - G(\varphi_{irs}^*)$  which is equal to the fraction of operating firms over total number of firms choosing to draw their productivity ( $\frac{N_{ir}}{M_{ir}}$ ). Using the Pareto cumulative distribution function (11) and inverting it we have

$$\varphi_{irs}^* = \frac{b}{\left(\frac{N_{irs}}{M_{ir}}\right)^{\frac{1}{a}}} \quad (17)$$

substituting (17) into (14) results in

$$\tilde{\varphi}_{irs} = b \left[ \frac{a}{(a + 1 - \sigma_{is})} \right]^{\frac{1}{1 - \sigma_{is}}} * \left(\frac{M_{ir}}{N_{irs}}\right)^{-\frac{1}{a}} \quad (18)$$

Next, the number of firms selecting to enter the market  $M_{ir}$  is determined. Based on the free entry condition, the last firm which enters has expected profits over its life time which just offset the sunk cost of entry. Industry entry of a firm requires a one-time payment of  $f^{ie}$ . An entered firm faces a probability of  $\delta$  in each future period to suffer a shock which forces its exit. Therefore,  $\delta M_{ir}$  firms are lost in each period. Based on Melitz (2003), in a stationary equilibrium, the number of aggregate variables must remain constant over time, including industry size. This requires that the number of new entrants in every period is equal to the number of firms lost  $\delta M_{ir}$ . Therefore, total entry cost is equal to  $c_{ir} \delta M_{ir} f^{ie}$ . Each firm faces the same expected share on that cost, i.e.  $c_{ir} \delta f^{ie}$  if risk neutral behavior and no time discounting is assumed. The firm's expected share on entry costs must be equal to the flow of expected profit on the condition that firm will operate.

$$\tilde{\pi}_{irs} = \frac{\widetilde{PF}_{irs} \tilde{Q}_{irs}}{\sigma_{is}} - c_{ir}f_{irs} \quad (19)$$

---

<sup>109</sup>  $\tilde{Q}_{irs} = \sum_a \tilde{Q}_{airs}$

The probability that a firm will operate on the r-s trade link is given by the ratio of  $\frac{M_{ir}}{N_{ir}}$ <sup>110</sup>. Thus, the free entry condition ensures that the expected industry profits, i.e. the profits summed up over all potential bilateral trade links, is equal to the annualized flow of the fixed costs of entry

$$c_{ir} \delta f^{ie} M_r = \sum_s N_{irs} \widetilde{PF}_{irs} \widetilde{Q}_{irs} \frac{\sigma_{is} - 1}{a\sigma_{is}} \quad (20)$$

where zero profit condition (16) is used to replace the fixed operating cost  $c_{ir}f_{irs}$ .

With the number of entered firm established in (20), we now turn to total composite input demand of the industry Y which consists of three components: sunk entry costs of all entrants ( $\delta M_{ir} f^{ie}$ ), operating fixed cost ( $\sum_s N_{irs} f_{irs}$ ) on each trade link and variable costs ( $\sum_s N_{irs} \frac{\tau_{hrs} \widetilde{Q}_{irs}}{\widetilde{\varphi}_{irs}}$ ). Therefore, composite input demand is defined as

$$Y_{ir} = \delta M_{ir} f^{ie} + \sum_s N_{irs} (f_{irs} + \frac{\tau_{irs} \widetilde{Q}_{irs}}{\widetilde{\varphi}_{irs}}) \quad (21)$$

This equation provides the link to the equations in the GTAP model describing the technology and related costs. Table 17 summarizes full set of Melitz equations which are introduced into the GTAP model.

Table 17: Equilibrium conditions in the Heterogenous Firm model

| Equation                                                   | Equilibrium condition             | Associated variable                           |
|------------------------------------------------------------|-----------------------------------|-----------------------------------------------|
| (7)                                                        | Sectoral Aggregation              | $P_{is}$ : Sectoral price index               |
| (8)                                                        | Firm-level demand                 | $\widetilde{PF}_{irs}$ : Average firm price   |
| (9)                                                        | Firm-level Pricing                | $\widetilde{Q}_{irs}$ : Average firm quantity |
| (16)                                                       | Zero cut off condition            | $M_{ir}$ : Number of operating firms          |
| (18)                                                       | CES Weighted Average productivity | $N_{ir}$ : Average firm productivity          |
| (20)                                                       | Free entry condition              | $N_{irs}$ : Number of entered firm            |
| (21)                                                       | Factor market clearing condition  | $c_{ir}$ : Sectoral composite input price     |
| Variables through which Melitz model is linked to the GTAP |                                   |                                               |
| $Q_{ais}$ : Sectoral Demand                                |                                   |                                               |

<sup>110</sup> The probability that the firm will operate is  $1 - G(\varphi_{irs}^*) = \frac{N_{ir}}{M_{ir}}$

$c_{ir}$ : Price of composite input

$Y_{ir}$ : Sectoral composite input demand

### Calibration of the model

In order to apply the above equation structure, the different parameters must be chosen such as to recover an observed benchmark. That benchmark consists of the global SAM provided by the GTAP data base<sup>111</sup> against which the remaining equations of the GTAP model are calibrated as well. It comprises values on domestic sales and on bilateral international trade expressed in USD million for each sector and region which are key data in the context of the Melitz module discussed above. Global detailed SAMs comprise many small entries both in relative and absolute terms which can affect the numerical stability during solution of a CGE. We therefore recommend filtering out in a systematic way small transactions in relative terms from the database when using the model with many sectors and regions, following by rebalancing the global SAM.

We focus here mainly on the calibration of the firm heterogeneity module while the calibration of Armington sectors are similar to the standard GTAP model and are not further discussed. In the following, a superscript “0” denotes a benchmark value, for instance,  $Q_{is}^0$  denotes the benchmark value of demand for commodity  $i$  in region  $s$ . In order to line up the variables in the Melitz module with the SAM, the following identities must hold:

$$c_{ir}^0 Y_{ir}^0 \equiv vom_{ir} \quad (22)$$

$$p_{is}^0 Q_{ais}^0 \equiv xafm_{ais} \quad (23)$$

where  $vom_{ir}$ , represent the production value of commodity  $i$  in region  $r$  at producer tax inclusive prices and  $xafm_{ais}$  represent the value of each agent’s demand for commodity  $i$  in region  $s$ , consumer tax inclusive. The first identity indicates that the cost of input supply must equal the value of output, and the second identity ensures that demand for all goods and factors is equal to supply at the benchmark. Accordingly, using the conventional choice of unity prices, i.e.  $p_{irs}^0 = 1$ , and  $c_{irs}^0 \equiv 1$ , total quantity demanded  $Q_{is}^0$  and total input supplied  $Y_{ir}^0$  is locked down.

It should be noted here again that the bilateral import demand in the SAM is aggregated over agents. The RHS entries in equation (19) above are hence constructed by splitting up that total by each agent’s share on total imports.

Given the agents’ demand for each commodity  $Q_{ais}^0$  and using the definition of the firm average in Melitz, we have

$$Q_{ais}^0 = \sum_r Q_{airs}^0 = \sum_r \tilde{Q}_{airs}^0 N_{irs}^0 \quad (24)$$

We now briefly compare the demand side based on the Armington assumption with the Melitz model to demonstrate that the choice of  $N_{irs}^0$  does not matter for the model’s simulation

<sup>111</sup> We currently use the GTAP9 Data Base which carries a snapshot of the 2011 world economy, covering 140 regions (aggregate of 226 countries) and 57 sectors. But all GTAP Data Bases share the basic structure such that the code can also be used with other releases.



behavior. Under the Armington assumption, a CES utility function is used to differentiate between origins. That implies that the (average) quality on the trade links can differ along with the price. As the resulting demand function is homothetic, expenditure shares are independent of the income level and solely depend on relative prices for these qualities. Conveniently, calibration is performed at given substitution elasticities by choosing share parameters such that given expenditure shares are recovered at given relative prices.

The Ditz-Stieglitz price index used in the Melitz (2003) is also based on a CES utility function; however, here we have a continuum of varieties. As share parameters are absent in the original model, the marginal utility of the first unit for each variety is the same. Each firm is assumed to produce its own variety and multiple firms are allowed to operate on a trade link. Thus, more firms imply more varieties and a higher utility per unit of traded output on that link. Total demand on a link is hence defined as the product of average output per operating firm and the number of firms operating on that link. Recovering expenditures shares at given prices for each trade link can hence be based either by deriving the number of firms on a link at given average firm output on that link or by pre-selecting the number of firms for each link and deriving average outputs. Share parameters are not needed; however, as a consequence, the resulting price index cannot be controlled. As only relative prices matter, that only affects readability and not simulation behavior.

In order to improve readability and provide a combined interpretation of the two models, we introduce share parameters in the CES-demand function used in the Melitz model. That allows a convenient interpretation of the extension introduced by the Melitz model: changes in the number of firms lead to preference shifts in the Armington model as they update the share parameters. The resulting price index  $P_{ais}$  and matching demands  $\tilde{Q}_{airs}$  for product  $i$  in region  $s$  from origins  $r$  by agent  $a$  can be depicted as

$$P_{ais}^0 = \left( \sum_r \lambda_{airs} \frac{N_{irs}}{N_{irs}^0} \widehat{P}_{A_{airs}}^0 \right)^{\frac{1}{1-\sigma_{is}}} \quad (25)$$

$$Q_{airs}^0 = \lambda_{airs} \frac{N_{irs}}{N_{irs}^0} Q_{ais}^0 \left( \frac{P_{ais}^0}{\widehat{P}_{A_{airs}}^0} \right)^{\sigma_{is}} \quad (26)$$

where  $N_{irs}$  depicts the number of operating firms on a trade link. Note that, as  $N_{irs}$  and  $\lambda_{airs}$  (the share parameter) are identical in the calibration point, we scale without any impact on the simulation behavior the lambdas  $\lambda_{irs}$  such that the price index  $P_{is} \equiv 1$  (or something else) at the benchmark. That allows to recover any given quantity index  $Q_{ais}$  and matching  $P_{ais}$ . Equations 25 and 26 show the link between the Armington and Melitz models. A given Armington model can hence be simply extended by adding changes in the number of operating firms. With the choice of  $\widehat{P}_{A_{airs}}^0 = 1$ ,  $P_{ais}^0 = 1$ , since we now have the observed value of  $Q_{airs}^0$ ,  $Q_{ais}^0$ , and estimated value  $\sigma_{is}$ <sup>112</sup>,

---

<sup>112</sup> In GTAP the first level Armington demand for each agent  $XA_{i,h}$ ,  $XA_{i,gov}$ ,  $XA_{i,inv}$  is a CES composite of domestic and aggregate imported good with the substitution elasticity  $\sigma_m$ ; and the second level Armington demand (aggregated import) is a CES composite of import demand by each region of origin with the substitution elasticity of  $\sigma_w$ . However, in Melitz structure represented consume has Dixit – Stiglitz preferences over the varieties including domestically produced commodities and imported commodities by source of origin with the substitution elasticity of  $\sigma_i$ . To be consistent with Dixit – Stiglitz framework,  $\sigma_i$  is derived based on the weighted average of the substitution elasticities of first and second level Armington nests where the weights given to

the value of  $\lambda_{airs}$  is calibrated in a way that we do not need the information on the benchmark number of firms. Accordingly, the values of  $\lambda_{airs}$  are recovered by inverting the demand functions.

$$\lambda_{airs} = \frac{Q_{airs}^0}{Q_{ais}^0} \left( \frac{\widetilde{PA}_{airs}^0}{P_{ais}^0} \right)^{\sigma_{is}} \quad (27)$$

Let  $\widetilde{PF}_{irs}$  denote the average firm's offer price on a specific trade link, from which  $\widetilde{PF}_{airs}$  is derived by considering export taxes (to arrive at cif), international transport margins (to arrive at fob), and import and consumption taxes.

The composite input demand linked to bilateral fixed cost  $f_{irs}$  is derived from the zero-profit cut off condition, where the values of  $c_{ir}^0$  and  $\widetilde{PF}_{irs}^0$  are set to unity at the benchmark,  $\widetilde{Q}_{irs}$  is estimated from (23), and the value of shape parameter  $a$  is taken from literature<sup>113</sup>:

$$f_{irs} = \frac{\widetilde{PF}_{irs}^0 \widetilde{Q}_{irs}^0}{c_{ir}^0} \frac{(a + 1 - \sigma_i)}{a\sigma_i} \quad (28)$$

Please note that  $Q_{irs}^0 = \widetilde{Q}_{irs}^0 N_{irs}^0$  implies that changing the number of firms at the benchmark updates  $f_{irs}$ , but not the total industry cost linked to bi-lateral fixed cost.

Similarly, the free entry condition allows deriving the composite input demand of sunk entry cost  $c_{ir} \delta f^{ie}$  at given price and quantity of the average firm:

$$c_{ir} \delta f^{ie} = \sum_s \frac{N_{irs}^0}{M_r^0} \widetilde{PF}_{irs}^0 \widetilde{Q}_{irs}^0 \frac{\sigma_{is} - 1}{a\sigma_{is}} = \sum_s \frac{1}{M_r^0} \widetilde{PF}_{irs}^0 Q_{irs}^0 \frac{\sigma_{is} - 1}{a\sigma_{is}} \quad (29)$$

Note again that the given values on the trade links  $Q_{irs}^0$  determine together with the chosen industry size the sunk entry costs such that, as above, either the number of operating firms or the average quantities can be chosen without affecting the calibration. Furthermore, the condition reveals that the choice of total industry size  $M_r^0$  does not matter either, as the sunk entry costs will adjust

---

$\sigma_m$  and  $\sigma_w$  are domestic and import purchases at agent price, respectively. We also insured that derived substitution elasticity is smaller than  $a + 1$  to ensure a finite average productivity level in the industry.

<sup>113</sup> Estimates of the value shape parameter vary and are conditional on a choice of elasticity of substitution. The choices of shape parameter and substitution elasticity are important as these key parameters have significant implications (for example on welfare). The importance of these variables is well discussed in Akgul et al. (2015, 2016) and Dixit (2016). Bernard et al. (2007) choose a shape parameter equal to 3.4, and estimates of Eaton et al. (2004) show  $a = 4.2$ , while Balistreri et al. (2011) find shape parameter of 5.17, 4.58 and 3.92 depending on different trade cost-distance elasticities but under the maintained assumption that  $\sigma = 3.8$ . Akgul, Z., N.B. Villoria, and T.W.Hertel.2016. "Theoretically-Consistent Parameterization of a Multi-sector Global Model with Heterogeneous Firms" Forthcoming calibrated shape parameter of 2.89 for manufacturing sector consistent with a shape parameter obtained Spearot, A. 2016. "Unpacking the long run effects of tariff shocks: New structural implications from firm heterogeneity models."AEJ Microeconomics,8(2):128–67. doi:10.1257/mic.20140015. In this study we used the weighted average of the substitution elasticities of first and second level Armington nests which is not necessary consistent with the shape parameter (4.6) which is taken from Balistreri et al.(2011). Once decided which sectors are treated as sectors with heterogeneous firms, the theoretically consistent sectoral elasticity of substitution and sectoral shape parameter should be obtained. The minimum productivity parameter is chosen as  $b = 0.2$  following Balistreri et al. (2011) from Bernard et al. (2007).

proportionally such that the industry's total annualized costs of industry entry solely depend on the shape parameter of the Pareto distribution, the substitution elasticity and the given SAM values.

Next, given a minimum productivity parameter of the Pareto distribution  $b$ , the average firm productivity on trade link can be initialized as:

$$\tilde{\varphi}_{irs} = b \left[ \frac{a}{(a + 1 - \sigma_{is})} \right]^{\frac{1}{1 - \sigma_{is}}} * \left( \frac{M_r^0}{N_{irs}^0} \right)^{-\frac{1}{a}} \quad (30)$$

Having determined the value of average productivity on each r-s link, and setting  $c_i \equiv 1$ ,  $\widetilde{PF}_{irs}^0 \equiv 1$ , the domestic input demand related to the transport margin cost can be computed by inverting the mark up equation (9) as follows:

$$\tau_{irs} = \frac{\sigma_{is}-1}{\sigma_{is}} * \tilde{\varphi}_{irs}^0 \quad (31)$$

Here, it is less obvious why  $M_r^0$  and  $N_{irs}^0$  can be freely chosen without affecting simulation behavior.  $\tilde{\varphi}_{irs}$  and  $\tau_{irs}$  enter the total industry cost and the price markup equations. Total industry costs on a trade link r-s are given as:  $N_{irs} \left( f_{irs} + \frac{\tau_{hrs} \tilde{Q}_{irs}}{\tilde{\varphi}_{hrs}} \right)$  which shows that changes in  $\tilde{\varphi}_{irs}^0$  and  $\tau_{irs}$  resulting from different benchmark values of  $N_{ir}$  and  $M_r$  do not matter. The same holds for the price markup equation:  $\widetilde{PF}_{irs} = \frac{\sigma_{is}}{\sigma_{is}-1} \frac{\tau_{irs} c_{ir}}{\tilde{\varphi}_{irs}}$ .

Accordingly, sensitivity analysis with different values for  $M_r^0$  and  $N_{irs}^0$  showed no differences in simulated results. However, the choice might affect numerical stability by affecting the overall scaling of the model. In order to improve readability, we have generally chosen  $N_{irs}^0 \equiv 1$  which gives average firm outputs on each trade link equal to given bi-lateral respectively domestic sales values at benchmark prices.

### Technology nesting

Similar the GTAP-HET model, we apply a different nesting for variable costs of trade and the fixed costs related to industry entry and operating on a link. The variable costs maintain whatever nesting is chosen originally by the analyst. The fix cost only use fixed value added, applying the same substitution elasticity as used for the value added nest in the original model. Note that this is a first implementation which should be improved as it implies for industries using land or natural resources that they also become part of the fixed costs. Technically, the flexible nesting structure by CGEBox is applied by defining a new technology nest termed “fcost”. At the same time, the implementation of the nesting was extended to allow for the case that several nests demand the same intermediate input or factor. Alternatively, the model can be simplified to use the same composite input for variable and fixed costs. The implementation allows combining the default Melitz model with its differentiation between variable and fixed costs composites with other nesting structures on the production side such as the implementation of GTAE-E or GTAP-AGR realized as modules of CGEBox.

#### 3.15.5 Numerical stability and domestic industry demand for its own output

Test runs with the model in different configurations with regard to sectoral and regional detail as well as differently deep multi-lateral trade liberalization shocks have consistently shown non-stable

behavior resulting in infeasibilities in cases where an industry's cost share of domestic intermediate input is relatively high, say around 30%. Under these conditions, the model easily ends up in a vicious circle where a cost increase on the domestic link amplifies itself such that corner solutions with zero industry sizes provoke infeasibilities. Changing the parameterization seems to help only in some selected cases.

A firm should not apply mark-up pricing for own produced and used intermediates if that decreases its own competitiveness, i.e. provokes real costs and is not a tax evasion tactic. However, the average firm model does not differentiate domestic intermediate demand inside a sector between a firm's demand for its own produce and demand by other firms. In order to allow for a numerical stable implementation, we therefore allow excluding domestic intermediate demand for the diagonal I/O element from the love of variety effect, i.e. applying Armington preferences. Accordingly, there is also no markup pricing involved if that option is chosen. That has proven to effectively prevent the down-spiraling solution behavior in most cases. However, it can still happen if a bundle of sectors shows quite high costs share of domestic demand for the same bundle of sectors. In that case, the defense by arguing over the single firm's behavior is not longer valid.

Potential improvements in that regard could encompass a split up of the diagonal domestic intermediate demand into a share subject to the love of variety assumption and a remainder treated as competitive. But before introducing further complexity into the framework, more testing is required.

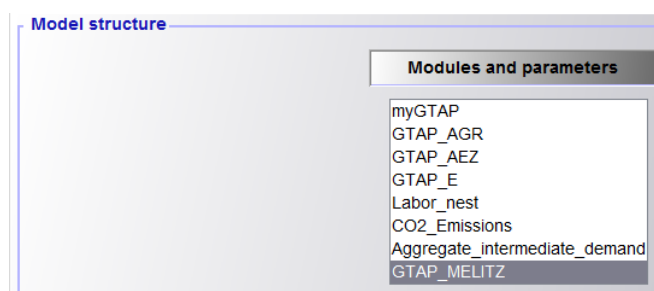
### 3.15.6 Technical implementation and an example application

#### GAMS code

##### Technical implementation

The necessary GAMS code for the calibration of the model can be found in "gams\GTAPMelitz\GTAPMelitz\_cal.gms". Besides the calibration steps discussed above, it is important to mention that certain flags are deleted to ensure that equations in the standard model are inactive for Melitz commodities.

The user can switch the module one on the Graphical User Interface under "Model Structure":



which generates a new tab with the following input possibilities.

Under Melitz commodities, the commodities / sectors can be chosen the parameter of the Pareto-distribution of the productivities defined.

The checkbox “Only aggregated demand, not heterogenous firms” changes the demand structure of the standard GTAP such that only one Armington nest shared by all agents is present. The substitution elasticity is defined as in the case when the full Melitz model is used. The remaining equations of the Melitz model are absent and replaced by the equations of the standard GTAP model. That allows disentangling impacts of the full Melitz model from the structural changes on the demand side.

The checkbox “Fix cost nest” introduces a differentiation between variable and fix costs input composites where the latter only comprises primary factors, at least as long as there is some minimal primary factor cost share left in the variable input composite. The spinner termed “Max cost share of HET domestic interm. demand” allows switching off love of variety for domestic intermediate demand by same industry depending on the costs share. Only sector-regional combinations with a share above the threshold will be excluded. Setting the threshold to unity will hence leave the love of variety effect switch on for all domestic intermediate demand, while zero will switch it off on all cases.

Detail on simulated value of variables found in the Melitz module can be found in two tables under the “Trade group” as shown below:



## An example application

We use a 50% reduction of all imports tariff and export subsidies globally with a 10x10 aggregation of the GTAP9 data base as an example application<sup>114</sup>, focusing on welfare changes and highlighting the newly available information from the model. We compare the standard GTAP model against two variants, one where only the demand structure of the Melitz model is used and a second one with its full implementation.

<sup>114</sup> Appendix 2 provides the sectoral and regional aggregation of GTAP sectors into the new mapping.

Table 18 below reports the Money metric for the full Melitz model (Tariffs\_m), the model where only the demand side of the Melitz model is used (Tariffs\_d) and the Standard GTAP model (Tariffs). First, it can be seen that aggregate welfare is almost not affected by choosing the more simple demand side with only one Armington nest shared by all agents. However, results for individual regions show some sensitivity.

As expected, adding the full Melitz model increases welfare, here, it more than quadruples the global welfare impact of multi-lateral trade liberalization. Interestingly, that impact is not uniform across regions: whereas North America suffers a small welfare loss in the standard configuration, it gains under the Melitz model. However, the welfare in Middle East and North Africa is the same under two structures. Generally the results are in line with finding of Balistreri et al. (2011) which also found welfare increases around factor four.

Note that, for our simulation exercise, we compare the welfare impact of the policy shock under different structures where more or less the same value of Armington elasticity is assigned in each structure. Specifically, our calibration code restricts the substitution elasticity to an interval of  $\pm \frac{1}{2}$  around the share parameter which for many sectors will yield more or less the same consumption quantity weighted elasticity as found in the two-level Armington system of GTAP.

Dixon et al. (2016) argues that in order to compare the welfare impacts of a policy shock under Armington structure and Melitz structure one should assign the substitution elasticity ( $\sigma$ ) in the Armington framework such that it simulated trade flows comparable to a Melitz framework. In the two commodity and two country model of Dixon et al. (2016); a substitution elasticity of 8.45 yielded trade flows similar to a Melitz model with a substitution elasticity of 3.8 when tariff in one of the countries increases by 7.18 percent. Dixon et al. (2016) obtained this equivalent Armington value of substitution using a trial and error approach. The substitution elasticity in an Armington model which yields similar results depends not only on the parameterization of the Melitz model, but also on the type and magnitude of the shock and structure of the model as a whole. Using trial and error to an “equivalent” substitution elasticity to replicate the trade pattern after a given shock can only be used if the model is rather small. However, with a medium sized model, we found that once  $\sigma$  is adjusted in Armington framework to replicate one of the trade flows, the error in other bilateral trade flows might even become larger. Such finding is in line with that of Balistreri et al. (2011). Akgul (2015, 2016), who introduced a theoretically consistent framework from which one can calibrate the sector specific Armington elasticity given the obtained sector specific value of the shape parameter and econometrically estimated results. Given data limitations, we adjust the given Armington elasticities around a uniform Pareto shape parameter across regions and sectors. This assumption remains open to critic and the choices of theoretically consistent substitution elasticities across regions and sectors is

**Table 18:** Money metric in comparison of Melitz and Standard GTAP model

| Money metric decomposition [0] |           |           |             |
|--------------------------------|-----------|-----------|-------------|
| Sectors and institutions       |           | Total     | Years shock |
| Total [Bio \$]                 | Total     |           |             |
| Tariffs                        | Tariffs_d | Tariffs_m |             |
| World                          | 44.88     | 50.89     | 177.64      |
| Australia, New Zealand         | 1.63      | 1.77      | 4.23        |
| East Asia                      | 30.14     | 31.72     | 94.15       |
| Southeast Asia                 | 2.93      | 4.13      | 13.64       |
| South Asia                     | 3.08      | 1.34      | 10.19       |
| North America                  | -3.76     | -0.62     | 10.96       |
| Latin America                  | 0.80      | 0.23      | 5.01        |
| European Union 25              | 4.50      | 9.21      | 23.92       |
| Middle East and North Africa   | 0.53      | -0.02     | 0.53        |
| Sub-Saharan Africa             | 0.68      | 0.35      | 1.41        |
| Rest of World                  | 4.34      | 2.77      | 13.60       |

The next Table 19 shows the simulated overall price index and seems to indicate, as expected, that the Melitz model tends to amplify the impacts found under Armington model.

**Table 19:** Aggregate price index of the Armington agents

Income, compare across regions [0]

Total or household groups

Sectors and institutions

Years shock

Percentage diff. to Scenarios no Shock

Total

Price index [Index]

Tariffs

Tariffs\_d

Tariffs\_m

|                              |                |                |                |
|------------------------------|----------------|----------------|----------------|
| World                        | 0.99<br>-0.62% | 0.99<br>-0.89% | 0.99<br>-1.01% |
| Australia, New Zealand       | 1.00<br>0.10%  | 1.00<br>0.12%  | 1.00<br>0.02%  |
| East Asia                    | 1.01<br>0.68%  | 1.01<br>0.54%  | 1.01<br>0.61%  |
| Southeast Asia               | 1.00<br>-0.01% | 1.00<br>0.03%  | 1.00<br>0.18%  |
| South Asia                   | 0.99<br>-1.36% | 0.98<br>-2.17% | 0.98<br>-2.27% |
| North America                | 1.00<br>-0.48% | 1.00<br>-0.38% | 0.99<br>-0.67% |
| Latin America                | 0.99<br>-0.79% | 0.99<br>-0.86% | 0.99<br>-1.12% |
| European Union 25            | 1.00<br>-0.45% | 1.00<br>-0.28% | 0.99<br>-0.50% |
| Middle East and North Africa | 0.98<br>-1.80% | 0.98<br>-2.08% | 0.98<br>-2.35% |
| Sub-Saharan Africa           | 0.98<br>-1.74% | 0.98<br>-2.07% | 0.98<br>-2.41% |
| Rest of World                | 0.99<br>-0.93% | 0.99<br>-1.07% | 0.99<br>-1.29% |

The impacts on factor income corrected for changes in the overall price index are less clear with more diversified changes across the regions. In an environment with multiple factors, effects on factor prices differ which seems significant in our application.

**Table 20:** Factor income, corrected with price index aggregate Armington agent

| Income, compare across regions [0] |                   |                                                     |                   |                     |
|------------------------------------|-------------------|-----------------------------------------------------|-------------------|---------------------|
| Total or household groups          |                   | Sectors and institutions                            |                   | Percentage diff. to |
| Total                              |                   | Factor income, normalized with price index [Bio \$] | Years shock       | Scenarios no Shock  |
|                                    | Tariffs           | Tariffs_d                                           | Tariffs_m         |                     |
| World                              | 39756.41<br>0.69% | 39861.78<br>0.95%                                   | 39922.23<br>1.11% |                     |
| Australia, New Zealand             | 759.26<br>0.55%   | 759.35<br>0.56%                                     | 760.90<br>0.76%   |                     |
| East Asia                          | 7226.88<br>0.69%  | 7223.42<br>0.64%                                    | 7268.05<br>1.26%  |                     |
| Southeast Asia                     | 1081.11<br>1.23%  | 1081.91<br>1.30%                                    | 1089.57<br>2.02%  |                     |
| South Asia                         | 1242.61<br>1.52%  | 1239.76<br>1.29%                                    | 1242.17<br>1.48%  |                     |
| North America                      | 12335.02<br>0.10% | 12337.88<br>0.12%                                   | 12345.46<br>0.18% |                     |
| Latin America                      | 2051.05<br>0.63%  | 2050.54<br>0.60%                                    | 2054.21<br>0.78%  |                     |
| European Union 25                  | 10260.98<br>0.11% | 10266.71<br>0.17%                                   | 10276.61<br>0.27% |                     |
| Middle East and North Africa       | 448.36<br>1.67%   | 447.76<br>1.53%                                     | 447.79<br>1.54%   |                     |
| Sub-Saharan Africa                 | 682.86<br>1.39%   | 682.22<br>1.30%                                     | 682.14<br>1.29%   |                     |
| Rest of World                      | 3561.89<br>0.76%  | 3559.80<br>0.70%                                    | 3566.61<br>0.89%  |                     |

Table 21 below summarizes main variables found in the Melitz module. As indicated in the table, as expected, the total number of firms entered (number of domestic verities) decrease as result of liberalization while the number of operating firm indicating the sum of verities consumed increases.

**Table 21:** Example summary information from Melitz model

| Melitz, overview [0]           |                  |                     |             |                     |
|--------------------------------|------------------|---------------------|-------------|---------------------|
| Exporter                       |                  | Products            |             | Percentage diff. to |
| Australia, New Zealand         |                  | Heavy Manufacturing | Years shock | Scenarios no Shock  |
|                                | Tariffs_m        | no Shock            |             |                     |
| All firms entered              | 0.98<br>-2.12%   | 1.00                |             |                     |
| All firms operating            | 0.05<br>0.65%    | 0.05                |             |                     |
| Variable cost price            | 1.00<br>0.09%    | 1.00                |             |                     |
| Fix cost price                 | 1.00<br>0.09%    | 1.00                |             |                     |
| General fixed cost             | 27.51<br>-2.03%  | 28.08               |             |                     |
| Bi-lateral fixed cost          | 4.75<br>-0.57%   | 4.77                |             |                     |
| Variable cost                  | 126.57<br>-2.03% | 129.19              |             |                     |
| General fixed cost per unit    | 0.13<br>0.37%    | 0.13                |             |                     |
| Bi-lateral fixed cost per unit | 0.02<br>1.86%    | 0.02                |             |                     |
| Variable cost per unit         | 0.58<br>0.37%    | 0.58                |             |                     |

The last Table 22 below details information on each trade link. Note first the impact on the domestic sales: as the firm's price in the domestic market drops, the number of firms operating on the domestic link is reduced and average productivity and output per firm increases. In opposite to that, the firms prices of selling to "South Asia" increase by almost 1.3 % which increases the number of operating firms on that link by 1.12% and let average quantities and productivity drop by about -2.87%. Still, as a result of these effects, export increase by 8.7%.



**Table 22: Example information from Melitz model by trade link**

| Meltz, trade matrix [0]    |        |                        |                        |                     |                |            |               |               |                   |                              |                    | Percentage diff. to |  |
|----------------------------|--------|------------------------|------------------------|---------------------|----------------|------------|---------------|---------------|-------------------|------------------------------|--------------------|---------------------|--|
|                            |        | Exporter               |                        | Products            |                |            |               | Years         |                   | Scenarios                    |                    |                     |  |
|                            |        | Australia, New Zealand |                        | Heavy Manufacturing |                |            |               | shock         |                   | no Shock                     |                    |                     |  |
|                            |        | Domestic sales         | Australia, New Zealand | East Asia           | Southeast Asia | South Asia | North America | Latin America | European Union 25 | Middle East and North Africa | Sub-Saharan Africa | Rest of World       |  |
| Firm price                 | 1.00   | 1.00                   | 1.01                   | 1.01                | 1.03           | 0.99       | 1.01          | 1.00          | 0.99              | 0.98                         | 1.00               |                     |  |
|                            | -0.15% | -0.09%                 | 1.38%                  | 0.51%               | 3.05%          | -0.67%     | 1.30%         | -0.37%        | -0.58%            | -2.35%                       | -0.45%             |                     |  |
| Number of operating firms  | 0.97   | 0.97                   | 1.04                   | 1.00                | 1.12           | 0.95       | 1.03          | 0.96          | 0.95              | 0.87                         | 0.95               |                     |  |
|                            | -3.22% | -2.96%                 | 3.78%                  | -0.24%              | 11.92%         | -5.49%     | 3.41%         | -4.17%        | -5.12%            | -12.66%                      | -4.55%             |                     |  |
| Avg. output per firm       | 102.25 | 7.52                   | 18.54                  | 6.78                | 4.54           | 6.60       | 0.60          | 9.33          | 0.15              | 1.53                         | 4.20               |                     |  |
|                            | 0.25%  | 0.19%                  | -1.26%                 | -0.41%              | -2.87%         | 0.76%      | -1.19%        | 0.46%         | 0.68%             | 2.51%                        | 0.55%              |                     |  |
| Avg. productivity per firm | 0.34   | 0.34                   | 0.29                   | 0.34                | 0.29           | 0.30       | 0.30          | 0.35          | 0.35              | 0.35                         | 0.35               |                     |  |
|                            | 0.25%  | 0.19%                  | -1.26%                 | -0.41%              | -2.87%         | 0.76%      | -1.19%        | 0.46%         | 0.68%             | 2.51%                        | 0.55%              |                     |  |
| Total output sold          | 98.97  | 7.30                   | 19.24                  | 6.76                | 5.08           | 6.24       | 0.62          | 8.94          | 0.14              | 1.34                         | 4.01               |                     |  |
|                            | -2.98% | -2.77%                 | 2.47%                  | -0.65%              | 8.71%          | -4.77%     | 2.18%         | -3.73%        | -4.48%            | -10.47%                      | -4.03%             |                     |  |

### 3.15.7 Conclusion

While the Armington specification based on regionally differentiated goods provides a popular and robust specification for numerical simulations of trade policy, it fails to explain empirical observations at firm-level in newer international trade literature. Recent models of international trade with heterogeneous firms overcome the limitations of the Armington specification and can at the same time be relatively easy integrated into aggregated equilibrium analysis. That has opened up the opportunity for CGE models to better depict and analyze mechanisms through which productivity and number of varieties impact the extensive margin of trade. This paper discusses an operational implementation of the firm heterogeneity theory of Melitz (2003) into the CGEBox model which addresses the shortcoming of the Armington specification while being relatively simple. It has proven numerically stable at least for medium sized aggregations of the GTAP data base.

### Appendix 1: Price linkages

The following section presents the price linkages between firms and agents. Firm's price  $PF_{irs}$  (inclusive of domestic transport margin) is defined as the price received by producers in region  $r$  for commodity  $i$  to be shipped to the sink region. If the commodity is shipped to the domestic market, the agent (purchase) price is

$$PA_{airs} = PF_{irs} (1 + \tau_{air}^{cd} + \varsigma_{air}^{cd}) \quad (32)$$

where  $\tau_{air}^c$  denote the consumption (sale) tax on each specific commodity for each agent while  $\varsigma_{air}^c$  is a uniform consumption tax across commodities and/or agents.

If the commodity is shipped to the other region/country, a bilateral export subsidy or tax ( $\tau_{irs}^e$ ) is applied to the firm offer price and determines the free on board (fob) price. An additional tax  $\varsigma_{irs}^e$  is also introduced into (33) representing the uniform export tax across destinations.

$$PE_{irs}^{FOB} = PM_{irs} (1 + \tau_{irs}^e + \varsigma_{irs}^e) \quad (33)$$

The FOB price  $PE_{irs}^{FOB}$  is augmented by the international transport margin  $t_{irs}^{tmg}$  (observed form GTAP and endogenous to the model) to establish the cif (cost –insurance and freight) price:

$$PM_{irs}^{CIF} = PE_{irs}^{FOB} + t_{irs}^{tmg} \quad (34)$$

The bilateral import tax ( $\tau_{irs}^m$ ) converts the cif price into the bilateral import price, and  $\varsigma_{irs}^m$  reflect a uniform tax shift across source countries

$$PM_{irs} = P_{irs}^{CIF} (1 + \tau_{irs}^m + \varsigma_{irs}^m) \quad (35)$$

and finally the resulting bilateral import prices are converted to the agent prices by adding a consumption tax on imported commodities:

$$PA_{airs} = P_{irs}^{CIF} (1 + \tau_{air}^{cim} + \varsigma_{air}^{cim}) \quad (36)$$

While uniform shift parameters are initially set to zero, the value of all other parameters are observed from GTAP 8 database. Potential users are strongly urged to consult Hertel (1997) and Mensbrugghe (2015).

### 3.15.8 Acknowledgments

Authors would like to thank Thomas W. Hertel, Edward J. Balistreri, and Dominique van der Mensbrugghe for their helpful comments.

### 3.16 GTAP-AEZ

#### 3.16.1 Summary

The GTAP-AEZ module allows to break land-use in a region to different Agro-Ecological Zones (AEZ) using a multi-tier land-supply approach. The necessary data are available from GTAP and covered by a GTAP data base license and are added during the data preparation step. The GUI gives limited methodological choice. The exploitation part allows generating global maps which depict land use (changes).

The module draws on Huey-Lin, Incorporating Agro-Ecologically Zoned Land Use Data and Land-based Greenhouse Gases Emissions into the GTAP Framework (<https://www.gtap.agecon.purdue.edu/resources/download/2245.pdf>) and uses the following logic. Each region (nation, group of nation or NUTS2) in the model is split into Agro-Ecological Zones (AEZs, set *aez*). The total land use of an activity at regional level is a CES-aggregate of the land use in the different AEZs, i.e. all other costs shares do not differ across AEZs. In each AEZ, individual land use activities compete for available total managed land which is not fully mobile among individual activities, described by a four-tier or six-tier CET factor supply system.

The relation between the regions and the AEZs depicted by the cross-set *aezFlag*:

```
set aezFlag(r,aez);
```

In each AEZ, the total endogenous land use is dis-aggregated to the following three top-level land categories and an agricultural aggregate *graz\_agr*:

```
set landCat(aezNest) "Land categories" / forest,PastureLand,CropLand,Graz_Agr /;
```

The different activities in the GTAP Data Base are mapped to these to land categories, i.e. forestry comprises managed forest (*frs*), the grazing categories comprises all land used by animal production activities while agriculture covers all crops.

```
set landCat_a0(landCat,i0) "Map between land categories and activities"/
    forest.frs,
    PastureLand.(ctl,oap,rmk,wol)
    CropLand.(set.icrops)
/;
```

The code does not require that the SAM comprises the full detail of activities as defined above, but assigns potentially higher aggregated activities to the land categories based on the mapping defined above. However, it will throw an error if the sectoral aggregation leads to the case where one activity in the SAM would be assigned to multiple land categories as, for instance, in the case where crop and livestock activities in the GTAP data base are aggregated.

```
set landCat_a(landCat,a);
landCat_a(landCat,a) $ sum(mapa(i0,a) $ landCat_a0(landCat,i0),1) = YES;

abort $ (sum(landCat_a(landCat,a),1) ne sum(a $ sum(landCat_a(landCat,a),1),1))
    "Not enough detail for agriculture / forestry, several sector assigned to same land category",
    landCat_a;
```

#### 3.16.2 Model equations

The model equations are defined in “*gams\GtapAEZ\gatpAez\_model.gms*”. The implementation differs in five important aspects from the AEZ model as coded in GEMPACK:

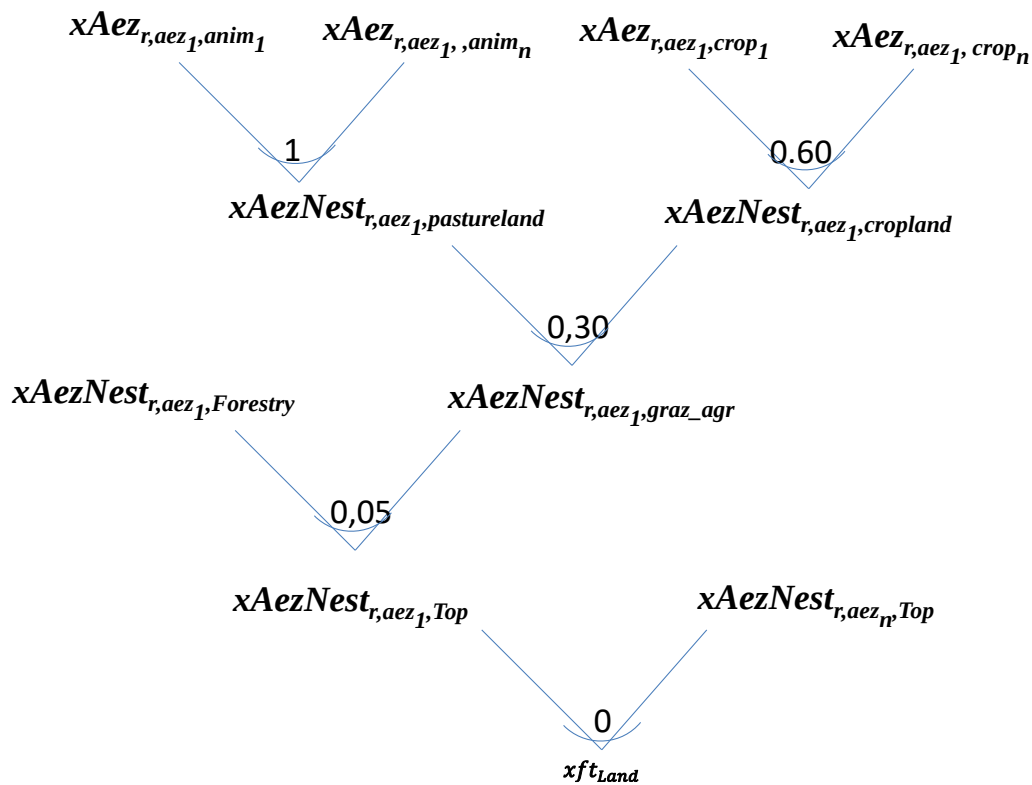
1. It comprises on demand a four or six instead of three-tier CET distribution of total managed land: the upper nest distributes total managed land to forestry and all agricultural uses, the middle nest from all agricultural uses to pasture, i.e. land used in animal, and cropland, i.e. crop activities, and the lowest tier distributes pasture to the different animal activities to cropland to the individual crop activities. The middle tier is not present in the original implementation. The introduction of the additional nest allows making managed forest cover more “sticky” compared to switching between aggregate animal and crop use.
2. Additionally, the crop nests can be broken down to two nests with typical arable crops and horticulture in the other.
3. The two upper tiers use a volume preserving form of the CET (van der Mensbrugghe, D., & Peters, J. C. (2016, April). Volume preserving CES and CET formulations. In 2016 GTAP Conference Paper. GTAP Resource (Vol. 5070)).
4. Land supply elasticities relate to the benchmark total land stock.
5. If NUTS2 regions for European countries as sub-national regions are present, they are also on demand broken down to individual AEZs.
6. It is possible to aggregate all AEZ into one, either for all regions or for sub-national regions, only.

It is hence recommended switching the following options on via GUI controls. Switching the land supply off gives a model closer to the GEMPACK implementation:

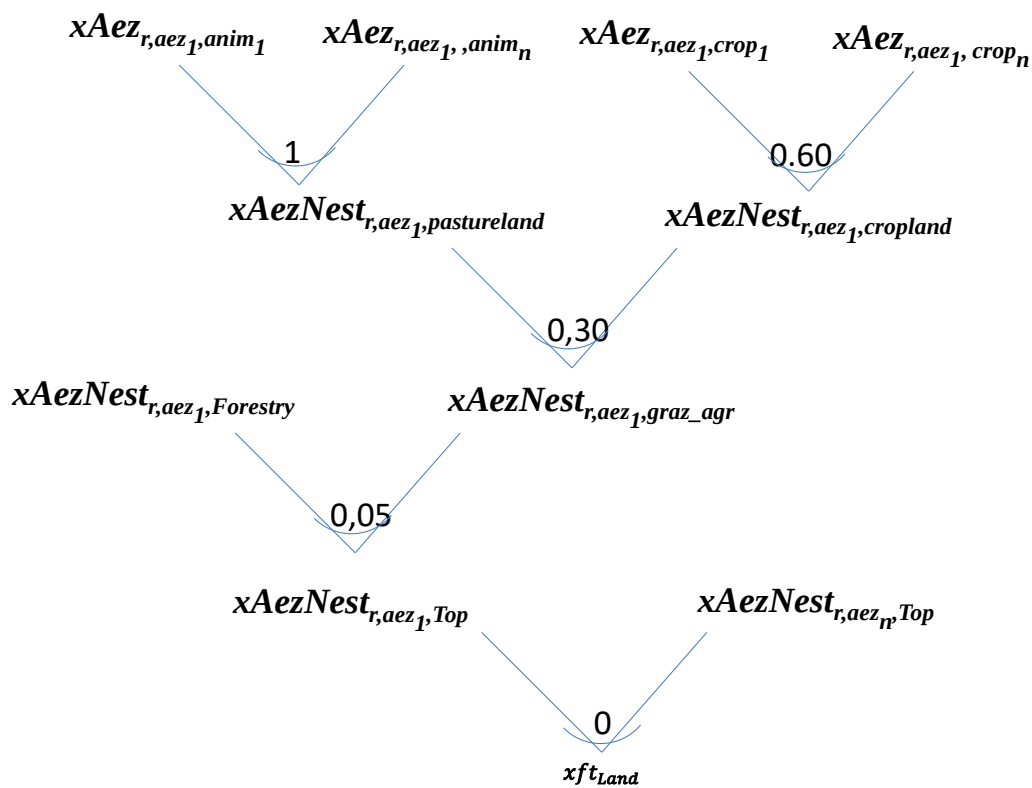
The screenshot shows a GUI control panel with the following settings:

- Land supply elasticities drive land buffer: ☒
- Aggregate all AEZ into one: ☐
- Aggregate all AEZ into one for sub-regions, only: ☒
- Nesting structure: 4 Nests (Top - Graz\_Agr - Agriculture - Grazing) (dropdown menu)

The figure below gives an overview on the CET based transformation of the total land stock in a region to the land stock in the different AEZs and to the four layers.



The version with the six layers looks as follows:



### Total economically used land stock

The model can be used in two variants to depict the total land stock in economic use in each AEZ:

1. A factor stock equation with a supply elasticity *etafAez*. The land stock at the benchmark *xAezNestBase* is updated with the factor supply elasticity *etafAez* applied to the returns to land *pAezNest*. As this price is not equal to unity at the benchmark, it is normalized with the benchmark price. Larger price changes and/or higher elasticities could provoke a solution where all available unmanaged land would be exhausted such that unmanaged land could become negative. To avoid this, a fugged min operator based on a *ncpVuSin* function is used. It restricts the expansion. The terms will be motivated when discussing the equation defining the stock of unmanaged land by category in the next equation. The *e\_xftTopStd* equation is removed in baseline construction with G-RDEM if FAO projections for cropland are used.

```

--- land supply to agriculture from stock with security bounds to keep unmanaged land positive

e_xftTopStd(aezFlag(rs,aez),ts(tRun)) $ (p_xAezNestBase(rs,aez,"top",tRun) $ (v_topLandChange.range(rs,aez,tRun) ne 0)
$ (pdAezNest.range(rs,aez,"top",tRun) ne 0) $ hasGtapAez) ..

--- total managed land relative to benchmark: (change + benchmark)/benchmark, change is used in e_unManagedLnd below
v_topLandChange(rs,aez,tRun)/p_xAezNestBase(rs,aez,"top", "%t0%")

=E=

--- smoothed min of the two below expressions:
ncpVUSin(
  --- (1) simulated relative change in managed land at benchmark according to given land supply elasticity
  ((pAezNest(rs,aez,"top",tRun)/(pwfact(tRun)*pAezNest.l(rs,aez,"top", "%t0%")))**p_xAezNestBase(rs,aez)-1)
  * p_xAezNestBase(rs,aez,"top", "%t0%") + p_xftTopCorr(rs,aez,tRun),

  --- (2) maximum possible change (can be negative) defined from remaining unmanaged land and the exogenous change in
  built-up land, relative to total managed land at benchmark
  smin(unManaged $ (unManagedLnd.l(rs,aez,unManaged, "%t0%") $ (p_bufferShare(rs,aez,unManaged) gt eps)),
  unManagedLnd.l(rs,aez,unManaged, "%t0%") * 0.90
  / [ (p_landUse(rs,aez,unManaged,"ha")*p_bufferShare(rs,aez,unManaged))
  /sum(unManagedl,p_landUse(rs,aez,unManagedl,"ha")*p_bufferShare(rs,aez,unManagedl))]*1.0
  - min(xAezNest.l(rs,aez,"top",tRun-1)*0.10,p_landUse(rs,aez,"builtUpLand","ha") * p_builtupLandChange(rs,ts) * 1.E-6)

  --- fugging parameter of ncpVUSin function
  , max(1.E-7,p_xAezNestBase(rs,aez,"top", "%t0%")*1.E-7)
)/p_xAezNestBase(rs,aez,"top", "%t0%");

```

2. The amount of unmanaged land by category *unmanaged* is driven by the absolute changes in the managed land *v\_topLandChange* defined above, plus exogenous relative changes in built up land in *p\_builtupLandChange*. The total change in managed land is distributed to the different categories according to the parameter *p\_bufferShare* and the stock at the benchmark. If, for instance, all managed land would have the same buffer share, an additional hectare of managed land would be distributed according to the share of each unmanaged land category on total unmanaged land. Assume that one of these shares would be 20% and the benchmark stock of this category would be 1,000 ha. It would hence be possible to expand the managed stock by 1,000 ha / 0.2 = 5,000 ha without violation the non-negativity condition. This calculation motivates the *smin* operator in the *xftTopStd* equation.

```

--- total stock of unmanaged land by type (unManagedForest, savannas ...)

e_unManagedLnd(aezFlag(rs,aez),unManaged,ts(tRun)) $ (
  (p_xAezNestBase(rs,aez,"top",ts) gt eps) $ (p_bufferShare(rs,aez,unManaged) gt eps)
  $ (p_landUse(rs,aez,unManaged,"ha") gt eps) $ hasGtapAez) ..

unManagedLnd(rs,aez,unManaged,ts) s (unManagedLnd, 'rs,aez,unManaged,ts')
=E=

(p_landUse(rs,aez,unManaged,"ha") * 1.E-6

--- exogenously determined changes in builtUpLand in ha (+ = Increase in built-up land in relative terms)
- [ min(xAezNest.l(rs,aez,"top",tRun-1)*0.10,p_landUse(rs,aez,"builtUpLand","ha") * p_builtupLandChange(rs,ts) * 1.E-6)

--- changes in total managed land in 1000ha (either from equation e_xftTopStd, or from G-RDEM, or fixed from user)
+ v_topLandChange(rs,aez,ts)
]

--- distributed according to shares on total buffer of unmanaged land (= different types such as nat.forest, savannah ...)
*[
  (p_landUse(rs,aez,unManaged,"ha")*p_bufferShare(rs,aez,unManaged))
  /sum(unManagedl,p_landUse(rs,aez,unManagedl,"ha")*p_bufferShare(rs,aez,unManagedl))]
) s (unManagedLnd, 'rs,aez,unManaged,ts')
;

```

### Distribution of total land demand from regional to AEZ level

The land demand  $x_{Aez}$  for each activity  $a$  in each Agro-Ecological Zone  $aez$  is defined in the equation  $e\_pAez$  based on a CES function. Factor returns at market level  $pf$  are used as agent specific ad-valorem tax rates on land which determine the differences between returns at market prices  $pf$  and at agent prices  $pfa$  are not differentiated by AEZ:

```

--- Land Demand for each activity in each aez, drives distribution in space
    (relation between average firm's price for land for the activity
    and the firm's price at AEZ level drives the distribution)

e_pAez(rs(r),aez,aIn(aezA),ts(t)) $(aezFlag(r,aez) $ p_andAez(r,aez,aezA,t) $ hasGtapAez) ..
x_Aez(r,aez,aezA,t)s(x_Aez,'r,aez,aezA,t')
=e= p_andAez(r,aez,aezA,t)
    * xf(r,lnd,aezA,t)s(x_Aez,'r,aez,aezA,t')
    * (1 $(not p_sigmaAez(r,aezA)) + ((pf(r,lnd,aezA,t)/pAez(r,aez,aezA,t))**p_sigmaAez(r,aezA)) $ p_sigmaAez(r,aezA));

```

This equation can be understood as a cost minimization problem to allocate crops such as to minimize the land rents to pay, taking however differences in land qualities into account. The total to distribute is equal to the total land use of that activity at regional level  $xf$ .

The average return  $pf$  uses the usual dual price aggregator defined in the equation  $e\_pfaAez$ . Note again that ad-valorem factor taxes are not present as they appear as a multiplicative term on both sides of the equation. The equation considers the simpler case of a Leontief relation with fixed allocation shares separately.

```

--- average factor price per activity aggregated over AEZs in each region

e_pfAez(rs(r),lnd(f),aIn(aezA),ts(t)) $(sum(aezFlag(r,aez), p_andAez(r,aez,aezA,t)) $(pf.range(r,lnd,aezA,t) ne 0) $ hasGtapAez) ..
pf(r,lnd,aezA,t)
=e= [sum(aezFlag(r,aez)
    %scale% $ p_andAez(r,aez,aezA,t)
    ,p_andAez(r,aez,aezA,t) * pAez(r,aez,aezA,t)**(1-p_sigmaAez(r,aezA))**(1/(1-p_sigmaAez(r,aezA))))] $ p_sigmaAez(r,aezA)
+ sum(aezFlag(r,aez)
    %scale% $ p_andAez(r,aez,aezA,t)
    ,p_andAez(r,aez,aezA,t) * pAez(r,aez,aezA,t)) $(not p_sigmaAez(r,aezA));

```

The volume preserving formulation used for the top-level land use nests maximizes the revenue of a CET aggregator to distribute a physical land stock  $V$  among competing uses  $X$ :

$$\max R = A \left[ \sum_i g_i (\lambda_i P_i X_i)^\nu \right]^{1/\nu} \quad (3)$$

Where  $g_i$  are share parameters,  $P_i$  are given prices,  $A$  is a scaling factor and  $X_i$  are the land use quantities to be optimized and  $\nu$  refers to the transformation elasticity. The A-CET introduces the condition that the sum of  $X$  is fixed to the given volume  $V$ , in this case, a land stock measured in physical units:

$$V = \sum_i X_i \quad (4)$$

With a given transformation elasticity  $\varpi$ , this generates the following solutions for the competing land uses  $X_i$  and the dual price aggregator  $P^c$

$$X_i = \gamma_i (A\lambda_i)^\varpi \left( \frac{P_i}{P^c} \right)^\varpi \quad (5)$$

$$P^c = A \left[ \sum_i \gamma_i (\lambda_i P_i)^\varpi \right]^{1/\varpi} \quad (6)$$

The advantage of using the modified A-CET formulation is twofold. Firstly, the parameters provided by the GTAP-AEZ model – or by alternative models based on the standard CET – can still be used for benchmarking. Secondly, structural changes to the equations in the GTAP-AEZ model are relatively minor. Specifically, equation (5) is indeed identical to the original CET formulation while equation (6) only changes the exponent. However, the price index in (6) is no longer equal to the average revenue exhaustion price  $P^e$  which has to be defined in a separate equation:

$$P^e = \sum X_i P_i / V \quad (7)$$

Accordingly, the dual price aggregator from individual land use activities and land use sub-nest to a land use nest are defined as follows in equation `e_pDAezNest`, where the differences between the volume preserving and standard CET are solely in the exponents. The equation considers three cases: (1) the finite case with a non-zero transformation elasticity in the first block, (2) a fixed allocation with a transformation elasticity of zero, and (3) infinite transformation, implying a linear aggregation.

```

---- Dual price aggregator for CET, either default or for volume preserving)

e_pDAezNest(aezFlag(rs,aez),aezNest,ts) $ ( ( sum(aezNest_a(aezNest,aezA), p_gAez(rs,aez,aezA,ts))
                                         or sum(aezNest_n(aezNest,aezNest1), p_gAez(rs,aez,aezNest1,ts)))
                                         $ (xNest.range(rs,aez,aezNest,ts) ne 0) $ hasGtapAez ) ..

0 =e=
(

    --- sluggish land mobility in aez nest, dual price aggregator for nests
    (CET is not infinite)

    pDAezNest(rs,aez,aezNest,ts)s(pDAezNest,'rs,aez,aezNest,ts')

- [[
    --- activities linked into nest (note the missing 1 if a volume preserving CET is used)
    (this is for case where p_omegaAez <> 0

    sum(aezNest_a(aezNest,aezA)
      %scale% $ p_gAez(rs,aez,aezA,ts)
    ,p_gAez(rs,aez,aezA,ts) * pAez(rs,aez,aezA,ts)**(1 $ aezTyp(aezNest,"CET") + p_omegaAez(rs,aez,aezNest)))

    --- subnests linked into sub nest

+ sum(aezNest_n(aezNest,aezNest1)
      %scale% $ p_gAez(rs,aez,aezNest1,ts)
    ,p_gAez(rs,aez,aezNest1,ts)*(1+v_gAez(rs,aez,aezNest1,ts))
      * (pAezNest(rs,aez,aezNest1,ts)**(1 $ aezTyp(aezNest,"CET") + p_omegaAez(rs,aez,aezNest)))
    ]*(1/(1 $ aezTyp(aezNest,"CET") + p_omegaAez(rs,aez,aezNest)))s(pDAezNest,'rs,aez,aezNest,ts')) $ p_omegaAez(rs,aez,aezNest)

```



```

- {
    --- activities linked into nest, p_omegaAez == 0

    sum(aezNest_a(aezNest,aezA)
        %scale% $ p_gAez(rs,aez,aezA,ts)
    ,p_gAez(rs,aez,aezA,ts) * pAez(rs,aez,aezA,ts))

    --- subnests linked into sub nest, p_omegaAez = 0

    + sum(aezNest_n(aezNest,aezNest1)
        %scale% $ p_gAez(rs,aez,aezNest1,ts)
    ,p_gAez(rs,aez,aezNest1,ts) * (1+v_gAez(rs,Aez,aezNest1,ts)) * pAezNest(rs,aez,aezNest1,ts))
    ]s(pDAezNest,'rs,aez,aezNest,ts') } $ (not p_omegaAez(rs,aez,aezNest))

) $ (p_omegaAez(rs,aez,aezNest) ne inf)

--- fully mobile land across uses in land supply nest, uniform prices => physical aggregation
+ ( xAezNest(rs,aez,aezNest,ts)s(xAezNest,'rs,aez,aezNest,ts')

    --- the share parameter act here as unit converting factors, from ha to const. USD

- sum(aezNest_a(aezNest,aezA)
    %scale% $ p_gAez(rs,aez,aezA,ts)
    ,p_gAez(rs,aez,aezA,ts)*xAez(rs,aez,aezA,ts))s(xAezNest,'rs,aez,aezNest,ts')
- sum(aezNest_n(aezNest,aezNest1)
    %scale% $ p_gAez(rs,aez,aezNest1,ts)
    ,p_gAez(rs,aez,aezNest1,ts)*(1+v_gAez(rs,Aez,aezNest1,ts))*xAezNest(rs,aez,aezNest1,ts))s(xAezNest,'rs,aez,aezNest,ts')

) $ (p_omegaAez(rs,aez,aezNest) eq inf)
;

```

The CET dual price index  $pDAezNest$  is defined by the following equation. For the case of standard CET (including the Leontief and infinite transformation cases), this is a just copy.

```

--- Average land rent in aezNest (exhaustion), needed for volume preserving nests
e_pAezNest(aezFlag(rs,aez),aezNest,ts(t)) $ (( sum(aezNest_a(aezNest,aezA), p_gAez(rs,aez,aezA,ts))
or sum(aezNest_n(aezNest,aezNest1), p_gAez(rs,aez,aezNest1,ts))) $ hasGtapAez ) ..

--- for the default CET formulation, this is a just a copy of the dual aggregator
pAezNest(rs,aez,aezNest,ts)s(pAezNest,'rs,aez,aezNest,t')

=e=

--- for the default CET formulation, this is a just a copy of the dual aggregator
+ (pDAezNest(rs,aez,aezNest,ts)s(pAezNest,'rs,aez,aezNest,t') )
$ (aezTyp(aezNest,"CET") or (p_omegaAez(rs,aez,aezNest) eq inf) or (p_omegaAez(rs,aez,aezNest) eq 0))

--- volume preserving: dual of x*p at xTot==1
+ [
    sum(aezNest_n(AezNest,aezNest1) $ p_gAez(rs,aez,aezNest1,ts),
    (p_gAez(rs,aez,aezNest1,ts)*(1+v_gAez(rs,Aez,aezNest1,ts))s(pAezNest,'rs,aez,aezNest,t')
    * pAezNest(rs,aez,aezNest1,ts)*(1+p_omegaAez(rs,aez,AezNest))))
    /pDAezNest(rs,aez,aezNest,ts))*p_omegaAez(rs,aez,AezNest)
] $ (aezTyp(aezNest,"volPreserv") $ ((p_omegaAez(rs,aez,aezNest) ne inf) and (p_omegaAez(rs,aez,aezNest) ne 0)));

```

The land supply from a land use nest  $xAezNest$  to the individual land use activities  $xAez$  is defined by the equation  $e\_xAez$ . No difference between the two CET formulations is found. As in the equations above, the case with fully mobile allocation needs to be treated differently:

```

--- land supply to activity from land supply nest
e_xAez(aezFlag(rs,aez),ain(aezA),ts) $ (p_andAez(rs,aez,aezA,ts) $ (pAez.range(rs,aez,aezA,ts) ne 0) $ hasGtapAez) ..

0 =e= [xAez(rs,aez,aezA,ts)s(xAez,'rs,aez,aezA,ts')
- sum(aezNest_a(aezNest,aezA),
    (p_gAez(rs,aez,aezA,ts)*xAezNest(rs,aez,aezNest,ts))s(xAez,'rs,aez,aezA,ts')
    * (pAez(rs,aez,aezA,ts)/pDAezNest(rs,aez,aezNest,ts))*p_omegaAez(rs,aez,aezNest)) ]
    $ (sum(aezNest_a(aezNest,aezA), p_omegaAez(rs,aez,aezNest)) ne inf)

+ [pAez(rs,aez,aezA,ts)/p_gAez(rs,aez,aezA,ts)s(pAez,'rs,aez,aezA,ts')
- sum(aezNest_a(aezNest,aezA), pDAezNest(rs,aez,aezNest,ts))s(pAez,'rs,aez,aezA,ts') ]
    $ (sum(aezNest_a(aezNest,aezA), p_omegaAez(rs,aez,aezNest)) eq inf);

```

The same relation holds for the allocation of sub-nests:

```

--- land supply to sub-nest from higher level land-supply nest
e_xAezNest(aezFlag(rs,aez),aezNest,ts) $ (sum(aezNest_n(aezNest1,aezNest), p_gAez(rs,aez,aezNest,ts))
$ (pdAezNest.range(rs,aez,aezNest,ts) ne 0) $ hasGtapAez ) ..

0 =e=

--- "sluggish" factor supply from another nest AezNest1 to the current nest
+ sum(aezNest_n(AezNest1,aezNest) $ p_gAez(rs,aez,aezNest,ts),
xazNest(rs,aez,aezNest,ts)s(xAezNest,'rs,aez,aezNest,ts')
- (p_gAez(rs,aez,aezNest,ts)*(1+v_gAez(rs,aez,aezNest,ts))* xazNest(rs,aez,aezNest1,ts))s(xAezNest,'rs,aez,aezNest,ts')
* (pAezNest(rs,aez,aezNest,ts)/pDAezNest(rs,aez,aezNest1,ts))*p_omegaAez(rs,aez,AezNest1))
$ ( sum(aezNest_n(aezNest1,aezNest) $ (p_omegaAez(rs,aez,AezNest1) ne inf),1))

--- perfect mobility: prices are equal
+ sum(aezNest_n(aezNest1,aezNest) $ (p_omegaAez(rs,aez,aezNest1) eq inf),
pAezNest(rs,aez,aezNest,ts)
- pAezNest(rs,aez,aezNest1,ts))s(pAezNest,'rs,aez,aezNest,ts')
;

```

## Link into global model

As land is immobile across the AEZ, economy-wide total factor use of land  $xft$  and the related price  $pft$  are defined by the following three equations. The first equation  $e\_xftLand$  sums the total managed land  $xftAez_{top}$  in each  $aez$  to the national or sub-national total  $xft$ . The second equation  $e\_vftLand$  sums up total land rents and the third one  $e\_pftLand$  defines the average factor price for land,  $pft$ :

```

--- aggregate land use over aez and calculate average price,
    feeds into regional income / tax equations in standard GTAP model
e_xftLand(rs,lnd,ts) $ ( (Xft.range(rs,lnd,ts) ne 0) $ sum(aezFlag(rs,aez),1) $ xftFlag(rs,lnd) $ hasGtapAez ) ..
xft(rs,lnd,ts)s(xft,'rs,lnd,ts')
=E= sum(aezFlag(rs,aez) $ xAezNest.l(rs,aez,"top","%t0%"), xAezNest(rs,aez,"top",ts))
/sum(aezFlag(rs,aez), p_xAezNestBase(rs,aez,"top","%t0%")) * p_xftBase(rs,lnd)s(xft,'rs,lnd,ts');

e_vftLand(rs,ts) $ (xftFlag(rs,lnd) $ sum(aezFlag(rs,aez),1) $ hasGtapAez ) ..
vftLand(rs,ts)
=E= sum(aezFlag(rs,aez) $ xAezNest.l(rs,aez,"top","%t0%"),
xAezNest(rs,aez,"top",ts)*pAezNest(rs,aez,"top",ts))s(vftLand,'rs,ts');

e_pftLand(rs,lnd,ts) $ ((pft.range(rs,lnd,ts) ne 0) $ sum(aezFlag(rs,aez),1) $ xftFlag(rs,lnd) $ hasGtapAez ) ..
pft(rs,lnd,ts) * xft(rs,lnd,ts)s(xft,'rs,lnd,ts')
=E= vftLand(rs,ts)s(xft,'rs,lnd,ts');

```

## Carbon stock accounting

The carbon stock accounting equation  $e\_dcarbon$  (shown in full below) comprises in a first block changes between the three economic use aggregates forest, cropland and pastureland for which carbon stock data are available. It multiplies the absolute change in each aggregate in the current simulation year (or against the benchmark in a comparative static run), normalized with the stock in the benchmark to define a relative change, with the stock in ha stored on  $p\_landUse$  times the carbon content  $p\_carbon$  in metric tons per ha, converted to millions metric tons, and multiplied with the overall scaler used in the model:

```

--- change in ha managed top categories to which carbon stocks are attached
sum(r_r(rs,rNat),aez,aezNest) $ ( (p_xAezNestBase(rs,aez,aezNest,"%t0%") gt eps) $ p_carbon(rNat,aez,aezNest,"carbon")),
(xAezNest(rs,aez,aezNest,tRun) - xAezNest.l(rs,aez,aezNest,tRun-1))
/p_xAezNestBase(rs,aez,aezNest,"%t0%")

--- map into hectare changes times carbon stock per ha
* p_landUse(rs,aez,aezNest,"ha")*p_carbon(rNat,aez,aezNest,"carbon") * 1.E-3 * gblScale

```

The second block is only used if land supply elasticities drive the different categories of unmanaged land reflects equation  $e\_unManagedLnd$  discussed above. We assign carbon stock  $p\_carbon$  to the changes in the unmanaged land extent  $unManagedLnd$  against last year's result or the benchmark:

```

$$iftheni.landSupply "%landSupplyEla%"=="on"
+ sum((r_r(rs,rNat),aez,unManaged) $ ( (p_landUse(rs,aez,unManaged,"ha") gt eps)
$ (p_xAezNestBase(rs,aez,"top",ts) gt eps)
$ (p_bufferShare(rs,aez,unManaged) gt eps) $ aezFlag(rs,aez)),

--- price induce released of un-managed land categories in ha, times carbon stock per ha

+ (unManagedLnd(rs,aez,unManaged,tRun) -unManagedLnd.l(rs,aez,unManaged,tRun-1))
* p_carbon(rNat,aez,unManaged,"carbon")) * 1000 * gblScale
$$endif.landSupply

```

Combined, the equation looks as follows:

```

dCarbon(rNat,tRun)
=scale% /sum((r_r(rs,rNat),aez,land_Cover) $ aezFlag(rs,aez),p_carbon(rNat,aez,land_cover,"carbon"))
=E= [

--- change in ha managed top categories to wich carbon stocks are attached

sum((r_r(rs,rNat),aez,aezNest) $ ( (p_xAezNestBase(rs,aez,aezNest,"%t0%") gt eps) $ p_carbon(rNat,aez,aezNest,"carbon")),
(xAezNest(rs,aez,aezNest,tRun) - xAezNest.l(rs,aez,aezNest,tRun-1))
/p_xAezNestBase(rs,aez,aezNest,"%t0%")

--- map into hectare changes times carbon stock per ha

* p_landUse(rs,aez,aezNest,"ha")*p_carbon(rNat,aez,aezNest,"carbon")) * 1.E-3 * gblScale

$$iftheni.landSupply "%landSupplyEla%"=="on"
+ sum((r_r(rs,rNat),aez,unManaged) $ ( (p_landUse(rs,aez,unManaged,"ha") gt eps)
$ (p_xAezNestBase(rs,aez,"top",ts) gt eps)
$ (p_bufferShare(rs,aez,unManaged) gt eps) $ aezFlag(rs,aez)),

--- price induce released of un-managed land categories in ha, times carbon stock per ha

+ (unManagedLnd(rs,aez,unManaged,tRun) -unManagedLnd.l(rs,aez,unManaged,tRun-1))
* p_carbon(rNat,aez,unManaged,"carbon")) * 1000 * gblScale
$$endif.landSupply

]
$$ifi defined p_timeStep / p_timeStep(tRun)

%scale% /sum((r_r(rs,rNat),aez,land_cover) $ aezFlag(rs,aez),p_carbon(rNat,aez,land_cover,"carbon"))
;

```

The equation also calculates the average per year in case that a recursive-dynamic simulation does not use annual time steps.

We add an equation *e\_dCarbonGblCo2* which defines the global change in CO2 equivalents:

```

e_dCarbonGblCo2(ts) $ ((card(rsNat) gt 0) $ (not sameas(ts,"%t0%")) $ hasGtapAez) ..

dCarbonGblCo2(ts) s(dCarbonGblCo2,'ts')
=E= sum(rsNat $ (sum((r_r(rs,rNat),aez,land_Cover)
$ aezFlag(rs,aez),p_carbon(rsNat,aez,land_cover,"carbon"))),
dCarbon(rsNat,ts)) s(dCarbonGblCo2,'ts') * 3.67;

```

An additional equation *e\_dCarbonBuiltUp* calculate the carbon stock losses resulting from changes in built-up land. These are comprised in the *dCarbon* variable.

```

e_dCarbonBuiltUp(rsNat(rNat),ts(tRun)) $ ( (not sameas(tRun,"%t0%")) $ sum((r_r(rs,rNat),aez,land_Cover)
$ aezFlag(rs,aez),p_carbon(rNat,aez,land_cover,"carbon")) $ hasGtapAez) ..

dCarbonBuiltUp(rNat,tRun)
=scale% /sum((r_r(rs,rNat),aez,land_Cover) $ aezFlag(rs,aez),p_carbon(rNat,aez,land_cover,"carbon"))

--- carbon stock changes from built-up land

=E= sum((r_r(rs,rNat),aez,unManaged) $ (
%scale% (p_landUse(rs,aez,unManaged,"ha") gt eps)
%scale% $ (p_xAezNestBase(rs,aez,"top",ts) gt eps)
%scale% $ (p_bufferShare(rs,aez,unManaged) gt eps) $
aezFlag(rs,aez)),

- min(xAezNest.l(rs,aez,"top",tRun-1)*0.10,p_landUse(rs,aez,"builtUpLand","ha") * (p_builtUpLandChange(rs,ts)-p_builtUpLandChange(rs,tRun-1)) * 1.E-6)
/sum(unManaged1,p_landUse(rs,aez,unManaged1,"ha")*p_bufferShare(rs,aez,unManaged1))*p_carbon(rNat,aez,unManaged,"carbon")) * 1000 * gblScale
$$ifi defined p_timeStep / p_timeStep(tRun)
%scale% /sum((r_r(rs,rNat),aez,land_Cover) $ aezFlag(rs,aez),p_carbon(rNat,aez,land_cover,"carbon"))
;

```

## MCP parings

The resulting pairings for all model equations in MCP format are shown below:

```

model GTAP_AEZ
/
  $$iftheni.landSupply "%landSupplyElas%"=="on"
    pftTopPriceEffect1Eq.pftTopPriceEffect1
    pftTopPriceEffectEq.pftTopPriceEffect
    unManagedLndeq.unManagedLnd
  $$endifi.landSupply

  dCarbonEq.dCarbon
  dCarbonGblCo2eq.dCarbonGblCo2

  xftTopEq.pdAezNest
  vAezNestEq.vAezNest
  xAezNestEq.pdAezNest
  pAezNestEq.pAezNest
  pAezNestEq.xAezNest

  xAezEq.pAez
  pAezEq.xAez
  pAezEq.pf

  pftLandEq.pft
  xftLandEq.xft
  vftLandEq.vftLand
/;

```

### 3.16.3 Parameterization

The necessary data for the AEZ module must be loaded during the data preparation step. The *aezFlags* are set depending on finding any land use activities in the AEZ data base (see *gams/gtapAez/gtapAez\_cal.gms*). The data driver ensures both consistency in physical terms and each land using category in the SAM shows some physical land use, with rents that exhaust the total land rents reported in the SAM. Accordingly, the code takes over the data for land use activity by AEZ from the data base

```

*
* --- define land use at AEZ level in Mio ha
*
xAez.l(aezFlag(r,aez),aezA,t0) = p_landUse(r,aez,aezA,"ha") * 1.E-6;
*
* --- price are equal to the reported rents
*
pAez.l(aezFlag(r,aez),aezA,t0) = p_landUse(r,aez,aezA,"rent");

```

And aggregate from there to the nests as defined in the code of the model equations:

```

*
* --- aggregate quantities and prices over single activities to nests
*
set aezNestToDo(aezNest),curAezNestToDo(aezNest);
alias(aezNestToDo,aezNestToDo1);
aezNestToDo(aezNest) = YES;

option kill=pAezNest.l;

While( card(aezNestToDo),
*
* --- we can calculate all nests which do not require information on nests
*   not yet calculated
*
option kill=curAezNestToDo;
curAezNestToDo(aezNestToDo) $ (not sum(aezNest_n(aezNestToDo,aezNestToDo1),1)) = YES;

xAezNest.l(r,aez,curAezNestToDo,t0) $ ( (not xAezNest.l(r,aez,curAezNestToDo,t0)) $ aezFlag(r,aez) )
= sum(aezNest_n(curAezNestToDo,aezNest1), xAezNest.l(r,aez,aezNest1,t0))
+ sum(aezNest_a(curAezNestToDo,a), xAez.l(r,aez,a,t0));

pAezNest.l(r,aez,curAezNestToDo,t0) $ (xAezNest.l(r,aez,curAezNestToDo,t0) $ aezFlag(r,aez))
= [ sum(aezNest_n(curAezNestToDo,aezNest1), xAezNest.l(r,aez,aezNest1,t0) * pAezNest.l(r,aez,aezNest1,t0))
+ sum(aezNest_a(curAezNestToDo,a), xAez.l(r,aez,a,t0) * pAez.l(r,aez,a,t0))
] / xAezNest.l(r,aez,curAezNestToDo,t0);

aezNestToDo(curAezNestToDo) = NO;
);
*
pAezNest.l(r,aez,aezNest,t0) $ aezFlag(r,aez) = pAezNest.l(r,aez,aezNest,t0);

xAezNestBase(r,aez,aezNest,t0) $ aezFlag(r,aez) = xAezNest.l(r,aez,aezNest,t0);
xAezNests.l(r,aez,aezNest,t0) $ aezFlag(r,aez) = xAezNest.l(r,aez,aezNest,t0);

```

If the user decides to aggregate over the AEZ (either for all model regions or for sub-national units, only), a new AEZ type called *aezAgg* is defined to which all data are aggregated:

```

$$onMulti
  set aez / aezAgg /;
$$offMulti

p_landUse(rs,"aezaggg",a,"vfm") = sum(aez, p_landUse(rs,aez,a,"vfm"));
p_landUse(rs,"aezaggg","land","vom") = sum(aez, p_landUse(rs,aez,"land","vom"));
p_landUse(rs,"aezaggg",lt,"ha") = sum(aez, p_landUse(rs,aez,lt,"ha"));
p_landUse(rs,"aezaggg",a,"ha") = sum(aez, p_landUse(rs,aez,a,"ha"));
p_landUse(rs,"aezaggg",a,"evfa") = sum(aez, p_landUse(rs,aez,a,"evfa"));

p_carbon(rs,"aezaggg",lt,"carbon") $ p_landUse(rs,"aezaggg",lt,"ha")
= sum(aez, p_carbon(rs,aez,lt,"carbon")*p_landUse(rs,aez,lt,"ha"))/ p_landUse(rs,"aezaggg",lt,"ha");

p_carbon(rs,"aezaggg",lt,"foreGoneSeqPerYear") $ p_landUse(rs,"aezaggg",lt,"ha")
= sum(aez, p_carbon(rs,aez,lt,"foreGoneSeqPerYear")*p_landUse(rs,aez,lt,"ha"))/ p_landUse(rs,"aezaggg",lt,"ha");

```

For post-model reporting purposes, original flags are stored, and then removed:

```

aezFlagOri(rs,aez) = aezFlag(rs,Aez);
aezFlag(rs,aez) = no;

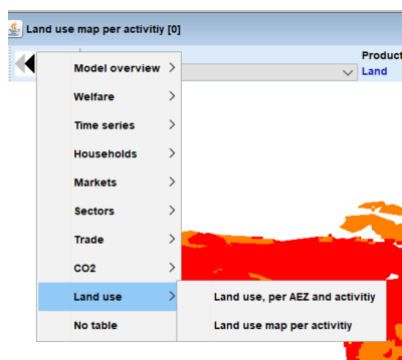
aezFlag(rs,"aezaggg") $ (sum(a, p_landUse(rs,"aezaggg",a,"vfm"))
and p_landuse(rs,"aezaggg","land","vom")) = Yes;

```

If this aggregation mechanism is active, proportionality assumptions are used to generate aez-specific results for reporting.

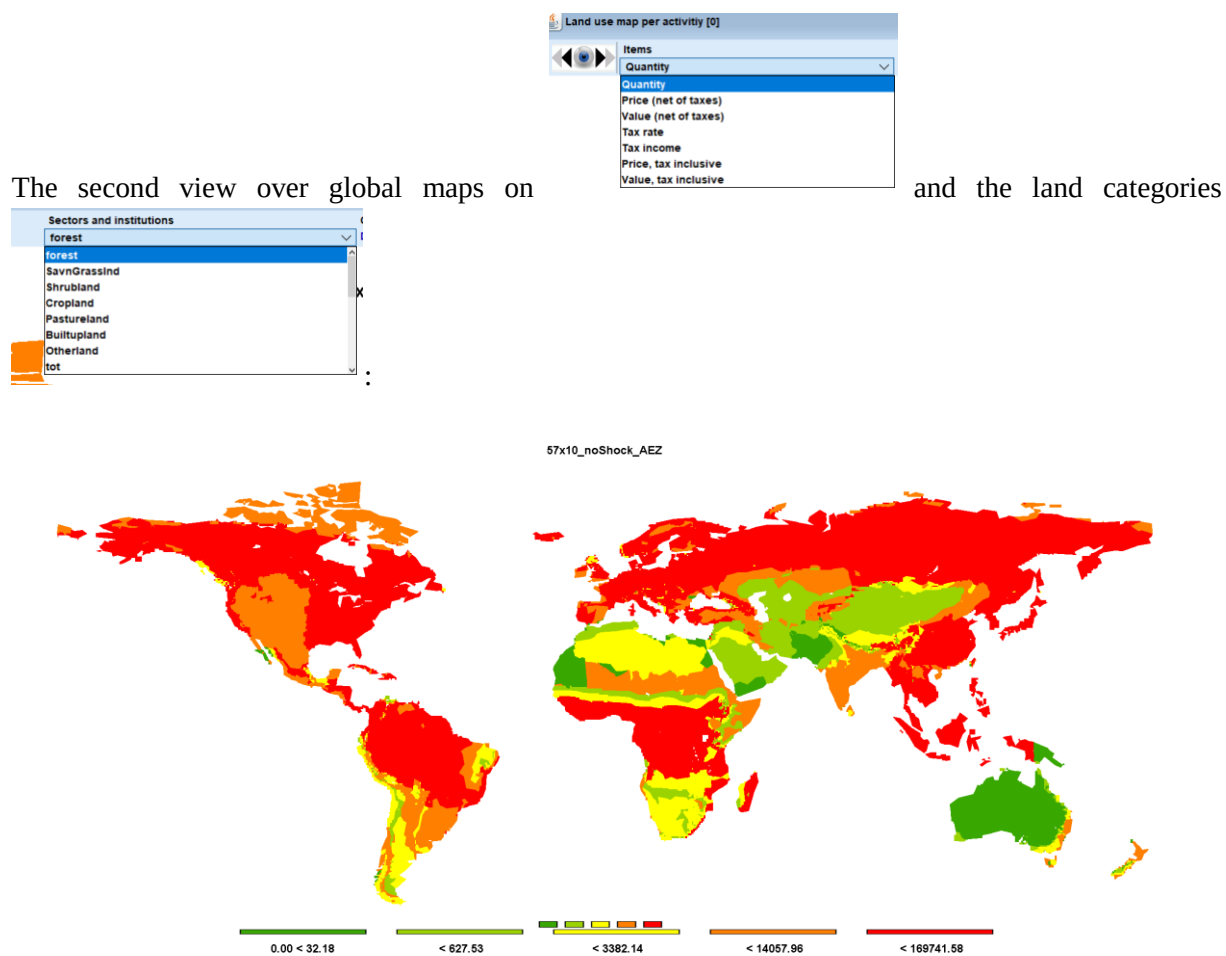
### 3.16.4 Visualization of GTAP-AEZ results

The views relating to results of GTAP-AEZ can be found under



The first view provides a tabular overview at global level, i.e. results are aggregated over regions to the global 18 AEZs:

| Land use, per AEZ and activity [0] |               |            |           |             |             |           |            |             |           |  |
|------------------------------------|---------------|------------|-----------|-------------|-------------|-----------|------------|-------------|-----------|--|
| Regions                            |               | Items      |           | Origins     |             | Years     |            |             |           |  |
| wor                                |               | Quantity   |           | Domestic    |             | shock     |            |             |           |  |
| 87x10_noShock_AEZ                  |               |            |           |             |             |           |            |             |           |  |
| forest                             | SavnGrassland | Shrubland  | Cropland  | Pastureland | Builtupland | Otherland | tot        | Paddy rice  | Wheat     |  |
| AEZ1                               | 5368.78       | 35709.01   | 149810.88 | 27488.91    | 202805.66   | 1110.16   | 199946.16  | 622239.56   | 1414.23   |  |
| AEZ2                               | 2200.48       | 74089.58   | 48676.56  | 89038.25    | 114788.45   | 527.87    | 97412.90   | 396734.09   | 1606.34   |  |
| AEZ3                               | 13716.38      | 164351.48  | 42867.55  | 122297.73   | 133454.83   | 1824.37   | 190384.25  | 668896.56   | 16434.10  |  |
| AEZ4                               | 64931.06      | 188750.91  | 26263.92  | 125355.00   | 184610.78   | 1972.18   | 236235.67  | 828119.50   | 30520.65  |  |
| AEZ5                               | 152295.12     | 165305.42  | 9033.02   | 122840.49   | 190275.94   | 2996.64   | 259542.78  | 902289.44   | 26842.60  |  |
| AEZ6                               | 339716.09     | 54816.60   | 4113.99   | 136558.34   | 98362.40    | 3871.14   | 371389.38  | 1008827.94  | 30220.06  |  |
| AEZ7                               | 8699.00       | 186182.55  | 483962.00 | 103135.23   | 759016.75   | 4380.51   | 429032.88  | 1974408.88  | 2856.97   |  |
| AEZ8                               | 25573.48      | 130645.14  | 72681.83  | 182671.56   | 273203.09   | 5679.49   | 294637.41  | 965092.00   | 3781.65   |  |
| AEZ9                               | 95863.79      | 87963.52   | 47098.09  | 173573.83   | 117137.34   | 5902.01   | 226083.73  | 763642.31   | 2612.37   |  |
| AEZ10                              | 231230.83     | 58136.07   | 32001.07  | 228214.50   | 117277.01   | 14034.57  | 245123.62  | 926017.69   | 3331.23   |  |
| AEZ11                              | 132749.08     | 36259.89   | 13443.31  | 107017.04   | 72205.89    | 9985.93   | 122886.27  | 484547.41   | 6821.41   |  |
| AEZ12                              | 148138.67     | 49226.12   | 2903.87   | 95872.81    | 110211.05   | 7391.83   | 168611.97  | 583354.12   | 15362.91  |  |
| AEZ13                              | 13291.09      | 43104.69   | 15982.81  | 29135.98    | 145489.50   | 278.95    | 63618.25   | 304891.25   | 110.68    |  |
| AEZ14                              | 200944.56     | 69155.51   | 3094.43   | 16597.48    | 97490.17    | 494.72    | 244624.50  | 632451.38   | 109.58    |  |
| AEZ15                              | 242743.47     | 57460.41   | 1673.26   | 31892.38    | 62462.44    | 1146.59   | 218235.50  | 615414.12   | 452.41    |  |
| AEZ16                              | 13472.57      | 12304.24   | 2363.53   | 2380.94     | 16928.22    | 95.36     | 3030.18    | 50575.04    | 107.92    |  |
| AEZ17                              | 1893.54       | 1672.16    |           | 32.22       | 861.61      | 1.35      | 248.43     | 4709.30     | 1.38      |  |
| AEZ18                              |               |            |           | 1.26        | 158.97      | 0.21      | 594.89     | 755.32      |           |  |
| tot                                | 1690826.00    | 1415133.25 | 955970.00 | 1563904.00  | 2703738.00  | 61693.88  | 3361690.75 | 11752966.00 | 142587.47 |  |



### 3.17 Climate Change Damages

The implementation here follows Roson & Sartori 2016 which develop a data base with Climate Change (CC) impacts for the GTAP V9 Data Base. An application to version 10 seems straightforward as the sole relevant difference seems the addition of Tajikistan in version 10 for which a proxy region from Version 9 can be defined. The implementation follows rather strictly what is proposed by Roson & Sartori 2016, exemptions are motivated below.

#### 3.17.1 Data input

##### Raw data processing

The basis for the current implementation is the supplementary material from Roson & Sartori 2016. In order to allow for a smoother integration into GAMS, the data were copy-and-pasted in a different presentation to an EXCEL work book and afterwards stored in GDX container called “ccDamage.gdx”.

The following screenshot shows how the data are organized. For each country, six different impact categories are present, which in case of labour productivity (*prodLab*) and changes in the energy consumption of household (*Energy\_HSLD*) are broken down further to sub-categories. In case of sea-level rise impacts at the same temperature increase, a differentiation between 2050 and 2100 is found in the raw data. To ease processing, these two years have been introduced for all impact categories. The last dimension depicts the temperature deltas.

| Plane Index (empty) |              |         |      |                      |                       |                       |                       |                      |
|---------------------|--------------|---------|------|----------------------|-----------------------|-----------------------|-----------------------|----------------------|
|                     |              |         |      | +1°C                 | +2°C                  | +3°C                  | +4°C                  | +5°C                 |
| AUS                 | prodLab      | AGR     | 2050 |                      |                       | -0,0109142280147764   | -0,0285779350094674   | -0,0554673197099693  |
|                     |              |         | 2100 |                      |                       | -0,0109142280147764   | -0,0285779350094674   | -0,0554673197099693  |
|                     | MAN          |         | 2050 |                      |                       |                       | -0,000211522238644735 | -0,00867679339769511 |
|                     |              |         | 2100 |                      |                       |                       | -0,000211522238644735 | -0,00867679339769511 |
|                     | landStock    | <empty> | 2050 | -0,00069213547798508 | -0,00114231521645835  | -0,00159249495493162  | -0,00204267469340489  | -0,00249285443187817 |
|                     |              |         | 2100 | -0,00138427095597016 | -0,0022846304329167   | -0,00318498990986325  | -0,00408534938680979  | -0,00498570886375633 |
|                     | Energy_HHSLD | Electr. | 2050 | -2,00960622378688E-5 | -3,15989420083961E-5  | -3,51893139802473E-5  | -3,14931270413276E-5  | -2,10864874671552E-5 |
|                     |              |         | 2100 | -2,00960622378688E-5 | -3,15989420083961E-5  | -3,51893139802473E-5  | -3,14931270413276E-5  | -2,10864874671552E-5 |
|                     |              | Gas     | 2050 | -0,00026309919323963 | -0,000524470808453779 | -0,000784470779228347 | -0,00104339298101878  | -0,00130147806289453 |
|                     |              |         | 2100 | -0,00026309919323963 | -0,000524470808453779 | -0,000784470779228347 | -0,00104339298101878  | -0,00130147806289453 |
|                     |              | Oil.P.  | 2050 | -0,0403849240953697  | -0,0796601317361653   | -0,117871475953865    | -0,15506227794245     | -0,191273502086637   |
|                     |              |         | 2100 | -0,0403849240953697  | -0,0796601317361653   | -0,117871475953865    | -0,15506227794245     | -0,191273502086637   |
|                     | Yield        | <empty> | 2050 | -0,019051744         | -0,040090688          | -0,063116832          | -0,088130176          | -0,11513072          |
|                     |              |         | 2100 | -0,019051744         | -0,040090688          | -0,063116832          | -0,088130176          | -0,11513072          |
|                     | Tourism      |         | 2050 | -0,0013958180893753  | -0,0030919127591279   | -0,00502925661768942  | -0,00714971781765354  | -0,00940100988336589 |
|                     |              |         | 2100 | -0,0013958180893753  | -0,0030919127591279   | -0,00502925661768942  | -0,00714971781765354  | -0,00940100988336589 |
|                     | Health       |         | 2050 | -0,00148875971692301 | -0,00148875971692301  | -0,00148875971692301  | -0,00148875971692301  | -0,00148875971692301 |
|                     |              |         | 2100 | -0,00148875971692301 | -0,00148875971692301  | -0,00148875971692301  | -0,00148875971692301  | -0,00148875971692301 |

#### 3.17.2 Aggregation

These raw data are read in “gtapAgg.gms” which prepares the data base for benchmarking of CGEBox. If the file with the damages is found on disk, sets are defined from the labels in the parameter above and the parameter is inputted:



```

$iftheni.ccDamages exist 'cc_Damag/cc_damages.gdx'

*
* --- read data from Roson-Sartori from GDX container
*
set ccShocks;
set ccSubShocks;
set temp;
set ccShockYears;

$$GDXIN 'cc_damag/cc_damages.gdx'
  $$load ccShocks < p_ccDamages.dim2
  $$load ccSubShocks < p_ccDamages.dim3
  $$load ccShockYears < p_ccDamages.dim4
  $$load temp < p_ccDamages.dim5

  parameter p_ccDamages(*,ccShocks,ccSubShocks,ccShockYears,temp);
  $$load p_ccDamages
$$GDXIN

```

Afterwards, the raw impact data are aggregated from single GTAP region (V9, 140 regions) to the model regions defined by the user. The following weights are used:

1. For labour productivity and health impacts, population:

```

set popWeighted(ccShocks) / prodLab,Health /;
*
* --- labour productivity and health impacts weighted with population
*
p_ccDamages(rNat,ccShocks,ccSubShocks,ccShockYears,temp) $ popWeighted(ccShocks)
= sum( rr(rdat,rnat), p_ccDamages(rDat,ccShocks,ccSubShocks,ccShockYears,temp) * popGtap(rDat))
/ pop(rnat);

```

2. Yield impacts are aggregated using crop land acreages:

```

*
* --- yield impacts weighted with crop land acreage
*
p_ccDamages(rNat,"yield",ccSubShocks,ccShockYears,temp) $ sum(aez,landUse(rNat,aez,"cropLand","ha"))
= sum( rr(rdat,rnat), p_ccDamages(rDat,"yield",ccSubShocks,ccShockYears,temp)
* sum(aez,landUse(rDat,aez,"cropLand","ha")) / sum(aez,landUse(rNat,aez,"cropLand","ha")) );

```

3. Land stock losses due to sea level rise are weighted with the sum of land-covers reported in the AEZ data base:

```

*
* --- landstock losses weighted with total area (including unmanaged land)
*
p_ccDamages(rNat,"landStock",ccSubShocks,ccShockYears,temp) $ sum(aez,land_cover),landUse(rNat,aez,land_cover,"ha"))
= sum( rr(rdat,rnat), p_ccDamages(rDat,"landStock",ccSubShocks,ccShockYears,temp)
* sum( (aez,land_cover),landUse(rDat,aez,land_cover,"ha"))
/ sum( (aez,land_cover),landUse(rNat,aez,land_cover,"ha")) );

```

4. Tourism impacts are weighted with the production value of the “hotels and restaurants” sector in GTAP:

```

*
* --- tourism (gdp% effect) weighted with total production value of hotel and restaurants
*
p_ccDamages(rNat,ccShocks,ccSubShocks,ccShockYears,temp) $ sum(rr(rDat,rNat),tVom("hht",rDat))
= sum( rr(rdat,rnat), p_ccDamages(rDat,ccShocks,ccSubShocks,ccShockYears,temp)
* tVom("hht",rDat))
/ sum(rr(rDat,rNat),tVom("hht",rDat));

```

5. The energy demand changes for the private household are weighted with the household demand at agent prices for the related GTAP products:



```

*
* --- energy effects weighed with energy demands at agent prices
*
p_ccDamages(rNat,"energy_hhsld","gas",ccShockYears,temp) $ sum(rr(rDat,rNat), vdpaGtap("gas",rDat)+vipaGtap("gas",rDat))
= sum( rr(rdat,rnat), p_ccDamages(rDat,"Energy_hhsld","gas",ccShockYears,temp)
      * (vdpaGtap("gas",rDat)+vipaGtap("gas",rDat)))
/sum(rr(rDat,rNat), vdpaGtap("gas",rDat)+vipaGtap("gas",rDat));

p_ccDamages(rNat,"energy_hhsld","oil.P.",ccShockYears,temp)
$ sum(rr(rDat,rNat), vdpaGtap("oil",rDat)+vipaGtap("oil",rDat)+vdpaGtap("p_c",rDat)+vipaGtap("p_c",rDat))
= sum( rr(rdat,rnat), p_ccDamages(rDat,"Energy_hhsld","oil.P.",ccShockYears,temp)
      * (vdpaGtap("oil",rDat)+vipaGtap("oil",rDat)+vdpaGtap("p_c",rDat)+vipaGtap("p_c",rDat)))
/sum(rr(rDat,rNat), vdpaGtap("oil",rDat)+vipaGtap("oil",rDat)+vdpaGtap("p_c",rDat)+vipaGtap("p_c",rDat));

p_ccDamages(rNat,"energy_hhsld","Electr.",ccShockYears,temp) $ sum(rr(rDat,rNat), vdpaGtap("oil",rDat)+vipaGtap("oil",rDat)
= sum( rr(rdat,rnat), p_ccDamages(rDat,"Energy_hhsld","Electr.",ccShockYears,temp)
      * (vdpaGtap("ely",rDat)+vipaGtap("ely",rDat)))
/sum(rr(rDat,rNat), vdpaGtap("ely",rDat)+vipaGtap("ely",rDat));

$endif.ccDamages

```

The resulting aggregated (and raw) CC impacts are then stored in the GDX container along with other data used for benchmark in a GDX container.

### 3.17.3 Graphical User Interface

The CC damage modules can be added to other modules via the following control:

| Modules and parameters     |                     |                       |             |
|----------------------------|---------------------|-----------------------|-------------|
| myGTAP                     | GTAP_AGR            | GTAP_AEZ              | GTAP_E      |
| CO2_Emissions              | Non_CO2_Emissions   | Air_Pollution         | MTOE        |
| Aggregate_GovToFinalDemand | Aggregate_Armington | Aggregate_firm_demand | GTAP_Melitz |
| ARM_LES_CES                | SPE                 | GTAP_MRIO             | TariffLines |
| CapVintages                | GTAP_NUTS2          | Parameters            | Nestings    |
| G-RDEM                     | GMIG                | CC_Damages            |             |

The interface allows to selected which of the six impact pathways are used

Impacts

Labour productivity

Human health

Available land

Crop yields

Household energy demand

Tourism

☒
☒
☒
☒
☒
☒

If the model is run in comparative-static mode, the temperature impact is defined by setting an increase for single shock:

Temperate increase

Shock [delta Degree Celsius]

For recursive-dynamic runs, three temp delta are given, the (already existing) base year increase, and the increase until 2050 and until 2101. There is difference in subtracting for the 2050 and 2100

number the base year effect. The chosen implementation should make it easier to input data from other sources:

| Temperate increase               |     |
|----------------------------------|-----|
| Base year [delta Degree Celsius] | 0.6 |
| 2050 [delta Degree Celsius]      | 1.0 |
| 2100 [delta Degree Celsius]      | 1.5 |

### 3.17.4 Shock definition

Currently, the code implementation does not yet support a combined use with G-RDEM. For baseline generation with G-RDEM, introducing the CC shocks makes limited sense as G-RDEM takes real GDP changes as given. Combining the CC-Damages with the information from the IAASA repository of GDP changes would assume that these GDP changes reflect these CC-Damages. It is therefore potentially more useful to first develop a baseline without CC damages in G-RDEM and next simulate the impact of CC-Damage against this baseline. This is the next step in code development. Currently, the code throws an error if G-RDEM is switched on.

#### Defining impacts for a specific year and temperature change

Identical to the data base processing step, first, the aggregated CC damages are inputted:

```
parameter p_ccDamages(*,*,*,*,*)
           p_ccDamagesApplied(*,*,*,*);

set ccShocks;
set ccSubShocks;
set temp;
set ccShockYears;
$$GDXIN "%resdir%/build/%dataset%.gdx"
  $$load ccShocks      < p_ccDamages.dim2
  $$load ccSubShocks   < p_ccDamages.dim3
  $$load ccShockYears  < p_ccDamages.dim4
  $$load temp          < p_ccDamages.dim5
  $$load p_ccDamages
$$GDXIN
```

Based on the information on the interface, we define the temperature increase in each year:

```
$$iftheni.recDyn "%dynMode%"=="Recursive dynamic"

  p_tempIncrease(t) $ (t.pos le 2050-%baseYear%)
    = %tempIncreaseBase% + (%tempIncrease2050%-%tempIncreaseBase%)/(2050-%baseYear%-1)*(t.pos-1);

  p_tempIncrease(t) $ (t.pos gt 2050-%baseYear%)
    = %tempIncrease2050%
    + (%tempIncrease2100%-%tempIncrease2050%)/50*(t.pos+%baseYear%-2050);

  t_ccShockYears(t,"2050") $ (t.pos le 2050-%baseYear%) = yes;
  t_ccShockYears(t,"2100") $ (t.pos gt 2050-%baseYear%) = yes;

$$else.recDyn

  p_tempIncrease(tRun) $ (not sameas(tRun,"%t0%"))
    = %tempIncreaseShock%;

  t_ccShockYears(t,"2050") $ (%tempIncreaseBase% le 2050) = yes;
  t_ccShockYears(t,"2100") $ (%tempIncreaseBase% ge 2050) = yes;

$$endif.recDyn
```

For the recursive-dynamic case, we consider the years between the base year and the year 2050 and the years between 2050 and 2100 differently, reflecting that the user can define different temperature changes on the interface for these two years. For the years up to 2050, the difference between the delta in 2050 and the already present temperature increase in the base year is distributed in equal steps over the simulation period up to 2050. Any remaining simulation years beyond 2050 receive a 1/50 of the difference in the increase in 2100 and 2050.

In case of a comparative static run, the inputted temperature delta is stored for the shock.

The above simple mapping for all years about 2050 to the 2100 sea level rise could be improved further with an additional interpolation.

As damages are reported for 1 degree deltas in temperatures, we interpolate linearly between the lower and upper deltas for a user inputted temperature delta:

```
*
* --- interpolate between two temperatures where shocks are given (e.g. 0.4 for 2.4, and the range 2+ and 3+)
*
p_interPolateTemp(tRun) = p_tempIncrease(tRun) - floor(p_tempIncrease(tRun));
*
* --- lower and higher temperature set element (e.g. 2+ and 3+ for a 2.4 increase)
*
tempIncrease(tRun,"lo",temp) $ ((p_tempIncrease(tRun) ge p_dtemp(temp)) $ (p_tempIncrease(tRun) lt p_dtemp(temp+1))) = Yes;
tempIncrease(tRun,"hi",temp) $ ((p_tempIncrease(tRun) lt p_dtemp(temp)) $ (p_tempIncrease(tRun) gt p_dtemp(temp-1))) = Yes;
```

Based on this information, the damages for the different years are mapped to a four dimensional parameter which is subsequently used to define the shocks:

```
*
* --- calculate damages for different years
*
p_ccDamagesApplied(rNat,ccShocks,ccSubShocks,tRun)
= sum(t_ccShockYears(tRun,ccShockYears),
      sum( tempIncrease(tRun,"lo",temp), p_ccDamages(rNat,ccShocks,ccSubShocks,ccShockYears,temp))
      + p_interPolateTemp(tRun)
      * [ sum( tempIncrease(tRun,"hi",temp), p_ccDamages(rNat,ccShocks,ccSubShocks,ccShockYears,temp))
          - sum( tempIncrease(tRun,"lo",temp), p_ccDamages(rNat,ccShocks,ccSubShocks,ccShockYears,temp)) ]
      );
```

## Mapping the impacts into shocks

### Labour productivity

For labour productivity, impacts are available for three broad sectors. These are linked to the detailed sectors in the model based on a cross-set (partly shown below):

```
set globSecD_iOri(globSecD,iOri) "Mapping of three broad sectors to original sectors in data base" /

AGR.    (set.icrops)
AGR.    ctl          "Cattle,sheep,goats,horses"
AGR.    oap          "Animal products nec"
AGR.    rmk          "Raw milk"
AGR.    wol          "Wool, silk-worm cocoons"
AGR.    frs          "Forestry"
AGR.    fsh          "Fishing"
MAN.    coa          "Coal"
MAN.    oil          "Oil"
MAN.    gas          "Gas"
$$ifthene.gtap10 %GtapNO%>9
MAN.    oxt          "Minerals nec"
$$else.gtap10
MAN.    omn          "Minerals nec"
$$endif.gtap10
MAN.    cmt          "Meat: cattle,sheep,goats,horse"
MAN.    omt          "Meat products nec"
MAN.    vol          "Vegetable oils and fats"
```

The shock changes the labour productivity shifter *lambdaf*:

```
$$iftheni.labProd "%CC_damage_labProd%"=="on"

  --- labour productivity shifts mapped from three broad sectors
      to individual ones, all labour categories

  lambdaf.fx(rs,l,a,tSim) = 1 + %rr%sum( (globSecD_iOri(globSecD,iOri),mapa(iOri,a)),
                                          p_ccDamagesApplied(%rPrice%,"prodLab",globSecD,tSim))
                                          /sum( (globSecD_iOri(globSecD,iOri),mapa(iOri,a)),1);
```

The set *l* relates to all labour categories in the benchmarking data set, *rs* for the (current) regions (nations and sub-regions thereof), *a* are the production activities and *tSim* is the current simulation year. For a comparative-static experiment, this refers to the label “*shock*”. The %rr% macro either is empty or depicts rr(rs,rsNat) in case that sub-national units are present in the model. The set mapa links original GTAP sectors to the activities in the model according to the make matrix

### Health impacts

Health impacts (if active) are added on top of these labour productivity shocks:

```
  lambdaf.fx(rs,l,a,tSim) = 1;
  $$endif.labProd
  $$iftheni.health "%CC_damage_health%"=="on"
*
*   --- add health effect on labour productivity
*
  lambdaf.fx(rs,l,a,tSim) = lambdaf.l(rs,l,a,tSim)
    + %rr% p_ccDamagesApplied(%rPrice%,"Health","",tSim)

  $$endif.health
```

### Tourism

Roson and Sartori (2016) develop the tourism impact based on estimates in changes of people travelling to specific destination which suggests in change in foreign income contributions. We therefore map there estimates in %GDP changes in changes in international transfers which enter the BOT==BOP balance of CGEBox. A loss of tourism revenues therefore implies that the net imports need to be reduced (or net exports to be expanded), interacting with the global bank mechanism of the GAP model:

```
  $$iftheni.tourism "%CC_damage_tourism%"=="on"
*
*   -- change in income inflow from abroad
*
  p_capTrans(rsNat,tSim) = rGdpmp(rsNat,tSim) * p_ccDamagesApplied(rsNat,"Tourism","",tSim);

  $$endif.tourism
```

### Energy demand changes for households

These are implemented as shifts of the CES share parameters in the energy nests which requires the nesting of the GTAP-E module. The code therefore aborts if these are not active.

```

$iftheni.hed "%CC_damage_hed%"=="on"
*
* --- change in household energy demand
*
$ifnot "%modulesGTAP_E%"=="on" $abort "CC impact on energy demand of household require nestings from GTAP-E, please switch on"

alphaa.fx(rsNat,"gas-c",h,tSim) = alphaa(rsNat,"gas-c",h,tSim) * (1+p_ccDamagesApplied(rsNat,"energy_hhslld","gas",tSim));
alphaa.fx(rsNat,"gdt-c",h,tSim) = alphaa(rsNat,"gdt-c",h,tSim) * (1+p_ccDamagesApplied(rsNat,"energy_hhslld","gas",tSim));
alphaa.fx(rsNat,"oil-c",h,tSim) = alphaa(rsNat,"oil-c",h,tSim) * (1+p_ccDamagesApplied(rsNat,"energy_hhslld","oil.p.",tSim));
alphaa.fx(rsNat,"p_c-c",h,tSim) = alphaa(rsNat,"p_c-c",h,tSim) * (1+p_ccDamagesApplied(rsNat,"energy_hhslld","oil.p.",tSim));
alphaa.fx(rsNat,ely,h,tSim) = alphaa(rsNat,ely,h,tSim) * (1+p_ccDamagesApplied(rsNat,"energy_hhslld","electr.",tSim));

$endif.hed

```

In the above,  $h$  are different households and  $\alpha_{haa}$  the cost shares in the CES-subnests in the demand system.  $Ely$  is the set of electricity products, which comprises either one element in case of the standard GTAP Data Base or multiple electricity carriers if the GTAP-Power Data Base is used.

### Crop yield changes

map crop yield changes into a change in tfp:

```

$iftheni.yields "%CC_damage_yields%"=="on"
*
* -- change in tfp of crops
*
axp.fx(rs,crop_a,tSim) = (1+ %rr%p_ccDamagesApplied(%rPrice%,"yield","",tSim));

$endif.yields

```

Where  $axp$  is the production shifter at the top of the production structure.

### Land loss due to sea level rise

This is implemented as a shifter of total land stock in economic use. Land in GTAP is only used in agriculture and forestry, the shock does not hence not affect the production possibility set of manufacturing or sectors directly.

The actual implementation differs depending on whether the AEZ module is switched on or not, and on detail of how the AEZ module is used:

```

*
* --- change in economic land stock
*
$iftheni.aez "%modulesGTAP_AEZ%"=="on"

p_xAezNestBase(rs,aez,"top",tSim)
= p_xAezNestBase(rs,aez,"top",t0) * (1+ %rr%p_ccDamagesApplied(%rPrice%,"landstock","",tSim));

$iftheni.landSupply "%landSupplyElas%"=="on"

v_topLandChange.fx(rs,aez,tSim) = xAezNest(rs,aez,"top",tSim)
* %rr%p_ccDamagesApplied(%rPrice%,"landstock","",tSim);

$endif.landSupply
$else.aez

p_aft(rs,lnd,tSim) = p_aft(rs,lnd,tSim) * (1+ %rr%p_ccDamagesApplied(%rprice%,"landstock","",tSim));
xft.fx(rs,lnd,tSim) $ (xft.range(rs,lnd,tSim) eq 0) = p_aft(rs,lnd,tSim);

$endif.aez
$endif.landstock

```

## 3.17.5 References

Roson, R., & Sartori, M. (2016). Estimation of climate change damage functions for 140 regions in the GTAP9 database. *Policy Research Working Paper 7728, World Bank, Available at:*

<https://openknowledge.worldbank.org/bitstream/handle/10986/24643/Estimation0of00n0the0GTAP90database.pdf?sequence=1>

## 4. Data base generation

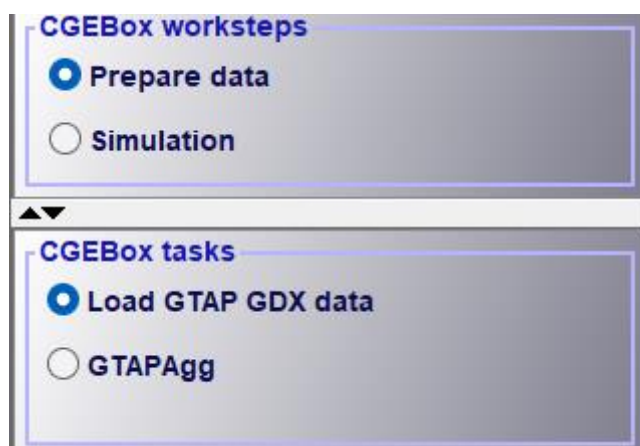
A CGEBox simulation run cannot be started from a GTAP data base downloaded from the GTAP center's website. The structure of this data base follows rather strictly the approach in GEMPACK, i.e. for each endogenous and exogenous variables in the GTAP model defined in GEMPACK, matching starting values need be defined. Even if the GTAP Standard model is the core of the equations in CGEBox, it defines benchmark values mainly from a global SAM defined from the various arrays that jointly defined a GTAP data base. The data base driver GTAPAgg (gams/gtapAgg.gms) defines this SAM plus various additional data such as emission factors, behavioral and other parameters, data related to the Armington decomposition and bi-lateral trade not comprised in the SAM. GTAPAgg combines functionalities of multiple GEMPACK utilities, such as for aggregation (GTAPAgg), splitting (splitcom), data base correction (GTAPAdjust), and adds additional ones, such as filtering out small transactions or deriving MRIO split shares.

During the data base generation, the user decides:

- which version of the GTAP data base to start from. This includes the option to use a data base in GTAP format prepared by GTAPAgg, for instance, to chain different split steps.
- which regional, product, sector and factor resolution to use, i.e. how to aggregate the data base
- if tiny costs shares are to be removed, called filtering (recommended);
- which additional data to introduce (sub-regional data for non-EU countries, MRIO split factors, gender differentiated labor accounts);
- if specific sectors and products are subject to a further split, based on user supplied split shares
- If the resulting data should be stored again as a data base which can be subject to the same steps

Generating benchmark data sets requires selecting the work step "Prepare data" on the GUI. Before data can be subject to the steps above, they need to be downloaded as GDX containers from the GTAP Website and imported by the task "Load GTAP data". Thanks to harmonization of the GTAP Data since Version 9, this streamlined process to integrate various data sources into a CGEBox installation is now available as documented next. It feeds into the task "GTAPAgg" which is run afterwards.

The first version of CGEBox were based on import from GTAPAGG. It was removed as it cannot handle additional sets (AEZ, Non-Co2 emissions, etc.) in a streamlined format.

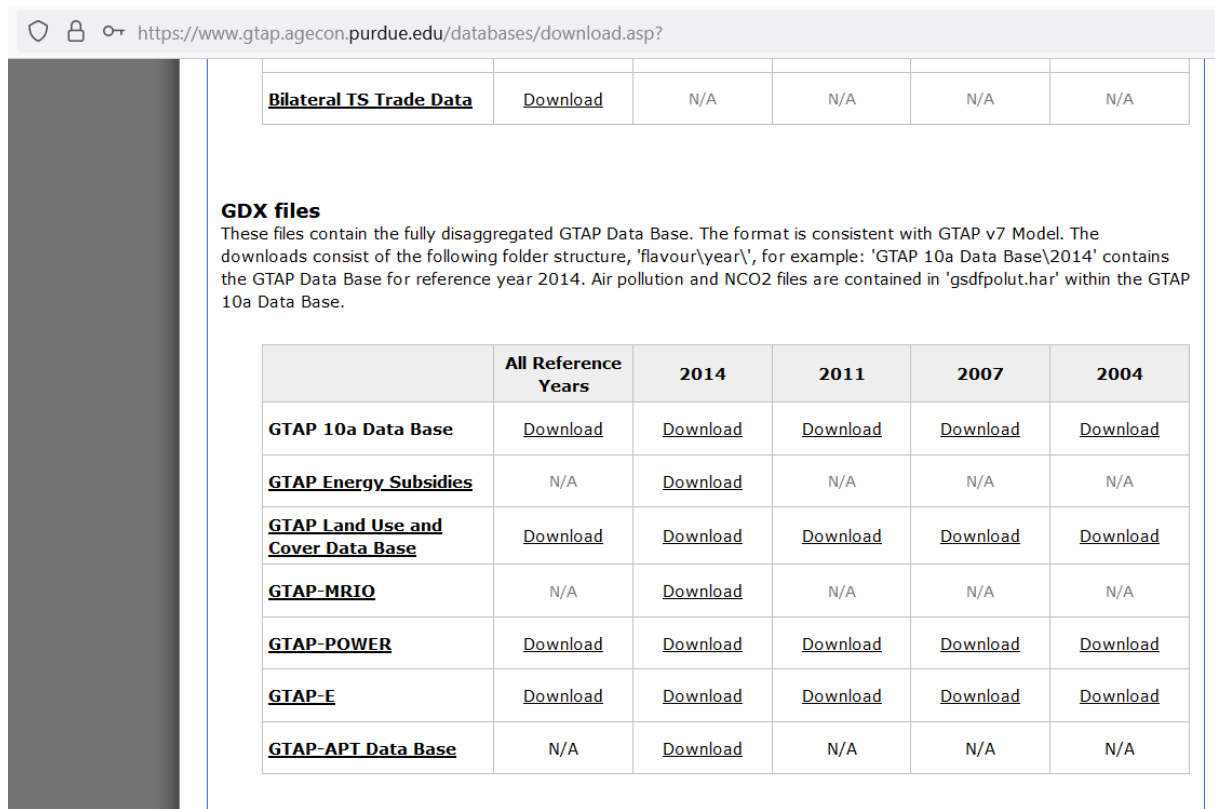




## 4.1 Loading a GTAP Data Base

### 4.1.1 Background

Since 2022, the GTAP center distributes almost all data also as GDX containers (access to the page shown in the screenshot below requires a GTAP license):



The screenshot shows the GTAP website's download page. At the top, there is a table with one row: **Bilateral TS Trade Data** with a [Download](#) link and 'N/A' in the other columns. Below this is a section titled **GDX files** with a paragraph explaining that these files contain the fully disaggregated GTAP Data Base in a format consistent with GTAP v7 Model. It mentions that downloads consist of a folder structure 'flavour\year\' and gives an example: 'GTAP 10a Data Base\2014' contains the GTAP Data Base for reference year 2014. It also notes that air pollution and NCO2 files are in 'gsdflolut.har' within the GTAP 10a Data Base. Below the text is a table listing various data bases and their availability for different years.

|                                          | All Reference Years      | 2014                     | 2011                     | 2007                     | 2004                     |
|------------------------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <b>GTAP 10a Data Base</b>                | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> |
| <b>GTAP Energy Subsidies</b>             | N/A                      | <a href="#">Download</a> | N/A                      | N/A                      | N/A                      |
| <b>GTAP Land Use and Cover Data Base</b> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> |
| <b>GTAP-MRIO</b>                         | N/A                      | <a href="#">Download</a> | N/A                      | N/A                      | N/A                      |
| <b>GTAP-POWER</b>                        | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> |
| <b>GTAP-E</b>                            | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> | <a href="#">Download</a> |
| <b>GTAP-APT Data Base</b>                | N/A                      | <a href="#">Download</a> | N/A                      | N/A                      | N/A                      |

CGEBox supports now the direct imports of the core data base (first line), of the Land use (AEZ), MRIO and GMIG data. The core data also comprises the CO<sub>2</sub>, non-Co<sub>2</sub>, air pollution and energy accounting data that were in earlier version comprised in separate data bases.

## 4.2 Importing the data into CGEBox

In order to import the data into CGEBox, first generate a new folder under the data directory with a senseful name, such “GTAPV10\_2014” with, for instance, the Windows explorer. If, for instance, your cgebox installation is hosted on “D:/CGEBOX”, generate such a subfolder under “D:/CGEBOX/DATA”, for instance, “D:/CGEBOX/DATA/GTAPV10\_2014”. A number of folders with default aggregation files are already present with any CGEBox installation.

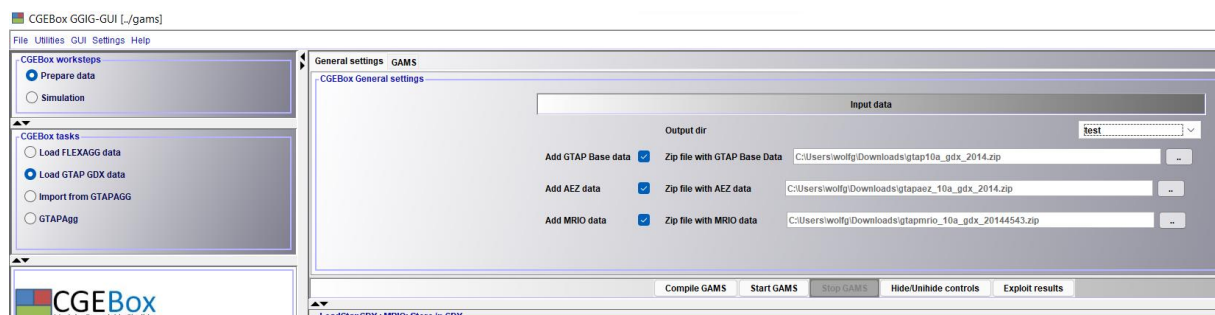
Next, download the core and AEZ Data from the GTAP website which are always processed by the data driver of CGEBox, choosing the single year edition. You should afterwards two zipped archives in the folder chosen for the download as shown in the example below:

## Importing the data into CGEBox

| Downloads |                          |                  |                      |            |
|-----------|--------------------------|------------------|----------------------|------------|
|           | Name                     | Änderungsdatum   | Typ                  | Größe      |
| Heute (3) |                          |                  |                      |            |
|           | gtap10a_gdx_2014.zip     | 08.03.2022 05:24 | ZIP-komprimierter... | 127.678 KB |
|           | gtapaez_10a_gdx_2014.zip | 08.03.2022 14:39 | ZIP-komprimierter... | 116.499 KB |

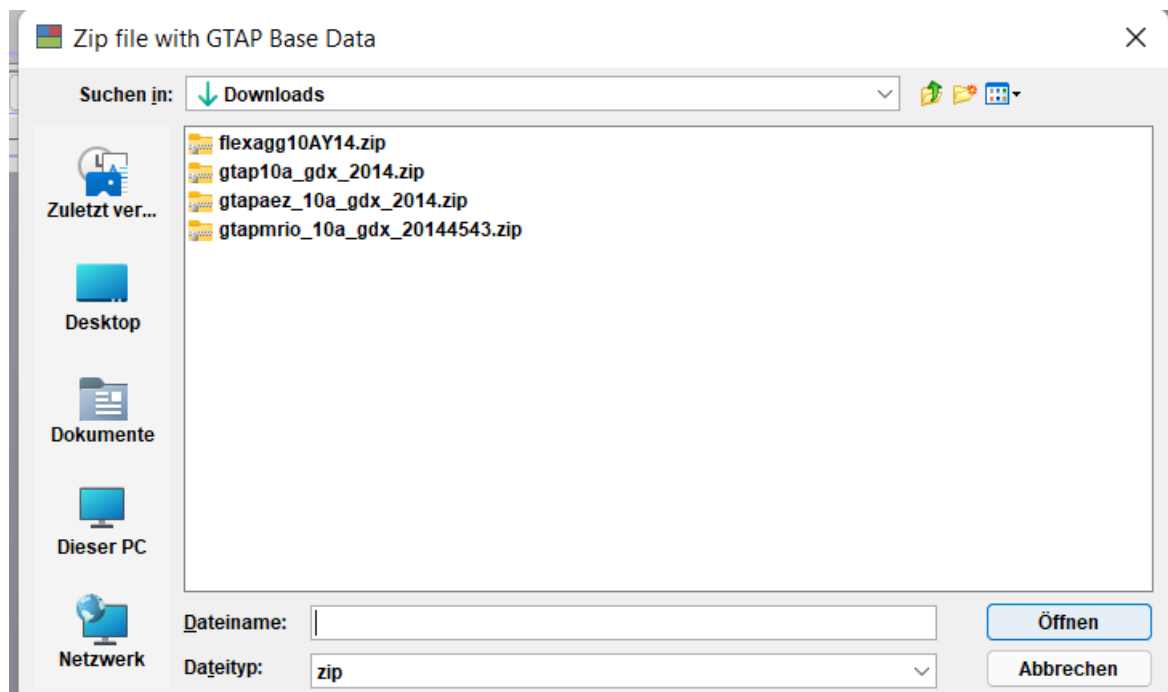
Remember into which directory you have downloaded the data, this location is needed for the next step, the import into CGEBox. You don't need to unzip the downloaded containers or move them to another location for the import step.

Next, open the CGEBox GUI, select the work step "Prepare data" and the task "Load GTAP GDX data":



As seen from the interface, you need to select in the first line the sub-directory in the data folder into which you will import, here "test". The folder should have been generated by you as proposed above or copies from the SVN repository when you installed CGEBox.

You can choose which data parts (base data, AEZ, MIO) to add based on the three check boxes. Pressing one of the button ".." to the right of the three text fields opens a file dialog as shown below. You use this dialog to move to the right folder and selected the zipped archive to load.





Afterwards, press “Start GAMS” button. The GAMS script will unzip the selected downloaded ZIP archives into the scratch directory, for instance:

```
--- call gmsunzip -j -o "C:\Users\wolfo\Downloads\gtapaez_10a_gdx_2014.zip" -d "C:\smdir\aez"
Archive:  C:/Users/wolfo/Downloads/gtapaez_10a_gdx_2014.zip
  inflating: C:/smdir/aez/gsdmdat.gdx
  inflating: C:/smdir/aez/gsdmiss.gdx
  inflating: C:/smdir/aez/gsdpar.gdx
  inflating: C:/smdir/aez/gsdSAM.gdx
  inflating: C:/smdir/aez/gsdset.gdx
  inflating: C:/smdir/aez/gsdshock.gdx
  inflating: C:/smdir/aez/gsdftax.gdx
  inflating: C:/smdir/aez/gsdview.gdx
  inflating: C:/smdir/aez/gsdvole.gdx
  inflating: C:/smdir/aez/readme.txt
```

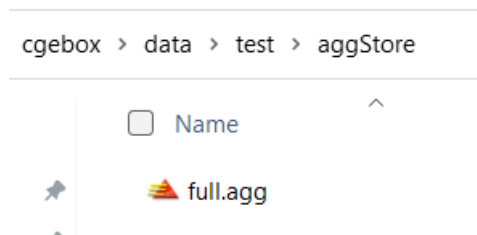
Please make sure that you have enough disk space in the scratch directory. If no scratch directory is explicitly set, the scratch directory of GAMS (225...) will be used. Note that the MRIO data require >1 GB of disk space for processing.

The process for all three data sources currently supported requires on a fast laptop less than a minute.

You can check if this step worked by opening with the Windows explorer or a similar tool the directory chosen for output, you should find the following files (the file mrio.. will be generated only if the MRIO data are added; the file aez\_d.gdx only if the AEZ data have been added.):

| Name                                                                                                                    | Änderungsdatum   | Typ         | Größe      |
|-------------------------------------------------------------------------------------------------------------------------|------------------|-------------|------------|
|  <b>aggStore</b>                       | 08.03.2022 17:11 | Dateiordner |            |
|  <b>aez_d.gdx</b>                    | 08.03.2022 17:11 | GAMS file   | 267 KB     |
|  <b>gsdmdat.gdx</b>                  | 18.01.2022 17:34 | GAMS file   | 106.325 KB |
|  <b>gsdmiss.gdx</b>                  | 18.01.2022 17:34 | GAMS file   | 867 KB     |
|  <b>gsdpar.gdx</b>                   | 18.01.2022 17:34 | GAMS file   | 547 KB     |
|  <b>gsdpsolut.gdx</b>                | 18.01.2022 17:34 | GAMS file   | 3.730 KB   |
|  <b>gsdSAM.gdx</b>                   | 18.01.2022 17:34 | GAMS file   | 61.325 KB  |
|  <b>gsdset.gdx</b>                   | 18.01.2022 17:34 | GAMS file   | 9 KB       |
|  <b>gsdshock.gdx</b>                 | 18.01.2022 17:34 | GAMS file   | 12.147 KB  |
|  <b>gsdftax.gdx</b>                  | 18.01.2022 17:34 | GAMS file   | 12.146 KB  |
|  <b>gsdview.gdx</b>                  | 18.01.2022 17:34 | GAMS file   | 84.675 KB  |
|  <b>gsdvole.gdx</b>                  | 18.01.2022 17:34 | GAMS file   | 2.266 KB   |
|  <b>mrio_splitfactor_v10_new.gdx</b> | 08.03.2022 17:11 | GAMS file   | 57.149 KB  |
|  <b>readme.txt</b>                   | 13.12.2021 20:36 | TXT-Datei   | 1 KB       |

The AggStore directory should comprise a file with a 1:1 aggregation called “full.agg”:



## 4.3 GTAPAGG: The data base generation utility of CGEBox

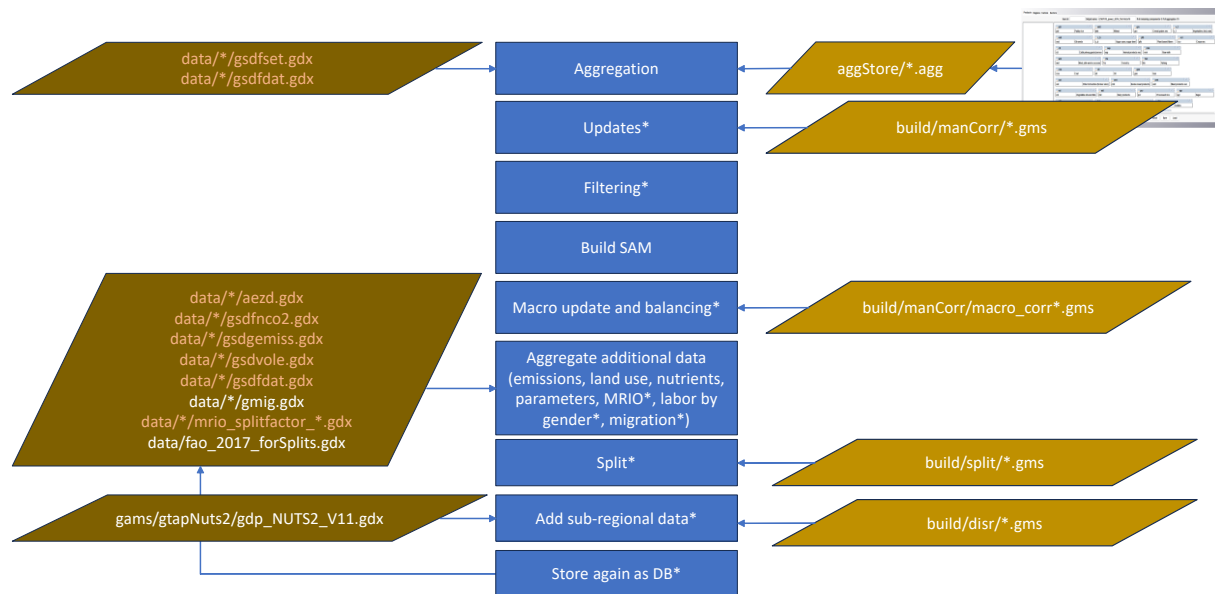
### 4.3.1 Background

GTAPAGG is a GAMS based utility that converts the GTAP data base plus additional data into a data set suitable for benchmarking a CGE model set-up in CGEBox. The data are stored as one GDX container and comprise, besides the data for myGTAP provided by the user when running a simulation, all necessary information for benchmarking.

The basic steps of GTAPAGG are as follows:

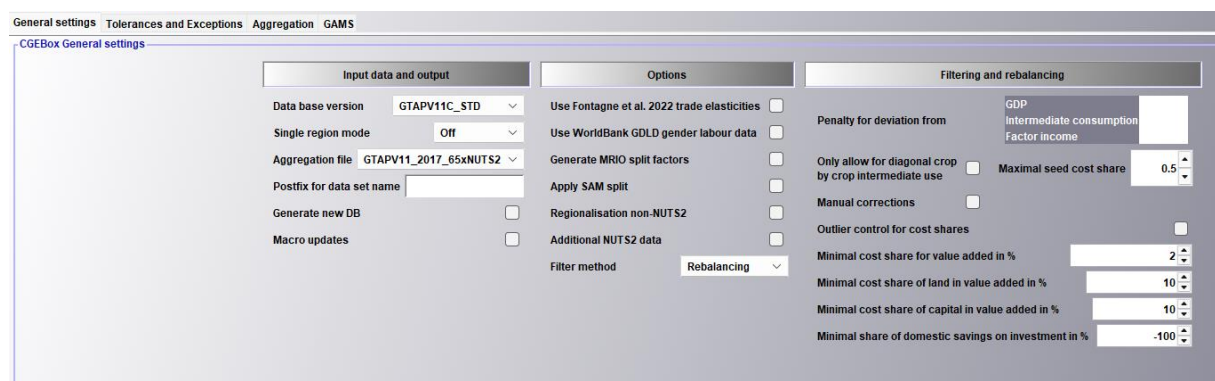
1. Define the desired aggregation by region, products, activities and factors
2. Aggregate the GTAP data based on this aggregation scheme
3. On demand, introduce changes to the GTAP data. These changes should be minor updates of individual transaction, not changes affecting macro totals (such as GDP, factor income, B.O.T etc.)
4. On demand, filter the resulting data to remove small cost or expenditures shares
5. Build a SAM from the aggregated (and updated) data
6. On demand, introduce macro total updates and re-balance the global data
7. Add additional data (emissions, land use, nutrients, MRIO, labor by gender, migration) and model parameters (trade elasticities, substitution elasticities, parameters of the CDE demand system), including on demand gender different data and labor and MRIO factors
8. On demand, split the data to additional product and sector detail
9. If sub-national data are available, render them consistent to the national totals and integrate them into the SAM
10. On demand, store the data back in the format of the GTAP data base
11. Store all data, sets, parameter in one GDX for benchmarking.

The graphic below shows these steps in the box in the middle. The boxes on the left-hand side show data that are either imported from the GTAP data base (in orange) or part of the repository of CGEBox. Note that a GDX version of GMIG is currently not available but expected to become available soon. Data on the right-hand-side are provided by the user, some examples are part of the repository. Steps and data that can be added on demand show an asterisk.



### 4.3.2 Overview

The screenshot below gives an overview on the core options available with data base generation utility of CGEBox, called GTAPAgg. It aggregates the GTAP Data Base (Standard, Power, CE) and additional data (CO<sub>2</sub>, Non-Co<sub>2</sub>, air pollution, migration, land use ...) according to user defined aggregation rules, filters out small numbers and allow further steps, such as applying a split or adding sub-national data.



The outcome of the data base utility is one single GDX container stored in the ‘results/build’ directory, under the name aggregation file plus a potential postfix entered by the user. It comprises about 80 different GAMS symbols. This GDX container along with code and data under version control comprises all necessary data and parameters to perform simulation runs with CGEBox, including its extensions.

Such GDX containers can hence also be exchanged between CGEBox users. For instance, one team member can generate a suitable data base with the GTAPAgg utility and share the container with all others. Installing the GTAP Data Base and other data sets from the GTAP center (AEZ, non-CO<sub>2</sub> etc.) is necessary only to build these GDX containers, not for simulation runs.

Please note that the GDX containers comprise the copy-right protected data from the GTAP center for which a license is required. The license for Version 10 states (April 2022): “Licensees are free to distribute small aggregations or RunGTAP applications for joint work or teaching purposes. However,

the licensee must obtain permission from the Center for the distribution of large aggregations. A large aggregation is defined as one in which there are more than 10 regions or more than 10 commodities.”<sup>115</sup> Adhering to these requirements, GDx containers up to a 10x10 aggregation might hence be shared with co-authors or students using CGEBox not having the required GTAP license. Sharing more detailed data with somebody who is not a licensee of the GTAP data base from which the data have been derived violates the license agreement. It deprives the center from necessary license revenues needed to produce future data bases and thus hurts the global community of CGE modelers.

#### 4.3.3 Defining your aggregation: The aggregator component

The first step usually consists of generating an aggregation definition file. This is done with the aggregator component found when the tab “Aggregation” is active. The aggregator component replaces the GTAPAgg utility form the GTAP center working with HAR-Files. It allows to generate non-diagonal aggregations where the sector and product aggregations differ. It outputs a file in a format similar to the one used for the GTAPAgg utility. The user can also generate such files with a text editor at its own risk as certain consistency checks cannot be easily performed on the aggregation files.

The component is integrated on a tabbed pane in the GUI controls of the data driver under the tab “Aggregation” and comprises itself a tabbed pane with the four tabs “Products”, “Regions”, “Factors” and “Sectors”. The aggregation control as the central control shown above in Figure 16 comprises on the LHS a field with the list of unassigned components (here empty) and on the RHS an area where new aggregates can be defined and existing ones are shown. The left area is empty in Figure 16 as the default aggregation rules, loaded at start from “*default.agg*”, correctly assigns already all components to an aggregate. While the functionality is similar to the user interface of GTAPAgg, the GUI component uses intensively “drag & drop” as a faster option with a more modern Look and Feel.

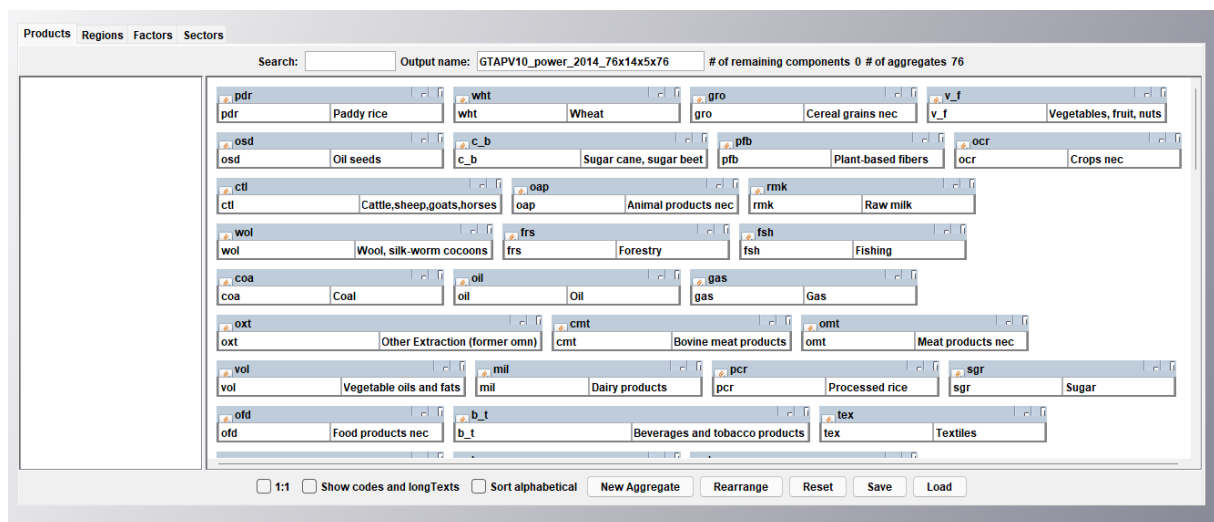


Figure 16: Aggregator component on a tab of the GUI

There are two ways to define aggregates:

1. Using the button “**New aggregate**” (in the control bar at the bottom) will generate a new window on the RHS to host components with a default label - “New” plus a number – as

<sup>115</sup> [https://www.gtap.agecon.purdue.edu/databases/documents/LicenseAgreement\\_v10.pdf](https://www.gtap.agecon.purdue.edu/databases/documents/LicenseAgreement_v10.pdf), visited 30.04.2022

shown in Figure 17 below. The user will overwrite this label in the label text field below the title base to give the new aggregate its desired name and potentially also provide a long text in the text field to the right of the label field. That is similar to the GTAPAgg solution. The field below the labels will hold the components, also similar to GTAPAgg. What differs is that each aggregate is visualized as a small window which can be dragged, defining the order of the aggregate, or can be minimized to leave more room to define larger aggregation. In order to populate an aggregate, components are dragged from the LHS into the area below the two text fields (which is empty in the screenshot below).

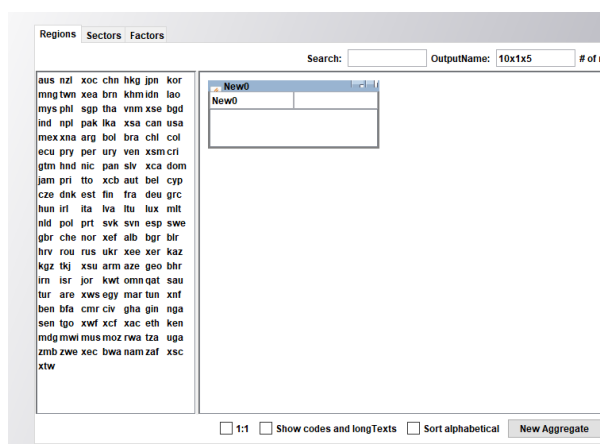


Figure 17: A new empty aggregate generated from “New Aggregate”

2. The second option **defines aggregates directly by drag and drop from the LHS or from existing aggregates**. In both cases, a list of components - a single one, one or more sequences etc. – is selected with the mouse (keep the CTRL key of the keyboard pressed to select multiple ones) and next jointly dragged to the RHS. If the drop ends in a window of an existing aggregate, the components are added to it. Otherwise, if the drop ends somewhere in the RHS outside of an aggregate window, a new aggregate is defined comprising the selected components:

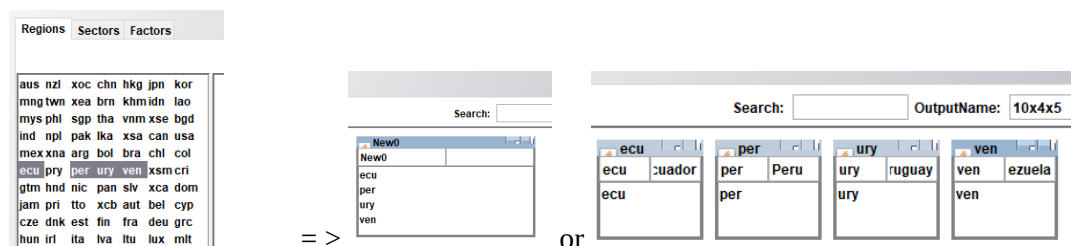


Figure 18: Aggregates defines by drag and drop examples

That behaviour changes if the user has activated the “1:1” option ☒ 1:1. In that case, (multiple) aggregates are defined from the dragged selection, each comprising a single component, and the label and long-text are automatically taken over from the component, as shown above on the right example. The drag & drop approach to generate new aggregates is probably the more frequently used one compared to first generating a new aggregate and next dragging components into it.

In order to generate a diagonal aggregation with the same aggregate definition for products and sectors, only define product aggregates and keep all single sectors in the LHS window on the “Sector” tab (or press “RESET” to delete all sector aggregate). Please check that the combination of sector and

product aggregates is useful. For instance, the default 10x10 aggregation uses ten world regions and ten aggregate products each produced by one sector. Solely dis-aggregating the products from this aggregation makes limited sense.

Further functionalities embedded in the aggregation tool cover:

- 3. Aggregates can be deleted with the “Delete” button found in their pop-up menu, their components are moved back into to LHS panel. The “Fit” button from the same pop-up menu will adjust the size of the aggregate window such that all components are visible.

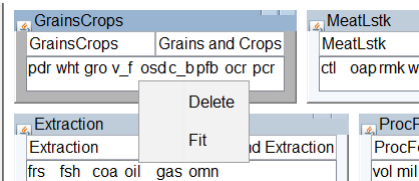


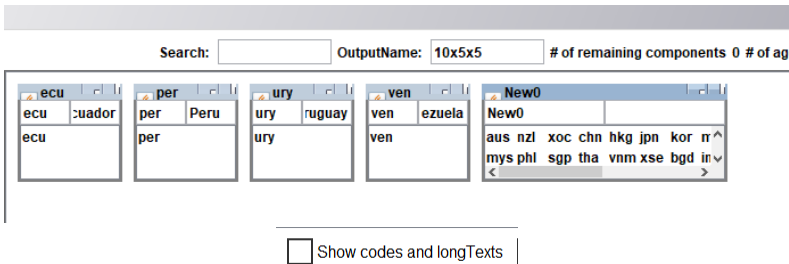
Figure 19: Pop-up menu for aggregate windows

- 4. Attention: the “Reset” button deletes all aggregates on the current tab; consequently, all components are moved back into the LHS. A typical application is to delete all sector aggregates as loaded from “default.agg”, to select next the ☒ 1:1 option and finally drag in one go all original data base sectors to the RHS to generate an aggregation definition will full sectoral detail. Subsequently, one could still drag some components into others to define aggregates. Note however that empty aggregates are not automatically deleted. That might be unwanted if one has edited the labels and long texts of aggregates already and next moves components around, generating temporary states where aggregates are empty.
- 5. The “Rearrange” button resizes all aggregates again and orders them (from top to down and left to right), that is also the order in which they are outputted and e.g. shown in model listings or later, for instance, in tables of the exploitation tool. One can hence simply drag an aggregate up or towards the left to move it up in the ordered list.
- 6. The number of aggregates and not allocated components are shown in a pane on top. It also comprises a text field with a generated default name of the output file – it can be overwritten by the user - and offers a search field (see further explanation to the search field below).



Figure 20: Text fields are search, name of output file and information label

- 7. The window of an aggregate can also be temporarily minimized. The title bar shows the current aggregate name in that case.
- 8. The user can switch between the wrapped lists with labels and lines each with a label – long text pair, as shown in Figure 21 below. 1:1 elements will hide their component area.



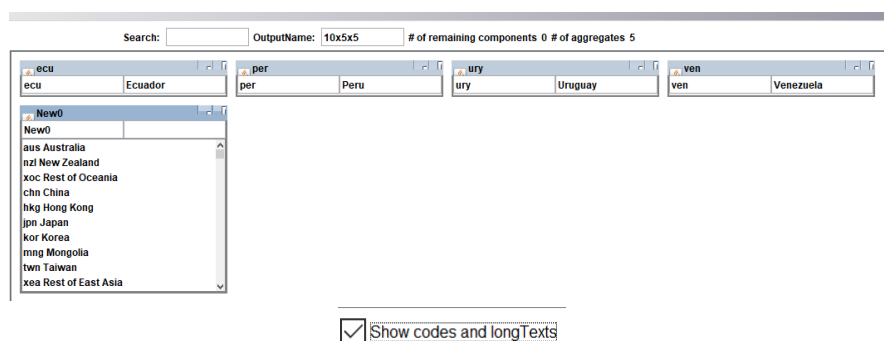


Figure 21: Aggregates shown with code or code and long texts

9. The “search” field allows finding a component, it becomes selected. If more than three characters are entered, the search will switch from a “Start with” to a “contains” type search in the long texts if these are shown.

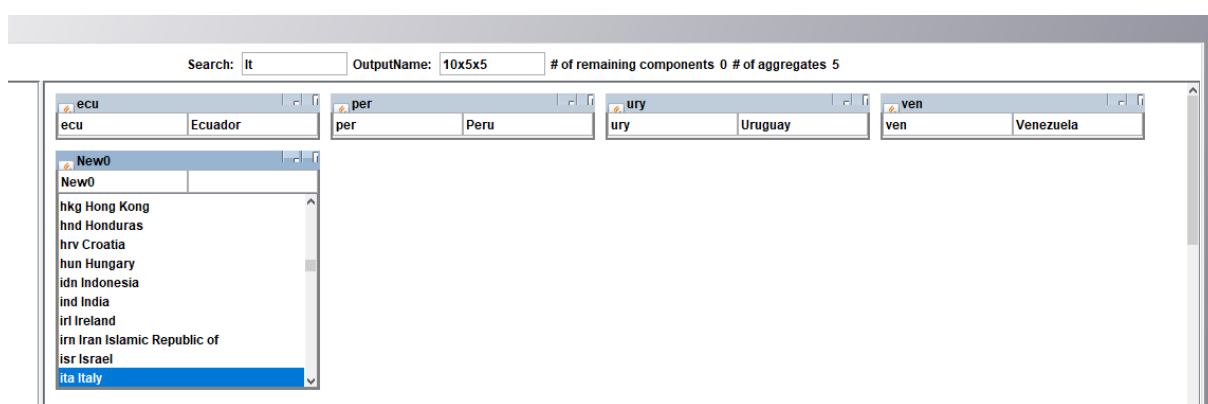


Figure 22: Example of search field

10. The user has the choice to show components either in the order as read from “default.agg” or sorted alphabetically:

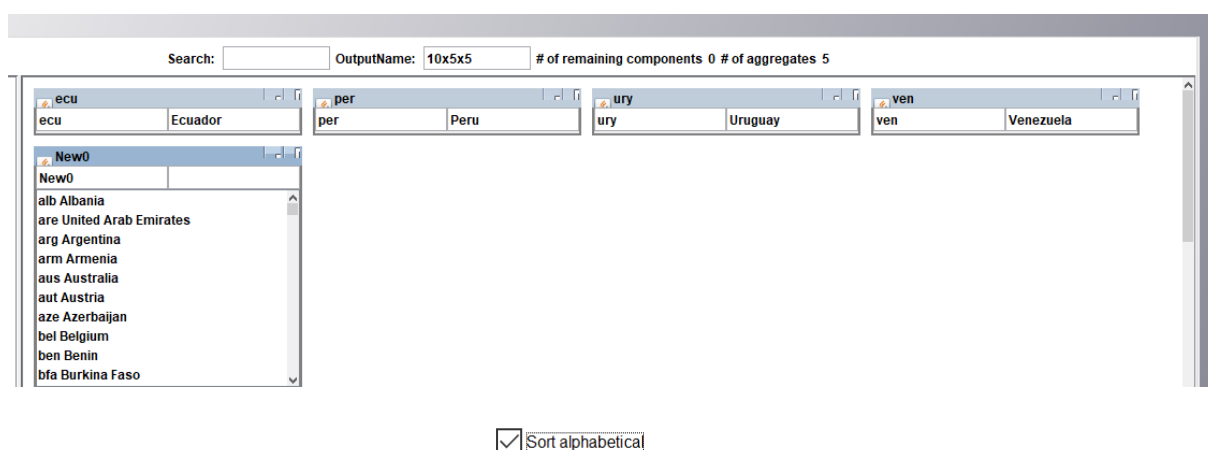


Figure 23: Alphabetical sorting

The “save” button  will save the new definition as an .agg file in the “aggstore” directory.

For each data base version, there is a separate folder in the data directory. Each such folder comprises its own “aggStore” folder for matching aggregation definitions.



For data base versions successfully tested with CGEBox, an empty folder with a sub-folder “aggStore” comprising a matching a default aggregation will be committed. The user can install a zipped GDX container as discussed above via the GUI in such pre-defined folders.

#### 4.3.4 Aggregation in GAMS

Once you have generated a suitable aggregation definition file and chosen it and you press start, the GAMS script will aggregate the data. If you are not interested on how this process works, you can skip this section.

The code of this script follows the set definitions and builds on the code of ENVSIAGE by Dominique Van der Mensbrugghe. In order to ease inspection and debugging, the code does not assign aggregated GTAP symbols directly to the SAM or other matrices or vectors used by CGEBox simulation codes. Instead, it defines first the GTAP symbols at the level of regional / sector / factor aggregates, similar to the output from GTAPAgg, e.g.

```
vdga(i,rNat) = sum( (ii(prod_comm,i),rr(rDat,rNat)) $ vdgaGtap(prod_comm,rDat), vdgaGtap(prod_comm,rDat));
vdgm(i,rNat) = sum( (ii(prod_comm,i),rr(rDat,rNat)) $ vdgmGtap(prod_comm,rDat), vdgmGtap(prod_comm,rDat));
vdpa(i,rNat) = sum( (ii(prod_comm,i),rr(rDat,rNat)) $ vdpaGtap(prod_comm,rDat), vdpaGtap(prod_comm,rDat));
vdpm(i,rNat) = sum( (ii(prod_comm,i),rr(rDat,rNat)) $ vdpmGtap(prod_comm,rDat), vdpmGtap(prod_comm,rDat));
```

Figure 24: Example of aggregation code

This approach eases debugging potential problems during aggregation.

Before assigning the aggregated base data entries to the SAM and other matrices / vectors, the filter program can be used to remove small entries as discussed below. It can be combined with manual correction of the data. It is based on linear programming for fast balancing and ensures a quite high feasibility accuracy. Accordingly, besides filtering, small numerical rounding errors as typically found in the original data are further reduced. Additionally, a Generalized RAS is run independent from filtering or not. The code will throw an error if numerical differences between column and rows sum greater 1E5 USD are found and the error in relative terms to the sum is higher than 0.01%.

The code aggregates in a similar fashion as shown for the parameters above

- The data from the GTAP data base and related parameters
- The CO2 emission data
- The NON-CO2 emission data
- The air pollution data
- The MTOE data
- The AEZ related data (here, only data at AEZ level related to land use are stored)
- GMIG related data
- MRIO data (on demand)

The data base attached per ha carbons stock to different land cover types. The underlying data were released with Version V10. All data necessary for a simulation run including the sets are stored into a GDX container in the build directory. The user can for debugging purposes store all output in a GDX but it is not guaranteed that such a container will work in simulation mode as the code renames some symbols if not all symbols are put into a GDX container. This option requires a user level of “debugger”.



#### 4.3.5 Backward compatibility

Old versions of the GTAP used partially different naming conventions etc. compared to the test data set produced by the center. The GTAP V9 release as available from the GTAP web site for users with a valid license was therefore as a test converted to the new structure:

1. For the parameters, that required some renames, which can be simply done with a GAMS script:

```
*
* --- get the parameters from default.prm and rename to new naming convention
*
parameter esubd,esubm,esubt,esubva,etrae,incpar,subpar,rorflex;
$GDXIN "default_prm.gdx"
  $$load esubd esubm esubt esubva etrae incpar subpar rorflex
$GDXIN
execute_unload "gtap_p.gdx" esubd=esbd,esubm=esbm,esubt=esbt,esubva=esbv,etrae=etre,incpar=incp,subpar=subp,rorflex=rflx;
```

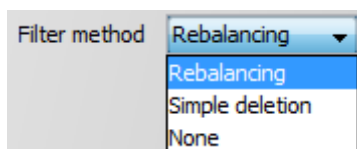
2. The core GTAP data format and names have not changed such that is sufficient to do a HAR2GDX with the new name “GTAP\_d.gdx”.
3. The new structure puts the sets in a separate container. GAMS code can easily accomplish that:

```
*
* --- the sets are in Angel's solution in a separate container:
*
set prod_comm,endw_comm,trad_comm,marg_comm,reg,tartype;
$GDXIN "gtap_d.gdx"
  $$load prod_comm endw_comm trad_comm marg_comm reg tartype
$GDXIN
execute_unload "gtap_s.gdx" prod_comm,endw_comm,trad_comm,marg_comm reg,tartype;
```

4. The structure of the aez data is different. The data had been distributed in two HAR files before, with different names and partly a different order of the indices. The GAMS code to map the old data into the new format looks like:

#### 4.3.6 Data base treatment after aggregation: Filtering or not

After aggregating the data to the desired region, product, sector and factor detail but before assigning it to a global SAM, you can use a facility to additionally filter out small values. Three options are available:



Choosing the option “None” will simply load the data base as it is, i.e., without filtering, which is the way the GEMPACK version operates. The following short section discusses the two other options. The recommended option is “Rebalancing” which filters out small values from the data base.

#### 4.3.7 Simple deletion

With simple deletion, transactions loaded from the different GDX container are removed from the data base if they are in absolute terms below the threshold entered on the interface under “Absolute tolerance”. With a value of 1.E-10, that deletion step is skipped. Afterwards, the transactions are formatted into a SAM structure. The resulting SAM is then cleansed with the chosen absolute tolerance. No attempt is made with “simple deletion” to maintain the resulting SAM balanced. That option is mostly maintained as a fall back, in case the more refined rebalancing step normally

recommended and discussed next should not work. Please note that there is no guarantee that a global SAM “cleansed” by simple delete will work with the model as such cleansing can lead not only to numerically, but also logically inconsistent data. Imagine the case where a tax rate is above 100%: cleansing might delete the tax base such as primary factor use in a sector, but keep the tax flow.

#### 4.3.8 Filtering and Rebalancing

##### Background

The “Rebalancing” option uses more advanced tactics to select transactions to delete, and perhaps more importantly, rebalances the resulting global SAM. As with simple deletion, first transactions in the original data base as loaded from the GDX containers which in absolute term are below the chosen absolute threshold are removed from the data base. With a value of 1.E-10, this preliminary deletion step is skipped. It is generally not recommended to use absolute deletion thresholds above 1.E-6 in combination with rebalancing as the subsequent relative thresholds will anyhow apply more refined rules. If available, it is recommended to use dedicated LP solvers such as CPLEX for the rebalancing step. For a discussion of the solver choice, see Britz, W. (2021): Comparing Penalty Functions in Balancing and Dis-aggregating Social Accounting Matrices, *Journal of Global Economic Analysis* 6(1): 34-81.

The consecutively applied filtering and rebalancing approach is an extension of the method and code developed by Tom Rutherford for “GTAPinGAMS” (see e.g. <http://www.mpsge.org/gtap6/>). It deletes component of the SAM depending on their shares on specific totals, according to the “Relative tolerance” entered on the interface:

- Domestic and imported intermediate demand of a commodity are dropped relative to its total output
- Private / government / investment domestic respectively import demand of a commodity are dropped relative to total private / government / investment domestic respectively import demand
- Trade flows of a product are dropped if both shares on total exports of that product and its exporter and on imports of that product and its importer are below the relative threshold
- Production is dropped if net production of a commodity, i.e. after intermediate use of that commodity in its own production is deducted, is below the relative threshold with regard to total net production

The filtering process imposes restrictions which should maintain the regional SAMs balanced. Additional constraints ensure that production activities require added value and intermediate inputs, if not already otherwise found in the data base.

##### User options during filtering

As filtering systematically removes elements from the SAM and the trade matrices, the process implies without further corrections shrinking the economies. During rebalancing, the algorithm can therefore add penalties for deviations from the following aggregate transactions:

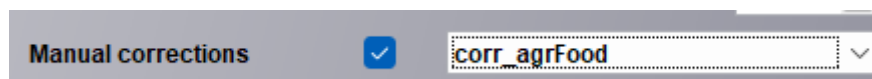
|                            |                          |                      |
|----------------------------|--------------------------|----------------------|
| Penalty for deviation from | GDP                      | <input type="text"/> |
|                            | Intermediate consumption |                      |
|                            | Factor income            |                      |

By adding these penalties terms, the non-deleted entries (and thus the more important transactions according to their value) tend to be scaled upwards. It is generally recommended to use these penalties terms. The code will also scale all non-deleted trade flows to approximately maintain the total volume of international trade and related international transport margins.

A further option allows removing non-diagonal intermediate crop use (with the exemption of the other crop product *ocr* which comprises seed production) and to control for the maximal seed cost share:

A screenshot of a software interface showing a checkbox labeled 'Only allow for diagonal crop by crop intermediate use' which is checked. To its right is a label 'Maximal seed cost share' followed by a numeric input field containing the value '0.5'.

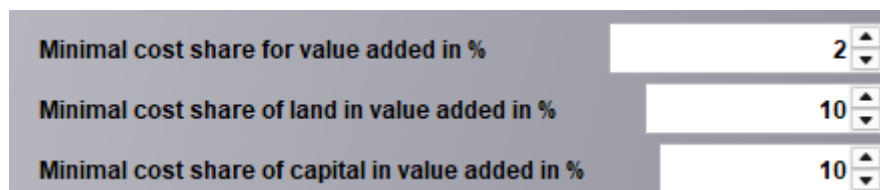
The user can add its own file with manual corrections to the raw data:

A screenshot of a software interface showing a checkbox labeled 'Manual corrections' which is checked. To its right is a dropdown menu with the text 'corr\_agrFood' selected.

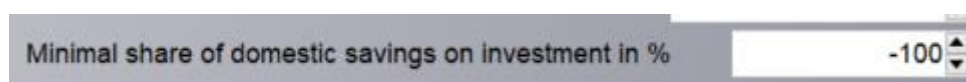
And use a statistical approach that uses the distribution of intermediate cost shares in the data base to remove potential outliers:

A screenshot of a software interface showing a checkbox labeled 'Outlier control for cost shares' which is unchecked.

Finally, a minimum cost share for the value added for each activity can be imposed. And minimal cost share of land and capital on value-added. Imposing these lower bounds can help in latter applications, for instance, when endogenous productivity are used to target land use developments.

A screenshot of a software interface showing three rows of controls. Each row has a label on the left and a numeric input field on the right. The first row is 'Minimal cost share for value added in %' with a value of '2'. The second row is 'Minimal cost share of land in value added in %' with a value of '10'. The third row is 'Minimal cost share of capital in value added in %' with a value of '10'.

And, a minimal share of domestic savings on investments can be imputed:

A screenshot of a software interface showing a label 'Minimal share of domestic savings on investment in %' followed by a numeric input field containing the value '-100'.

During filtering, the absolute and relative thresholds can be stepwise enforced. For the first few steps, exponential increases are used, starting with minus half the number of steps. For six steps, to give an example, the first thresholds applied will be the 1.E-3 of the final one, next 1.E-2 and finally 10%. The remaining steps will use equal linear increases between 10% and the desired final ones. Once the final thresholds are active, filtering is still applied several times until no small values are found any longer. The code should ensure that the resulting transactions are still fully consistent with each other, i.e. both the resulting trade matrices and the SAMs are balanced. The changes imposed by filtering and subsequent balancing are stored on the “*itrlog*” symbol in the GDX container with the final results. Inspecting how the stepwise enforcement of the thresholds impacts on the number of non-zero items can inform on an appropriate level for tolerances to be used.

The SAMs used during filtering are – as in GTAPAGG – defined in Million constant dollars. An absolute threshold of 1.E-6 will hence delete any economic transactions worth a single dollar or less. In SAMs with high regional and sectoral detail, even such tiny transactions might make up to 10% of

the non-zero entries. Increasing the threshold to 1000 \$ might remove  $\frac{1}{4}$  or more of all non-empty transactions. Similar results are found from using relative tolerances of 0.001 %.

Thanks to balancing, also rather dis-aggregated versions of the model with large number of sectors and regions can be used. The biggest impact of the filtering is typically on transactions related government demand. Here, often 50% or more of the flows account for only 1% of the total value of these transactions. Thus, tiny changes in the relative tolerance can have a considerable impact on the number of deleted transactions, and one might need to experiment with settings in the range around 1.E-1 to 1.E-4 to find a compromise between sparsity and the implied changes on the structure of the economies. For very large data sets (e.g. a 1:1 version) larger filtering thresholds such as 0.001 % might be needed to yield reasonable model sizes. The user can additionally define a minimum number of transactions to be kept, which reduces the need to experiment with different thresholds as the filtering process will stop once less than the desired number of transactions is reached. Tests with the model have shown that the model can be solved in full resolution of the GTAP 11c data base with moderate filtering, i.e. 65 sectors and 160 regions, in partial trade liberalization scenarios, even with an advanced set-up (GTAP-AEZ, GTAP-E, evolved nesting structures). But such runs require quite some RAM (30 GB and more) and users should consider their hardware before such tests to avoid damage. A closer look at the filtering statistics is recommended, to avoid sharp impacts on the structure of the economy. A more detailed discussion on the relation between model dis-aggregation, filtering, solution behavior and simulated welfare impacts provide Britz et al. 2015. However, the reported solution times are not certainly larger than what one would find today for a modern laptop.

### Structure of the filtering and balancing program

The basic structure of the programs consists of the following steps (see “*gtapAgg/filter.gms*”)

1. Store some original data, derive ad-valorem tax rates and calculate macro-totals to target (such as GDP, factor income).
2. In a loop
  - a. Delete small bi-lateral trade flows and define from their fixed total imports and exports by product for each model region as well as their international trade margin supply.
  - b. Flag other regional small transactions (intermediate and final demands, factor demands, output values) for deletion by setting their targeted value to zero.
  - c. Run a balancing program, separate for each model region (see next section)
  - d. If there is almost no change in the number of non-zero targeted items or if the maximum number of iterations is reached, break, otherwise, repeat a-c
3. Report statistics, such as number of removed transactions and changes in macro-totals and define tax-inclusive transactions

### Equations in the re-balancing problem

The balancing program is found in “*gtapAgg/filter\_model.gms*” and derived from the code by Tom Rutherford. To improve readability, endogenous variables are pre-fixed with *v\_*. As international transactions are fixed, the re-balancing problem can be applied separately to the model regions such that a relative compact QP (an optimization problem with a quadratic objective and linear constraints) problem is solved for each model region, usually in parallel on a grid. The problem comprises inequalities that are hard to implement in not optimization based balancing approaches such as RAS.

Two equations ensure that the endogenous total domestic use for each product  $v\_vdm$  plus fixed exports  $vxm$  is equal to the production output of the related activities (according to set *make*), and that the sum of production outputs  $v\_makem$  by activity, considering fixed ad-valorem output taxes  $txo$ , is equal to production cost. The production costs consist of endogenous imported  $v\_vifm$  and domestic  $v\_vdfm$  intermediate demands, considering exogenous use tax rates  $rtfi$  respectively  $rtfd$ , and endogenous factor demand  $v\_vfm$ , considering exogenous factor tax rates  $rtf$ :

```
*
* --- domestic market supply
*
e_make(i,rb(rNat)) $ sum(make(a,i),vom0(a,rNat)) ..
    sum(make(a,i) $ makem0(a,i,rNat),v_makem(a,i,rNat))
    =E= v_vdm(i,rNat) + vxm(i,rNat);
*
* --- revenue exhaustion (output value = production costs)
*
e_profit(a,rb(rNat)) $ vom0(a,rNat) ..
    sum(make(a,i) $ makem0(a,i,rNat),v_makem(a,i,rNat)) * (1-rto(a,rNat))
    =E=
        sum(j,
            v_vifm(j,a,rNat)*(1+rtfi(j,a,rNat)) $ vifm0(j,a,rNat)
            + v_vdfm(j,a,rNat)*(1+rtfd(j,a,rNat)) $ vdfm0(j,a,rNat))
        + sum(f, v_vfm(f,a,rNat)*(1+rtf(f,a,rNat)) $ vfm0(f,a,rNat));
```

The parameters ending with 0 such as  $vfm0$  depict the original, non-filtered transactions. Products are labelled with  $i,j$ , sectors with  $a$ , factors with  $f$ , model regions with  $rNat$ . The dynamic set  $rb$  depicts the model region currently balanced.

The following two equations ensure that factor use and intermediate input use cannot drop to zero as long as there is production, using the “BigM” approach found in integer programming.

```
*
*
e_fe(f,aIn(a),rb(rNat)) $ (
    (vfm0(f,a,rNat)*bigM>vom0(a,rNat))
    and vom(a,rNat) and vom0(a,rNat) and vfm0(f,a,rNat) and bigM)..
    sum(make(a,iIn(i)) $ makem0(a,i,rNat),v_makem(a,i,rNat) * (1-rto(i,a,rNat)))/BIGM
    =L= v_vfm(f,a,rNat)*(1+rtf(f,a,rNat))
    /sum(make(a,i),max(makem(a,i,rNat),makem0(a,i,rNat)));
*
* --- make sure that we don't have a intermediate demand if there is no production
*
e_nd(aIn(a),rb(rNat)) $ ( vom(a,rNat) and vom0(a,rNat) and bigM )..
    sum(make(a,iIn(i)) $ makem0(a,i,rNat),v_makem(a,i,rNat) * (1-rto(i,a,rNat)))/BIGM
    =L= v_nd(a,rNat) /sum(make(a,iIn(i)),max(makem(a,i,rNat),makem0(a,i,rNat)));
*
```

Three additional equations ensure a minimal cost share for value-added. Capital and labor; the related thresholds can be set by the user on the interface:

```
*
* --- minimal cost shares for va, capital and land
*
e_vaMin(aIn(a),rb(rNat)) $ ( vom(a,rNat) and vom0(a,rNat) and %minCostShareVA% )..
    sum(make(a,iIn(i)) $ makem0(a,i,rNat),v_makem(a,i,rNat) * (1-rto(i,a,rNat))) * (%minCostShareVA%/100)
    =L= v_va(a,rNat);
e_capMin(aIn(a),rb(rNat)) $ ( sum(capF,vfm0(capF,a,rNat)*(1+rtf(capF,a,rNat)))
    and vom(a,rNat) and vom0(a,rNat) and %minCostShareCap% )..
    v_va(a,rNat) * (%minCostShareCap%/100)
    =L= sum(capF $ vfm0(capF,a,rNat), v_vfm(capF,a,rNat)*(1+rtf(capF,a,rNat)));
e_lndMin(aIn(a),rb(rNat)) $ ( sum(lnd,vfm0(lnd,a,rNat)*(1+rtf(lnd,a,rNat))) $ (not excludeLndMin(a))
    and vom(a,rNat) and vom0(a,rNat) and %minCostShareLnd% )..
    v_va(a,rNat) * (%minCostShareLnd%/100)
    =L= sum(lnd $ vfm0(lnd,a,rNat), v_vfm(lnd,a,rNat)*(1+rtf(lnd,a,rNat)));
*
```

The maximal seed cost share  $\%maxSeedCostShare\%$  (default 50%), inputted via the GUI by the user, is controlled by the following inequality constraint  $e\_seed$ . It checks for intermediate input use of crop outputs  $cropsI$  produced by the activity  $a$  according to the make set:

```
*
* --- seed cost (= own use of crop output) cannot exceed a certain share of the production value
*
e_seed(a,rb(rNat)) $ ((vom(a,rNat) and vom0(a,rNat)) $ sum(make(a,cropsI,1)) ..
    sum(make(a,cropsI) $ makem0(a,CropsI,rNat),v_makem(a,cropsI,rNat) ) * %maxSeedCostShare% * (1-rto(a,rNat))
    =G= sum(make(a,cropsI), v_vifm(cropsI,a,rNat)*(1+rtfi(cropsI,a,rNat)) $ vifm0(cropsI,a,rNat)
        + v_vdfm(cropsI,a,rNat)*(1+rtfd(cropsI,a,rNat)) $ vdfm0(cropsI,a,rNat));
```

The following two equations impose upper and lower bounds on the intermediate cost shares. They are introduced if the outlier detection is active:

```
e_IntCorrUp(iIn(i),aIn(a),rb(rNat)) $ ( (vom(a,rNat) and vom0(a,rNat)) $ (not exempt(i,a))
    $ ( outliers(rNat,i,a) or
        ([ vifm(i,a,rNat)*(1+rtfi(i,a,rNat)) $ vifm0(i,a,rNat)
          + vdfm(i,a,rNat)*(1+rtfd(i,a,rNat)) $ vdfm0(i,a,rNat)]
          gt 0.001*sum(make(a,j) $ makem0(a,j,rNat),makem(a,j,rNat)))) ..

    sum(make(a,j) $ makem0(a,j,rNat),v_makem(a,j,rNat))
    * [p_stats(i,a,"avg") + 2.5*p_stats(i,a,"cv")*p_stats(i,a,"avg")]/100
    + v_slackIntCorrP(i,a,rNat)*sum(make(a,j),max(makem(a,j,rNat),makem0(a,j,rNat)))

    =G=
        v_vifm(i,a,rNat)*(1+rtfi(i,a,rNat)) $ vifm0(i,a,rNat)
        + v_vdfm(i,a,rNat)*(1+rtfd(i,a,rNat)) $ vdfm0(i,a,rNat);

e_IntCorrLo(iIn(i),aIn(a),rb(rNat)) $ ( (vom(a,rNat) and vom0(a,rNat)) $ (not exempt(i,a))
    $ ( outliers(rNat,i,a) or
        ([ vifm(i,a,rNat)*(1+rtfi(i,a,rNat)) $ vifm0(i,a,rNat)
          + vdfm(i,a,rNat)*(1+rtfd(i,a,rNat)) $ vdfm0(i,a,rNat)]
          gt 0.001*sum(make(a,j) $ makem0(a,j,rNat),makem(a,j,rNat)))) ..

    sum(make(a,j) $ makem0(a,j,rNat),v_makem(a,j,rNat))
    * [p_stats(i,a,"avg") - 2.5*p_stats(i,a,"cv")*p_stats(i,a,"avg")]/100
    - v_slackIntCorrN(i,a,rNat)*sum(make(a,j),max(makem(a,j,rNat),makem0(a,j,rNat)))

    =L=
        v_vifm(i,a,rNat)*(1+rtfi(i,a,rNat)) $ vifm0(i,a,rNat)
        + v_vdfm(i,a,rNat)*(1+rtfd(i,a,rNat)) $ vdfm0(i,a,rNat);
```

The domestic market balance in the equation  $e\_dommkt$  below ensures that domestic sales  $v\_vdm$  for each product  $i$  are exhausted by demands for the domestic origin of the different Armington agents, i.e. equal to the sum of government demand  $v\_vdgm$ , private demand  $v\_vdpm$ , investment demand  $v\_vdim$  and the summed up intermediate demands of the sectors  $v\_vdfm$ :

```
*
* --- domestic market
*
e_dommkt(i,rb(rNat)) $ vdm(i,rNat) ..

v_vdm(i,rNat)
    =E=
        v_vdgm(i,rNat) $ vdgm0(i,rNat)
        + v_vdpm(i,rNat) $ vdpm0(i,rNat)
        + v_vdim(i,rNat) $ vdim0(i,rNat)
        + sum(a $ vdfm0(i,a,rNat), v_vdfm(i,a,rNat)) ;
```

A similar equation for the import market  $e\_impmkt$  ensures that the given total imports  $vim$  are exhausted by the demand of the imported origin of the different agents:

```
*
* --- import market (vim is fixed during solve as trade flows are exogenous)
*
e_impmkt(i,rb(rNat)) $ vim(i,rNat) ..

vim(i,rNat)
    =E=
        v_vigm(i,rNat) $ vigm0(i,rNat)
        + v_vipm(i,rNat) $ vipm0(i,rNat)
        + v_viim(i,rNat) $ viim0(i,rNat)
        + sum(a $ vifm0(i,a,rNat), v_vifm(i,a,rNat));
```

Thinning out the SAM by removing small transaction without further correction would imply that the size of each regional economy would shrink. In order to avoid that, the program defines relative changes  $v\_keepCor$  for various economic totals, defined from the current endogenous solution and the totals calculated from the raw data stored on the parameter  $p\_keepTarget$ . The relative changes are minimized in the objective along with other terms.

The first such constraint  $e\_intKeep$  defines this correction factor for total intermediate use, i.e. the sum of the imported  $v\_vifm$  and domestic use  $v\_dfm$  by firms, aggregated over all products and activities. The RHS calculates this sum divided by the original sum, the LHS derives from there the relative change in percentage term.

```
*
* --- try to maintain benchmark intermediate consumption
*
e_intKeep(rb(rNat)) ..
    (1 + v_keepCor("int",rNat)/100)
    =E= sum( (j,a) $ vom0(a,rNat), v_vifm(j,a,rNat) $ vifm0(j,a,rNat)
              + v_vdfm(j,a,rNat) $ vdfm0(j,a,rNat) )
          / keepTarget(rNat,"int");
```

Similarly, the relative changes in GDP at market prices are defined by the equation  $e\_gdpKeep$  where the LHS defines the sum of the final demand at agent prices plus the B.o.T:

```
*
* --- try to maintain at benchmark gdp
*
e_gdpKeep(rb(rNat)) ..
    (1 + v_keepCor("gdp",rNat)/100)
    =E= sum(i,
        v_vdim(i,rNat)*(1+rtid(i,rNat)) $ vdim0(i,rNat)
      + v_viim(i,rNat)*(1+rtii(i,rNat)) $ viim0(i,rNat)
      + v_vdpm(i,rNat)*(1+rtpd(i,rNat)) $ vdpm0(i,rNat)
      + v_vipm(i,rNat)*(1+rtpi(i,rNat)) $ vipm0(i,rNat)
      + v_vdgm(i,rNat)*(1+rtgd(i,rNat)) $ vdgm0(i,rNat)
      + v_vigm(i,rNat)*(1+rtgi(i,rNat)) $ vigm0(i,rNat)

      + sum( rNat1 $ vxmd0(i,rNat,rNat1), vxmd(i,rNat,rNat1)*(1-rtxs(i,rNat,rNat1)))
      - sum( rNat1 $ vxmd0(i,rNat1,rNat), viws(i,rNat1,rNat) * vxmd(i,rNat1,rNat)/vxmd0(i,rNat1,rNat) )
      + vst(i,rNat) )
    /keepTarget(rNat,"gdp");
```

Additional equations target factor income in total and by each factor, as well as factor cost as shown below:

```
*
* --- try to maintain benchmark factor income
*
e_factinc(rb(rNat)) ..
    (1 + v_keepCor("facty",rNat)/100)
    =E= sum((f,a) $ vfm0(f,a,rNat), v_vfm(f,a,rNat))/keepTarget(rNat,"facty");

*
* --- try to maintain benchmark factor cost
*
e_vfmKeep(f,rb(rNat)) $ sum(a, vfm0(f,a,rNat)*(1+rtf(f,a,rNat))) ..
    (1 + sum(sameas(keepItem,f),v_keepCor(keepItem,rNat))/100)
    =E= sum(a $ vfm0(f,a,rNat), v_vfm(f,a,rNat) * (1+rtf(f,a,rNat)))/keepTarget(rNat,f);

e_vfmTotKeep(rb(rNat)) ..
    (1 + v_keepCor("vfm",rNat)/100)
    =E= sum((f,a) $ vfm0(f,a,rNat), v_vfm(f,a,rNat) * (1+rtf(f,a,rNat)))/keepTarget(rNat,"vfm");
```

Similar statements define correction factors for total government demand, investment demand, private consumption, i.e. the components GDP. Note that the B.o.T. is fixed in the constrained optimization problem as all individual trade flows are exogenous. Targeting the B.o.T. as a further element in the GDP makes hence no sense.

```

*
* --- try to maintain benchmark private consumption total
*
e_yp(rb(rNat)) ..
    (1 + v_keepCor("yp",rNat)/100)
    =E= sum(i,      v_vdpm(i,rNat)*(1+rtpd(i,rNat)) $ vdpm0(i,rNat)
              + v_vipm(i,rNat)*(1+rtpi(i,rNat)) $ vipm0(i,rNat))
    /keepTarget(rnat,"yp");
*
* --- try to maintain benchmark gov consumption total
*
e_yg(rb(rNat)) ..
    (1 + v_keepCor("yg",rNat)/100)
    =E= sum(i,      v_vdgm(i,rNat)*(1+rtgd(i,rNat)) $ vdgm0(i,rNat)
              + v_vigim(i,rNat)*(1+rtgi(i,rNat)) $ vigim0(i,rNat))
    /keepTarget(rnat,"yg");
*
* --- try to maintain benchmark investments
*
e_vdimKeep(rb(rNat)) ..
    (1 + v_keepCor("yi",rNat)/100)
    =E= sum(i,      v_vdim(i,rNat)*(1+rtid(i,rNat)) $ vdim0(i,rNat)
              + v_viim(i,rNat)*(1+rtii(i,rNat)) $ viim0(i,rNat))
    /keepTarget(rnat,"yi");

```

Which macro totals are targeted can be defined on the interface by the user as shown above.

### Overall solution logic

The program (see *gtapAgg/filter.gms*) stepwise increases the threshold applied to define small transactions over iterations,

```

curAbsTol      = absTol * tolMultiplier(itr+1);
curRelTol(td,rNat) = relTolSpec(td,rNat) * TolMultiplier(itr+1);
curRelTol(td,rNat) = round(curRelTol(td,rNat),8);

```

At the beginning of each iteration, it re-defines the aggregate values of imports for each product  $i$  and model region  $rNat$ :

```

* --- Aggregate value of imports (from bi-lateral exports, plus transport margins, plus import tariffs)
vim(i,rNat) $ vim0(i,rNat)
= sum(rNat1 $ vxmd(i,rNat1,rNat),
      (vxmd(i,rNat1,rNat)*(1-rtxs(i,rNat1,rNat))+sum(j,vtwr(j,i,rNat1,rNat))*(1+rtms(i,rNat1,rNat))));

```

And rescales the transport margins demand by product  $i$  (also called mode) for each model region  $rNat$  such that their global sum is not changing:

```

* --- Rescale transport margin demands: the sum over vtwr is the total transport margin demand for product i ,
* distributed by the share of the region on glboal transport margin supply of i
vst(i,rNat) $ sum(rNat1, vst(i,rNat1)) = vst(i,rNat) * sum((j,rNat1,rrr), vtwr(i,j,rNat1,rrr)) / sum(rNat1, vst(i,rNat1));

```

Similarly, it calculates the value of  $vxm$  and domestic sales  $vdm$  by product  $i$  and production values  $vom$  for each model region  $rNat$ :



```

* --- Value of exports by product and nation: bi-lateral exports at market prices plus margin supply
vxm(i,rNat) = sum(rNat1, vxmd(i,rNat,rNat1)) + vst(i,rNat);

* --- Production for the domestic market (final and intermediate domestic demand)
vdm(i,rNat) = vdpm(i,rNat) + vdgm(i,rNat) + sum(a, vdfm(i,a,rNat)) + vdim(i,rNat);
*
* --- aggregate total demand per product
vim(i,rNat) = sum(a, vifm(i,a,rNat)) + vipm(i,rNat) + vigm(i,rNat) + viim(i,rNat) ;
*
* --- Aggregate production flag (!): domestic and export demand
* (attention: in case multiple processes output i, this comprises a double counting)
vom(a,rNat) $ vom0(a,rNat) = sum(make(a,i) $ makem0(a,i,rNat), vdm(i,rNat) + vxm(i,rNat));

```

Next, it flags product demands for deletion by setting them to zero, according to the current active relative thresholds *curReltol* and the transaction specific total, for instance using total government demand *tvgm* and the imported share *tvgmi*

```

*
* --- delete import and domestic gov demand for products
* if expenditure share on total imported/domestic gov demand is below threshold
*
vigmi(i,rNat) $ ((vigmi(i,rNat) < curRelTol(i,rNat)/100 * tvgm(rNat) * tvgmi(rNat)/tvgm(rNat)) $ vigm(i,rNat)) = 0;
vdgmi(i,rNat) $ ((vdgmi(i,rNat) < curRelTol(i,rNat)/100 * tvgm(rNat) * (1-tvgmi(rNat)/tvgm(rNat))) $ vdgm(i,rNat)) = 0;

```

After flagging the agent specific demands for deletion, total import and domestic demand for each product *i* are re-aggregated:

```

*
* --- aggregate total import and domestic demand per product
*
vim(i,rNat) = sum(a, vifm(i,a,rNat)) + vipm(i,rNat) + vigm(i,rNat) + viim(i,rNat);
vdm(i,rNat) = sum(a, vdfm(i,a,rNat)) + vdpm(i,rNat) + vdgm(i,rNat) + vdim(i,rNat);

```

Exports of a product *i* by region *rNat* are dropped if simultaneously:

- (1) the product's exports value *vxm* falls below the threshold times the exporters' total exports (summed over product *j*)
- (2) this export value is below the exporters outputs value for the product times the threshold, according to the make matrix value *makem*
- (3) the bilateral import value (export flows at f.o.b. plus transport margins plus import tariffs) of trade flows of any region *rNat1* importing this product *i* from exporter *rNat* is below the importers total import demand times the threshold for all importers *rNat1*.

```

*
* --- drop export of product if share on total exports and
* max. import share on total import demand of product at destination is below threshold
* and exports don't exceed the threshold in production
*
if ( relTol > 1.E-10,
dropexports(i,rNat) $ ( (vxm(i,rNat) < curRelTol(i,rNat)/100*sum(j,vxm(j,rNat))) and
(vxm(i,rNat) < curRelTol(i,rNat)/100*sum(make(a,i),makem(a,i,rNat))) and
(sum(rNat1 $
( ( (vxmd(i,rNat,rNat1)*(1-rtxs(i,rNat,rNat1))+sum(j,vtwr(j,i,rNat,rNat1)))*(1+rtms(i,rNat,rNat1)) /
vim(i,rNat1)) < curRelTol(i,rNat1)/100)
$ vim(i,rNat1),1) eq sum(rNat1 $ vim(i,rNat1),1) and vxm(i,rNat)) = YES;
);

```

If total exports are dropped for a country and product, all related bi-lateral transactions need to be dropped as well:

```

*
* --- delete bilateral trade flow entries if exports are dropped
*
vxmd(i,rNat,rNat1) $ dropexports(i,rNat) = 0;
vtwr(j,i,rNat,rNat1) $ dropexports(i,rNat) = 0;
rtxs(i,rNat,rNat1) $ dropexports(i,rNat) = 0;
rtms(i,rNat,rNat1) $ dropexports(i,rNat) = 0;

```

Imports for a product  $i$  and region  $rNat$  are dropped if simultaneously

- (1) Their value is below the total import value (summed over products  $j$ ) times the threshold
- (2) The value of the bi-lateral exports defining these imports for each exporter  $rNat1$  is below the total out value of this product

```

--- drop import of product if share on total imports and
    max. export share on output at destination is below threshold

if ( (relTol > 1.E-10) or (absTol > 1.E-10),
    dropimports(i,rNat)$( (vim(i,rNat) < curRelTol(i,rNat)/100*sum(j,vim(j,rNat))) and
        (sum(rNat1 $ ( (vxmd(i,rNat1,rNat)/sum(make(a,i),vom(a,rNat1)) < curRelTol(i,rNat1)/100)
            $ sum(make(a,i),vom(a,rNat1))),1) eq sum((rNat1,make(a,i)) $ vom(a,rNat1),1)) and vim(i,rNat) = yes;
);
if ( abstol ne 1.E-10, dropexports(i,rNat) $ (vim(i,rNat) le curAbsTol) = YES);

```

Imports are also dropped if no import demand of any Armington agent is left. Excluded are cases without observed imports at the base or when imports are fixed.

```

*
* --- drop imports if no import demand
*
dropimports(i,rNat) $ ( (vigm(i,rNat)+vipm(i,rNat)+viim(i,rNat)+sum(a,vifm(i,a,rNat))) eq 0 ) = yes;
dropimports(i,rNat) $ (not (vigm0(i,rNat)+vipm0(i,rNat)+viim0(i,rNat)+sum(a,vifm0(i,a,rNat)))) = no;
dropimports(i,rNat) $ (not v_vdm.range(i,rNat))= no;

vim(i,rNat) $ dropimports(i,rNat) = 0;

```

Equally, production by activity  $a$  is dropped if the share of its output values stored on the make matrix *makem* – corrected for own intermediate consumption such as seed use – is below the share of economy-wide total production value (summation over the alias *aCur* for activities), corrected for own intermediate consumption:

```

*
* --- drop production if share of output net of own intermediate consumption
*   is below threshold on total output net of own intermediate consumption
*   and share on total factor use is below threshold
*
if ( (relTol > 1.E-10),
    dropprod(a,rNat) $ ( (sum(make(a,i),makem(a,i,rNat)-vdfm(i,a,rNat))<sum(make(a,i),curRelTol(i,rNat)/100
        *sum(make(aCur,j),makem(aCur,j,rNat)-vdfm(j,aCur,rNat)))
        $ vom0(a,rNat))= yes;
    if (card(dropProd), display dropProd);
);

```

Equally, production is dropped if no domestic plus export demand is left or if all intermediate and factor demand of the activity is zero:

```

*
* --- drop production if no domestic and export demand
*
dropprod(a,rNat) $ ( [sum(make(a,i) $ makem0(a,i,rNat), (vxm(i,rNat)+vdgm(i,rNat)+vdim(i,rNat)+vdpm(i,rNat)
    +sum(aCur $ (not sameas(aCur,a)),vdfm(i,aCur,rNat))) eq 0] $ vom0(a,rNat)) = yes;

dropprod(a,rNat) $ ((not dropprod(a,rNat)) $ vom0(a,rNat))
    = yes $ [(sum(f, vfm(f,a,rNat))
        +sum(j $ (not make(a,j)),vdfm(j,a,rNat)+vifm(j,a,rNat))) eq 0];

```

Some of these statements are repeated in a loop to ensure that all removals are consistent, not shown here.

Afterwards, bi-lateral flows are subject to a drop if their value on total imports is below the threshold. The code snippet also shows the related statement for absolute threshold checks.

```

*
* --- drop bilateral trade flow if value net of trade margins and taxes is below
* thresholds both of exports and imports
*
if ( relTol > 1.E-10,
    droptrade(i,rNat,rNat1) $ vxmd(i,rNat,rNat1)
        = yes $ (vxmd(i,rNat,rNat1)<min(curRelTol(i,rNat1),curRelTol(i,rNat))/100 * min( vxm(i,rNat), vim(i,rNat1))));
);
if ( absTol > 1.E-10,
    droptrade(i,rNat,rNat1) $ vxmd(i,rNat,rNat1) = yes $ (vxmd(i,rNat,rNat1) le curAbsTol);
);

```

If a trade flow is dropped, all related bi-lateral transactions have to be removed:

```

*
* --- if dropped, also delete related trade margin and export/import taxes
*
vxmd(i,rNat,rNat1) $ droptrade(i,rNat,rNat1) = 0;
vtwr(j,i,rNat,rNat1) $ droptrade(i,rNat,rNat1) = 0;
rtxs(i,rNat,rNat1) $ droptrade(i,rNat,rNat1) = 0;
rtms(i,rNat,rNat1) $ droptrade(i,rNat,rNat1) = 0;

```

In order to make sure that the bi-lateral trade balances are approximately maintained, the non-dropped trade flows are scaled upwards. The scaling factor for each pair of trading partners ( $rNat, rNat1$ ) is the average of the scaling factor for the exporters  $rNat$  original exports (postfix 0) over the current exports and the scaling factor for the importers  $rNat1$  original imports over the targeted imports. Imports and exports are summed up over all products. The resulting bi-lateral scaler is applied to the exports at market prices and the trade margins.

```

*
* --- aggregate value of imports and exports after dropping trade
*
vxm(i,rNat) = sum(rNat1, vxmd(i,rNat,rNat1)) + vst(i,rNat);
vim(i,rNat) = sum(rNat1, (vxmd(i,rNat1,rNat)*(1-rtxs(i,rNat1,rNat))+sum(j,vtwr(j,i,rNat1,rNat)))*(1+rtms(i,rNat1,rNat))));
*
* --- rescale vxmd to match approx. old exports of exporter and imports of importer
*
bilatTradeScale(rNat,rNat1) = ( [sum(j,vxm0(j,rNat))/sum(j,vxm(j,rNat))] $ sum(j,vxm(j,rNat))
    + [sum(j,vim0(j,rNat1))/sum(j,vim(j,rNat1))] $ sum(j,vim(j,rNat1)) )/2;

vxmd(i,rNat,rNat1) $ vxmd(i,rNat,rNat1) = vxmd(i,rNat,rNat1) * bilatTradeScale(rNat,rNat1);
vtwr(j,i,rNat,rNat1) $ vtwr(j,i,rNat,rNat1) = vtwr(j,i,rNat,rNat1) * bilatTradeScale(rNat,rNat1);

```

After these corrections, a global scaler is applied to ensure that global trade for each product  $i$  does not change too much. The scaling uses the value of exports at market prices.

```

*
* --- rescale global trade by product (up to 20%)
*
vxt(i) = sum( (rNat,rNat1), vxmd(i,rNat,rNat1));
vxt(i) $ vxt(i) = min(1.2,vxt0(i)/vxt(i));
vxmd(i,rNat,rNat1) $ vxmd(i,rNat,rNat1) = vxmd(i,rNat,rNat1) * vxt(i);
*
* --- apply same scaler to trade margins
*
vtwr(j,i,rNat,rNat1) $ vtwr(j,i,rNat,rNat1) = vtwr(j,i,rNat,rNat1) * vxt(i);

```

Similarly, also import and export tax rates are scaled to maintain total tariff income and export revenues for each country:

```

*
* --- rescale imports and import tax revenues
*
total(rNat,"new") = sum( (i,rNat1), (vxmd(i,rNat1,rNat)*(1-rtxs(i,rNat1,rNat))+sum(j,vtwr(j,i,rNat1,rNat)))*(rtms(i,rNat1,rNat)));
rtms(i,rNat1,rNat) $ rtms(i,rNat1,rNat) = rtms(i,rNat1,rNat) * rtms0(rNat) / total(rNat,"new");

vim(i,rNat) = sum(rNat1, vxmd(i,rNat1,rNat)*(1-rtxs(i,rNat1,rNat))+sum(j,vtwr(j,i,rNat1,rNat)))*(1+rtms(i,rNat1,rNat)));
vim(i,rNat) $ (sum(rNat1,vxmd(i,rNat1,rNat)) eq 0) = 0;
vipm(i,rNat) $ (vim(i,rNat) eq 0) = 0;
vigm(i,rNat) $ (vim(i,rNat) eq 0) = 0;
viim(i,rNat) $ (vim(i,rNat) eq 0) = 0;
vifm(i,a,rNat) $ (vim(i,rNat) eq 0) = 0;

```

The transport margin demand for each product  $i$  and model region  $rNat$  is afterwards scaled to maintain the global sum, and from there, the total export demand is defined that is fixed in the subsequent application of the re-balancing problem for each model region:

```
* --- Rescale transport demand:

vst(i,rNat) $ vst(i,rNat) = vst(i,rNat) * sum((j,rNat1,rrr), vtwr(i,j,rNat1,rrr)) / sum(rNat1, vst(i,rNat1));
vxm(i,rNat) = sum(rNat1, vxmd(i,rNat,rNat1)) + vst(i,rNat);
```

The aggregate product flag *vom* is updated according to the domestic and export demands by product *i* and the set *make*:

```
* --- Aggregate production flag:

vom(a,rNat) $ vom(a,rNat) = sum(make(a,i) $ makem0(a,i,rNat),vdm(i,rNat) + vxm(i,rNat));
```

The resulting different target values for the various transactions are used as starting points for endogenous variables in the next solve of the rebalancing problem:

```
* --- start values are equal to targets
*
*
v_vdm.l(i,rNat) = vdm(i,rNat);
v_vfm.l(f,a,rNat) = vfm(f,a,rNat);
v_vifm.l(i,a,rNat) = vifm(i,a,rNat);
v_vdfm.l(i,a,rNat) = vdfm(i,a,rNat);
v_vdgm.l(i,rNat) = vdgm(i,rNat);
v_vdpm.l(i,rNat) = vdpm(i,rNat);
v_vigm.l(i,rNat) = vigm(i,rNat);
v_vipm.l(i,rNat) = vipm(i,rNat);
v_vdim.l(i,rNat) = vdim(i,rNat);
v_viim.l(i,rNat) = viim(i,rNat);
```

Afterwards, each country is solved individually to rebalance the SAM. That is possible as bilateral transactions are fixed in the equation system. In order to provide a fast and accurate solution to the filtering problem, normalized absolute deviation are minimized. A block of equation therefore defines the positive and negative penalty terms entering the objective function:

```
--- define positive and negative deviations between solution variables such as v_vdpm and the target vdpm

e_vdpm(iIn(i),rb(rNat)) $ vdpm0(i,rNat) .. v_vdpm(i,rNat) =E= vdpm(i,rNat) - v_corrN(i,"", "vdpm",rNat) + v_corrP(i,"", "vdpm",rNat);
e_vipm(iIn(i),rb(rNat)) $ vipm0(i,rNat) .. v_vipm(i,rNat) =E= vipm(i,rNat) - v_corrN(i,"", "vipm",rNat) + v_corrP(i,"", "vipm",rNat);
e_vdgm(iIn(i),rb(rNat)) $ vdgm0(i,rNat) .. v_vdgm(i,rNat) =E= vdgm(i,rNat) - v_corrN(i,"", "vdgm",rNat) + v_corrP(i,"", "vdgm",rNat);
e_vigm(iIn(i),rb(rNat)) $ vigm0(i,rNat) .. v_vigm(i,rNat) =E= vigm(i,rNat) - v_corrN(i,"", "vigm",rNat) + v_corrP(i,"", "vigm",rNat);
e_vdim(iIn(i),rb(rNat)) $ vdim0(i,rNat) .. v_vdim(i,rNat) =E= vdim(i,rNat) - v_corrN(i,"", "vdim",rNat) + v_corrP(i,"", "vdim",rNat);
e_viim(iIn(i),rb(rNat)) $ viim0(i,rNat) .. v_viim(i,rNat) =E= viim(i,rNat) - v_corrN(i,"", "viim",rNat) + v_corrP(i,"", "viim",rNat);

e_vdfm(iIn(i),ain(a),rb(rNat)) $ vdfm0(i,a,rNat) .. v_vdfm(i,a,rNat) =E= vdfm(i,a,rNat) - v_corrN(i,a, "vdfm",rNat) + v_corrP(i,a, "vdfm",rNat);
e_vifm(iIn(i),ain(a),rb(rNat)) $ vifm0(i,a,rNat) .. v_vifm(i,a,rNat) =E= vifm(i,a,rNat) - v_corrN(i,a, "vifm",rNat) + v_corrP(i,a, "vifm",rNat);
e_vfm(f,aIn(a),rb(rNat)) $ vfm0(f,a,rNat) .. v_vfm(f,a,rNat) =E= vfm(f,a,rNat) - v_corrN(f,a, "vfm",rNat) + v_corrP(f,a, "vfm",rNat);
e_makem(aIn(a),iIn(i),rb(rNat)) $ makem0(a,i,rNat) .. v_makem(a,i,rNat) =E= makem(a,i,rNat) - v_corrN(a,i, "makem",rNat) + v_corrP(a,i, "makem",rNat);
```

In the above term, *v\_vdpm* is the endogenous variable, pulled to its target value *vdpm* as defined the statement above. The objective function minimizes the sum of these terms, normalized with the maximum of the targeted transaction and a minimum defined from size of the economy as shown for the first type of transactions below:

```
e_objeR(rb(rNat)) ..
objeR(rNat)/100
=e=

--- the first part of the objective relates to normalized absolute deviations. The normalization factor is the maximum of
of the target, the original target value and a minimum scaler that 1E-6 times GDP

sum(iIn(i),
((v_corrN(i,"", "vdpm",rNat)+v_corrP(i,"", "vdpm",rNat))/(max(vdpm0(i,rNat),vdpm(i,rNat))+minScale(rNat))) $ vdpm0(i,rNat)
```

In order to support sparsity, a second set of terms pulls transactions that are flagged for deletion to zero:

```
v_vdpm(i,rNat) $ ((vdpm(i,rNat) eq 0) $ vdpm0(i,rNat))
```

These sparsity terms are added up and enter the objective with a weight of 1,000. Adding these sparsity terms is useful if we consider a case of some tiny endogenous result of 1.E-8 for a transaction

flagged for deletion (targeted value is zero). The penalty term will hence be  $\text{abs}(1.E-8 - 0)/(\text{original value plus minimum transaction thresholds})$ , a quite small number.

In order to speed up processing, the inear problems can be solved on a grid. There are more technical details in the code not discussed in here. The rebalancing step produces output on screen which might be interesting to check. The following screen-shots show results from filtering out from a 57x20 data base with a maximal threshold of 0.001%:

The first block reports the global totals, these deviations should be small, in here, the maximal absolute correction in sum over all transactions appeared for domestic private consumption vdpm where around 974 Mio USD out of 35,035,323 Mio USD reported originally are lost:

```
-- loadGTAPAgg.gms (3289) 84 Mb
  item          start      end      abs diff  rel diff
-- loadGTAPAgg.gms (3290) 84 Mb
  vfm           58750721   58749829      -892      0.00 %
  GDP           68168524   68168548        24      0.00 %
  vom           141872873  141872304     -569      0.00 %
  vxmd          19129902   19129902         0      0.00 %
  vtwr           776031    776031         0      0.00 %
  vifm          13747932   13747675     -257      0.00 %
  vdfm          60683808   60683471     -337      0.00 %
  vipm          4313228    4313725       497      0.01 %
  vdpm          35035323   35034349     -974      0.00 %
  vigm          222442     222426       -16     -0.01 %
  vdgm          12306107   12306312       205      0.00 %
  viim          2349721    2349497     -224     -0.01 %
  vdim          13941703   13942239       536      0.00 %
-- loadGTAPAgg.gms (3298) 84 Mb
```

The second block shows the reduction in number of the non-zero SAM entries. For the given example, despite the rather small threshold, the non-zero SAM entries as defined in the program could be reduced by almost a quarter (-23.27% in total), clearly with some differences across countries:

| loadGTAPAgg.gms (3302) 84 Mb |        |        |          |          |
|------------------------------|--------|--------|----------|----------|
| item                         | start  | end    | abs diff | rel diff |
| loadGTAPAgg.gms (3303) 84 Mb |        |        |          |          |
| total                        | 216819 | 166359 | -50460   | -23.27 % |
| Oceania                      | 11915  | 9940   | -1975    | -16.58 % |
| Chn                          | 11540  | 9398   | -2142    | -18.56 % |
| Jpn                          | 11528  | 8172   | -3356    | -29.11 % |
| Kor                          | 11429  | 7859   | -3570    | -31.24 % |
| Tw                           | 11438  | 6877   | -4561    | -39.88 % |
| HKG                          | 11453  | 7363   | -4090    | -35.71 % |
| EastAsia                     | 11120  | 8754   | -2366    | -21.28 % |
| SEAsia                       | 11954  | 9936   | -2018    | -16.88 % |
| SouthAsia                    | 11600  | 9349   | -2251    | -19.41 % |
| Usa                          | 11612  | 8732   | -2880    | -24.80 % |
| Can                          | 11528  | 9682   | -1846    | -16.01 % |
| Mexx                         | 11345  | 8834   | -2511    | -22.13 % |
| Arg                          | 11171  | 7241   | -3930    | -35.18 % |
| Bra                          | 11312  | 8417   | -2895    | -25.59 % |
| LatinAmer                    | 11729  | 10638  | -1091    | -9.30 %  |
| Deu                          | 11513  | 9044   | -2469    | -21.45 % |
| EU_28                        | 11786  | 10482  | -1304    | -11.06 % |
| MENA                         | 11951  | 10839  | -1112    | -9.30 %  |
| Ken                          | 10958  | 7389   | -3569    | -32.57 % |
| SSA                          | 12002  | 10598  | -1404    | -11.70 % |
| RestofWorld                  | 11849  | 10876  | -973     | -8.21 %  |

A second list reports the changes for each product:

|     |      |      |       |          |
|-----|------|------|-------|----------|
| pdr | 3735 | 1855 | -1880 | -50.33 % |
| wht | 3816 | 2227 | -1589 | -41.64 % |
| gro | 4011 | 2722 | -1289 | -32.14 % |
| v_f | 4260 | 3208 | -1052 | -24.69 % |
| osd | 4134 | 2626 | -1508 | -36.48 % |
| c_b | 3168 | 1074 | -2094 | -66.10 % |
| pfb | 3921 | 2026 | -1895 | -48.33 % |
| ocr | 4311 | 3366 | -945  | -21.92 % |
| ctl | 3897 | 2476 | -1421 | -36.46 % |
| oap | 4236 | 3109 | -1127 | -26.61 % |
| rmk | 3087 | 1430 | -1657 | -53.68 % |

(...) and a last one for the different targeted total items, see below. Interestingly, perhaps different from expected, solely around 7% of the bi-lateral exports *vxmd* were thinned out and even less of the related transport margin demands *vtwr* (1.6%). This reflects the rather conservative way how the threshold is applied for trade flows as discussed above.

|      |       |       |        |          |
|------|-------|-------|--------|----------|
| vfm  | 3969  | 3844  | -125   | -3.15 %  |
| vom  | 1197  | 1165  | -32    | -2.67 %  |
| vxmd | 24717 | 22896 | -1821  | -7.37 %  |
| vtwr | 43296 | 42606 | -690   | -1.59 %  |
| vifm | 68229 | 40849 | -27380 | -40.13 % |
| vdfm | 68229 | 50520 | -17709 | -25.96 % |
| vipm | 1197  | 1033  | -164   | -13.70 % |
| vdpm | 1197  | 1076  | -121   | -10.11 % |
| vigm | 1197  | 658   | -539   | -45.03 % |
| vdgm | 1197  | 452   | -745   | -62.24 % |
| viim | 1197  | 559   | -638   | -53.30 % |
| vdim | 1197  | 701   | -496   | -41.44 % |

Overall, the balancing has proven as quite useful when solving models with both high regional and sector detail, and even more so, if subsequently, the data base is subject to the split.

### Special treatment for specific regions and sectors

When building a data base for a project, it might be desirable to apply less aggressive filtering thresholds for specific regions and / or sectors that are in the focus of the application. The algorithm therefore allows defining lists of regions / sectors with accompanying specific thresholds. The codes for regions / sectors need to be inputted in the two text fields. “*Reduced thresholds only in combination*” will apply the different threshold only to the intersection of the inputted regions and sectors, otherwise, all regions and sectors inputted will be receive different thresholds. Take an example where you enter for regions “*xoc*” and for sectors “*pdr*”. If “*Reduced thresholds only in combination*” is NOT switched on, all transactions of the region “*xoc*” and all transactions for the sector “*pdr*” will be treated differently. If the “*Reduced thresholds only in combination*” is active, only the transaction relating both to “*pdr*” and the region “*xoc*” are treated as exemptions.

However, filtering for the remaining sectors / regions has still an impact on these exemptions. For example, if production of a sector in a region is dropped, the related export flows need to be dropped as well, potentially transactions in regions and sectors are influenced where tighter thresholds are used. Tests have however indicated that very few transactions are lost in regions/sector where stringent thresholds are applied as long as the overall filtering thresholds are not too aggressive.

### Technical aspects

The global GTAP SAM is clearly not well scaled from a data balancing problem as very small and very large transactions must be hosted in the same matrix. That poses a problem for the solvers. The code uses a number of tactics to respond to that challenge. CPLEX as a specialized LP / QCP solver is extremely fast in solving large scale balancing problems with linear constraints. It is therefore the preferred choice to solve balancing problems. The quality of a QCP solution by CPLEX with regard to feasibility accuracy is however lower compared to a LP solution and generally also lower compared to

CONOPT4. Therefore, after the loop over the tests, a second go with CONOPT4 starting with the solution of CPLEX is appended if not CONOPT4 was already used. This additional solves require some extra time but prevent tiny imbalances in the SAM in later model runs. To improve the accuracy further, a generalized RAS is run over maximal 20 rounds if remaining differences in column and rows sum exceed 1.E-8. Generally, the imbalances in the resulting SAMs both in absolute and relative terms are quite low, as discussed next.

### 4.3.9 Rebooking of agri-food intermediate demand

Wolfgang Britz, November 2025

#### Background

The GTAP data base provides relatively high detail for agri-food sectors given their weight in the global economy with around 20 sectors out of 65 in the GTAP standard data base that are either primary agricultural or food processing sectors, or closely related to use of crop or animal products. Given underlying data and the way they are fused and balanced to produce the consistent global data base, the intermediate use matrix is typically quite dense such that almost all products are used by the different sectors.

Many of the reported intermediate uses of agri-food products in not directly related upstream sectors can hardly be interpreted from a production function perspective. The use of crop products, for instance, in the production of industrial goods could be explained by firms providing some form of food services, for instance, by staff restaurants. In these cases, it might be more elegant to let the firms buy these services from sector specialization in providing food services. Doing so reduces the size of the data base and avoids numerical issue with small cost shares.

In the following, drawing on this idea, an approach will be presented that uses consistent re-booking of these intermediate uses to arrive at a GTAP derived data base with lower intermediate use entries. The code is implemented in ‘*gams/build/mancorr/corr\_agrFood.gms*’.

#### Approach

Table 23: Didactic example

|            | Before |      | After  |      |
|------------|--------|------|--------|------|
|            | Sector | AFS  | Sector | AFS  |
| agriFood   | 5      | 600  | 0      | 605  |
| AFS        | 0      | 100  | 5      | 100  |
| rest       | 1995   | 300  | 1995   | 300  |
| Total cost | 2000   | 1000 | 2000   | 1005 |

Table 23 above exemplifies the basic approach. We move the intermediate use of agri-food products in sectors where they are likely not part of the core production process to another sector, here the “Accommodations and Food Service” sector AFS. This includes the intermediate input use in this sector, driving up its total cost or, equivalently, its production value. This additional output is then booked as new (or additional) intermediate use of AFS in such sectors. The approach increases the total intermediate input use in the economy as the agri-food products are now channeled through the AFS sector, but otherwise leaves the accounts unchanged. A re-balancing of the data is hence not necessary.



The transactions are rebooked in a GAMS program that is used in the data driver of CGEBox. It defines first the list of agri-food products to rebook to the AFS sector and next the sectors where direct use in the production process is unlikely:

```
set agrfood(i) / pdr-c,wht-c,gro-c,v_f-c,osd-c,c_b-c,ocr-c
               ctl-c,oap-c,rmk-c,fsh-c,
               cmt-c,omt-c,vol-c,mil-c,pcr-c,sgr-c,ofd-c,b_t-c/;

set removeIt(a) / frs-a,coa-a,oil-a,gas-a,
                 ppp-a,p_c-a,fmp-a,ele-a,eeq-a,
                 ome-a,mvh-a,otn-a,omf-a,tex-a,wap-a,lea-a,
                 gdt-a,wtr-a,cns-a,otp-a,wtp-a,atp-a,cmn-a,
                 ofi-a,ins-a,rsa-a,obs-a,ros-a,dwe-a /;
```

Exempt from the second list are food-processing sectors, chemical sectors and trade and warehousing activities. A number of exemptions defined:

```
set exemptIA(i,a) / (set.i)      . (whs-a,trd-a)
                   (omt-c,cmt-c) . (lea-a)
                   (wol-c,pfb-c) . (tex-a)
                   frs-c        . (lum-a,ppp-a)
                   /;
```

The standard case is to rebook the intermediate use (domestic and imported, both at market and agent prices) to the “afs” sector. The rebooked transactions are temporarily stored in a parameter called *p\_rebook*:

```
p_reBook(rnat,"afs-a",removeIt(a),agrfood(i),"vifm") $ (not exemptIA(i,a)) = vifm(i,a,rNat);
p_reBook(rnat,"afs-a",removeIt(a),agrfood(i),"vifa") $ (not exemptIA(i,a)) = vifa(i,a,rNat);
p_reBook(rnat,"afs-a",removeIt(a),agrfood(i),"vdfm") $ (not exemptIA(i,a)) = vdfm(i,a,rNat);
p_reBook(rnat,"afs-a",removeIt(a),agrfood(i),"vdfa") $ (not exemptIA(i,a)) = vdfa(i,a,rNat);
```

The intermediate use of wool (wol-c) and plant-based fibres (pfb-c) is replaced by use of textiles:

```
p_reBook(rnat,"tex-a",a,"wol-c","vifm") $ (not exemptIA("wol-c",a)) = vifm("wol-c",a,rNat);
p_reBook(rnat,"tex-a",a,"wol-c","vifa") $ (not exemptIA("wol-c",a)) = vifa("wol-c",a,rNat);
p_reBook(rnat,"tex-a",a,"wol-c","vdfm") $ (not exemptIA("wol-c",a)) = vdfm("wol-c",a,rNat);
p_reBook(rnat,"tex-a",a,"wol-c","vdfa") $ (not exemptIA("wol-c",a)) = vdfa("wol-c",a,rNat);

p_reBook(rnat,"tex-a",a,"pfb-c","vifm") $ (not exemptIA("pfb-c",a)) = vifm("pfb-c",a,rNat);
p_reBook(rnat,"tex-a",a,"pfb-c","vifa") $ (not exemptIA("pfb-c",a)) = vifa("pfb-c",a,rNat);
p_reBook(rnat,"tex-a",a,"pfb-c","vdfm") $ (not exemptIA("pfb-c",a)) = vdfm("pfb-c",a,rNat);
p_reBook(rnat,"tex-a",a,"pfb-c","vdfa") $ (not exemptIA("pfb-c",a)) = vdfa("pfb-c",a,rNat);
```

The use of raw forestry products is replaced by the use of wood:

```
p_reBook(rnat,"lum-a",a,"frs-c","vifm") $ (not exemptIA("frs-c",a)) = vifm("frs-c",a,rNat);
p_reBook(rnat,"lum-a",a,"frs-c","vifa") $ (not exemptIA("frs-c",a)) = vifa("frs-c",a,rNat);
p_reBook(rnat,"lum-a",a,"frs-c","vdfm") $ (not exemptIA("frs-c",a)) = vdfm("frs-c",a,rNat);
p_reBook(rnat,"lum-a",a,"frs-c","vdfa") $ (not exemptIA("frs-c",a)) = vdfa("frs-c",a,rNat);
```

In order to get a cleaner definition of the dairy sector, we moved the use of animal products that are no raw milk to the other food-processing sector:

```

set corIAnim(i) / ctl-c,oap-c,wol-c,fsh-c/;
p_reBook(rnat,"ofd-a","mil-a",corIAnim,"vifm") $ (not exemptIA(corIAnim,"mil-a")) = vifm(corIAnim,"mil-a",rNat);
p_reBook(rnat,"ofd-a","mil-a",corIAnim,"vifa") $ (not exemptIA(corIAnim,"mil-a")) = vifa(corIAnim,"mil-a",rNat);
p_reBook(rnat,"ofd-a","mil-a",corIAnim,"vdfm") $ (not exemptIA(corIAnim,"mil-a")) = vdfm(corIAnim,"mil-a",rNat);
p_reBook(rnat,"ofd-a","mil-a",corIAnim,"vdfa") $ (not exemptIA(corIAnim,"mil-a")) = vdfa(corIAnim,"mil-a",rNat);

```

After these transactions are stored, the actual updates of the data base are handled. First, we increase the intermediate input in the target sectors:

```

vifm(i,a,rnat) = vifm(i,a,rnat) + sum(aCur,p_reBook(rnat,a,aCur,i,"vifm"));
vifa(i,a,rnat) = vifa(i,a,rnat) + sum(aCur,p_reBook(rnat,a,aCur,i,"vifa"));
vdfm(i,a,rnat) = vdfm(i,a,rnat) + sum(aCur,p_reBook(rnat,a,aCur,i,"vdfm"));
vdfa(i,a,rnat) = vdfa(i,a,rnat) + sum(aCur,p_reBook(rnat,a,aCur,i,"vdfa"));

```

And update in parallel the make entries by the increases in the intermediate demands

```

makem(a,j,rNat) $ (sum((aCur,i),p_reBook(rnat,a,aCur,i,"vdfa")+p_reBook(rnat,a,aCur,i,"vdfa")) $ j_a(j,a)
= makem(a,j,rNat)
* (1+sum((aCur,i),p_reBook(rnat,a,aCur,i,"vifa")+p_reBook(rnat,a,aCur,i,"vdfa"))/sum(k $ j_a(k,a),makem(a,k,rnat)));

```

For the sectors where intermediate inputs are removed, we delete the entries:

```

vifm(i,a,rNat) $ sum(aCur,p_reBook(rnat,aCur,a,i,"vifm")) = 0;
vifa(i,a,rNat) $ sum(aCur,p_reBook(rnat,aCur,a,i,"vifa")) = 0;
vdfm(i,a,rNat) $ sum(aCur,p_reBook(rnat,aCur,a,i,"vdfm")) = 0;
vdfa(i,a,rNat) $ sum(aCur,p_reBook(rnat,aCur,a,i,"vdfa")) = 0;

```

And introduce respectively add the intermediate input use now delivered by other sectors:

```

vifm(j,a,rnat) = vifm(j,a,rnat) + sum((j_a(j,aCur),i),p_reBook(rnat,aCur,a,i,"vifm"));
vifa(j,a,rnat) = vifa(j,a,rnat) + sum((j_a(j,aCur),i),p_reBook(rnat,aCur,a,i,"vifa"));
vdfm(j,a,rnat) = vdfm(j,a,rnat) + sum((j_a(j,aCur),i),p_reBook(rnat,aCur,a,i,"vdfm"));
vdfa(j,a,rnat) = vdfa(j,a,rnat) + sum((j_a(j,aCur),i),p_reBook(rnat,aCur,a,i,"vdfa"));

```

#### 4.3.10 Macro updates

Wolfgang Britz, September 2026

##### Background

Building up the global GTAP Data base is a challenging exercise requiring to process, combine and correct various raw statistical data such that are complete and consistent. In order to allow for an efficient work flow where updates to individual countries are possible without affecting all others, the first step in building the global Data base consists in setting consistent data of bi-lateral imports at c.i.f. and exports at f.o.b. Both are reported by all countries to WTO at the HS6 level. But these partner notifications often do not agree on traded quantities and it is not uncommon to find that the c.i.f. value reported by the importer is below the f.o.b. value reported by the exporter. Potential reasons are, for instance, changes in exchange rates during the time of transport, transport losses, different HS6 classifications of the same cargo by the importer and exporter, shipments on sea during a reporting year such that are counted as exported but not imported or reporting errors.

Rendering reported quantities consistent and ensuring positive per unit trade margins in a reasonable range requires to decide which notification to trust (more). Whatever approach is taken, at least for one partner reported import or export values by HS6 will change. If this happens for imported export or import flows, the resulting B.o.T. might show larger deviations compared to what the countries own trade notification might suggest. In the GTAP data base, the B.o.P. is closed by setting foreign savings equal to -B.o.T. As later steps in the work flow target GDP, larger deviations from the B.o.T. for a country can then imply larger changes in the regional savings due to  $I = S = -B.o.T. + \text{regional saving}$

+ depreciation. An example for the GTAP12a data release provides Ghana. The GTAP data base reports for 2023 a B.o.T. (including margin exports) of around -7 Bio USD and regional savings of -4.3 Bio USD. The WITS portal instead shows a trade surplus of around 0.5 Bio USD. At the reported net investments of 2.7 Bio USD in the GTAP Data base, the trade surplus shown by WITS would suggest regional savings in the range of 3.2 Bio USD.

It is obvious that the team at Purdue cannot change its global data base each time a user wants some data corrected to match some other statistic. Global completeness and consistency come at the price of changing reported data for single countries. What CGEBox offers is to let the user correct the GTAP Data base while maintaining its internal consistency. The GEMPACK tool Altertax (Malcolm 1998) does the same but is designed to change tax rates. It uses a special version of the GTAP model different from what is used in here. The approach in CGEBox is quite different: it uses a constrained optimization approach without drawing on the equations of the model.

### **Approach**

The approach is based on Linear Programming framework minimizing a penalty function comprising absolute deviations from entries in a global SAM and domestic and import demands by agents at market prices. The various balancing conditions are depicted as equality constraints. They are to a large degree identical to what is also used in the split program of CGEBox, added are constraints related to macro accounts. In-equality constraints allow to control maximal changes in cost and expenditure shares while bounds or fixing of SAM, Armington entries and certain macro totals allow to control in detail the outcome of the balancing process.

The equality constraints cover the following balancing conditions and definitions of totals:

1. Definition of column and row sums for each SAM account in each model region. These sums can also be targeted, such as targeting production value by activity or total household income.
2. Equality of column and row sum for each SAM account in each model region, i.e. keeping the global SAM balanced.
3. Decomposition of demand by agent at market prices (SAM entries) to domestic and import demand.
4. Exhaustion of total import demand at market prices (sum over agents) for each product and region by the sum of bi-lateral imports at c.i.f. plus tariffs.
5. Total domestic demand at market prices (sum over agents) plus sum of exports at f.o.b. minus value of export taxes for each product and region is equal to production value (sum of make entries).
6. Each bi-lateral export flows at f.o.b. plus trade margins equal to mirror import flow at c.i.f. plus transport margins.
7. Definition of global transport margin demands from fixed per unit trade margins and bi-lateral trade flows at f.o.b.
8. Transport margin demand for each country as a fixed share of global transport margin demands.
9. Definition of bi-lateral total exports and imports for each country.
10. Definition of foreign saving's SAM entry as total imports minus total exports minus the country's transport margin demand (= margin exports).
11. Definition of total output value.
12. Definition of final demands (households, government, investments)
13. Definition of regional saving's SAM entry as total investment demand minus foreign savings minus depreciation.
14. Definition of GDP: sum over final demands minus foreign savings.

Based on the defined totals, the following inequality constraints allow to control their composition:

1. Minimum and maximum cost shares for intermediate demand by product at market prices and endowment demand by factor at market prices for each activity and country
2. Minimum and maximum expenditure shares for each product at market prices on each final demand (households, government, investments) for each country
3. Minimum and maximum share of activity output net of production taxes on total production value for each country

Ad-valorem tax rates and per unit transport margin per trade flows are considered constant so far. Changes in individual tax rates can be addressed by introducing manual corrections before filtering in CGEBox.

### ***Default bounds***

Default bounds are set as follows and define maximal deviations:

1. Column sum of endowment accounts in 90-110% range
2. Final demands in 90-100% range
3. Total output value in 95-105%
4. Depreciation is fixed
5. GDP can change by +/-0.1%

Lower and upper shares are defined as follows:

1. +/- 5% for cost share of activities
2. +/- 5% for expenditure shares of final demands
3. +/-5% for revenue shares of total production value

These bounds can be changed by a user provided file that also comprises target (either fixed or as bounds) to any variable in the balancing problem (SAM and Armington entries, final demands, GDP, output value). Note that the B.o.T. and regional savings are specific SAM entries.

### ***Solution approach***

The use of linear (in)equality constraints and a linear penalty function allows to use performant LP solvers to re-balance the problem above. The problem is first solved with a maximal infeasibility tolerance of 1.E-7 (0.1 USD) and slacks of 0.01 USD around the column sums. Should this fail, the tolerance is increase to 1 USD and the slacks to 0.1 USD. If this also fails, a final try uses a maximal infeasibility tolerance of 1.E-5 (10 USD) and increases the slacks to 1 USD. This last try uses CONOPT4 that will also report details of problems found by the triangular preprocessor. After a successful solve, the modified RAS as part of CGEBox is rerun to remove any tiny infeasibilities and as an independent check for correct balancing.

After the solution, the results are assigned to the SAM and the Armington demands at market prices and Armington demands at agent prices re-calculated at fixed tax rates. Equal, all tax revenue flows (production taxes by activity and output, bi-lateral import and export taxes, factor taxes and subsidy by factor and activity) are updated based on the results. All emission factors are updated such as to maintain reported emissions related to the economic transaction.

### ***Interface***

The macro updates are activated with a check box as shown below. In this case, a file selection panel becomes visible. The related search mask is *gams/build/macro\_corr\_\*.gms*. The user can add its own

correction files to this directory, they are automatically added to the file selection on the GUI as long as their name starts with *macro\_corr\_*.

The repository comprises an empty file *gams/build/macro\_corr\_empty.gms* that show how to target typical macro totals:

```
*
* --- example of targeting BOT
*
* v_sam0.lo(rNat,inv,"bop") = sam0(rNat,inv,"bop") * 1.10;
* v_sam0.up(rNat,inv,"bop") = sam0(rNat,inv,"bop") * 1.11;
*
* --- example of targeting GDP
*
* v_gdp.lo(rNat) = v_gdp.l(rnat) * 1.05;
* v_gdp.up(rNat) = v_gdp.l(rnat) * 1.051;
*
* --- example of targeting GDP components (inv,gov,h)
*
* v_finDem.lo(rNat,inv) = v_finDem.l(rnat,inv) * 1.05;
* v_finDem.up(rNat,inv) = v_finDem.l(rnat,inv) * 1.051;
```

The lower and upper bounds for shares are called *p\_shareLoA* and *p\_shareUpA* for costs shares of activities, *p\_shareLoF* and *p\_shareUpF* for expenditure shares of final demand and *p\_shareLoO* and *p\_shareUpO* for the composition of total output value. Their defined around unity. For instance, an upper factor of 1.1 means that the cost or expenditure share can deviate increase by up to 10% compared to the given GTAP data. It might be helpful to check the code of *gtapAgg/macroUpdate.gms* to learn about how these factors are defined if the user wants to change the defaults.

### Reporting

The user can switch on the solution listing via the interface to have a detailed look at all variables – check the output for model *m\_macro*. An alternative to this is to first generate a data base without corrections and then one with corrections (giving it a different name with the postfix option on the GUI) and comparing results with the GDXDIFF utility of GAMS Studio. The code also provides a summary of changes based on the relative changes in the column sums, GDP, X, M, X-M, regional savings and total output value. Values for final demands account refer to product demand, only, instead of full column sums. This output is discussed in the example below.

**An example application**

We apply the approach to change the B.o.T of Ghana to around 0.5 Bio USD in our manual correction file while the GTAP Data report around 7 Bio USD.

```

*
* --- The B.o.T. in GTAP is a large negative number in GTAP (~7 Bio USD)
*      reported as +0.5 Bio USD in WITS
*
  v_sam0.lo("gha",inv,"bop") = 0.49 * 1.E+3;
  v_sam0.up("gha",inv,"bop") = 0.51 * 1.E+3;

*
* --- somewhat tighter bounds for final demand
*
  v_finDem.lo(rnat,finDem) = v_finDem.l(rnat,finDem) * 0.92;
  v_finDem.up(rnat,finDem) = v_finDem.l(rnat,finDem) * 1.08;

```

The test is run on the GTAP V12a Power Data base with full activity detail and product detail and 10 World regions plus Ghana. The electricity products are aggregated into one the labor accounts summarized to low and high skilled. The data base is filtered at 1.E-4 without further corrections. The balancing problem has around 360,000 rows and 520,000 columns and solved by CPLEX in 16 seconds.

The results show very limited changes to accounts and other totals for other regions with the exemption of bi-lateral B.o.T. positions and some macro accounts in *SouthAsia* (screenshot from listing). For *SouthAsia*, an increase in the trade deficit of around 1.1% is reported, equivalent to a change of around -3 Bio USD. As the correction require a total change in X-M for Ghana by around 7.5 Bio USD, changes in these magnitudes for some trading partners are necessary.

```

---- 8058 PARAMETER p_resMacroUpdate

```

|                 |        | before      | after       | %       | diff      |
|-----------------|--------|-------------|-------------|---------|-----------|
| ----- 119 line( |        |             |             |         |           |
| EastAsia        | .GHA   | 13862.254   | 12206.748   | -11.943 | -1655.506 |
| EastAsia        | .bot   | 1122950.866 | 1120857.840 | -0.186  | -2093.025 |
| SEAsia          | .GHA   | 1976.362    | 1925.733    | -2.562  | -50.629   |
| SouthAsia       | .GHA   | 3219.216    | 6172.771    | 91.748  | 2953.555  |
| SouthAsia       | .v_f-c | 209142.743  | 212197.075  | 1.460   | 3054.332  |
| SouthAsia       | .hhsld | 2992778.751 | 2995849.060 | 0.103   | 3070.310  |
| SouthAsia       | .imptx | 53150.174   | 53224.791   | 0.140   | 74.618    |
| SouthAsia       | .bop   | 994606.091  | 997585.805  | 0.300   | 2979.714  |
| SouthAsia       | .rsav  | 618858.274  | 615855.335  | -0.485  | -3002.940 |
| SouthAsia       | .bot   | -255206.695 | -258209.635 | 1.177   | -3002.940 |
| SouthAsia       | .X     | 994606.091  | 997585.805  | 0.300   | 2979.714  |
| NAmerica        | .GHA   | 4964.759    | 4935.414    | -0.591  | -29.346   |
| LatinAmer       | .GHA   | 1058.448    | 1052.380    | -0.573  | -6.068    |
| EU_27           | .GHA   | 7953.104    | 8953.994    | 12.585  | 1000.890  |
| EU_27           | .ofd-c | 964742.720  | 965860.379  | 0.116   | 1117.658  |
| EU_27           | .bot   | 402651.437  | 400293.623  | -0.586  | -2357.814 |
| SSA             | .GHA   | 4857.401    | 4898.412    | 0.844   | 41.010    |
| RestofWorld     | .GHA   | 8285.045    | 8914.652    | 7.599   | 629.607   |
| RestofWorld     | .nfm-c | 463287.868  | 463855.081  | 0.122   | 567.213   |

Given a GDP of around 80 Bio USD for Ghana, the change in the B.o.T. by 7.5 Bio USD must require some larger changes in accounts at fixed GDP. The screenshot shows below filtering the macro accounts for Ghana shows the expected changes: as foreign savings are reduced by 7.5 Bio USD, regional savings increase by around 6 Bio USD, the remainder stems from a downward correction of total investments by around -0.6 Bio USD or -7.4%. As regional savings are sourced by the household account, the increase in regional savings is offset by a reduction in final demand of households by around -8.9%. This also reflects that forcing X-M to grow at fixed GDP implies a reduction in final demand by the same amount.

Total exports increase by 5 Bio USD and total imports decrease by -2.5 Bio USD, yielding the targeted changes by +7.5 Bio USD. While total output value (*outVal*) slightly increases (+0.8 Bio USD or 0.5%) and total intermediate consumption (*int*) is rather stable, the massive increase in exports implies a reduction in domestic sales by around -4.3 Bio USD. Changes in tax flows also imply a reduction in government demand by around 3.8%. Overall factor income by around 0.7 Bio USD despite a small reduction in extraction rents. This is mostly driven by a change in the output composition.

```

----- 8070 Changes in SAM accounts and GDP after macro update
----- 8070 PARAMETER p_resMacroUpdate

```

|                  |          | before     | after      | %        | diff      |
|------------------|----------|------------|------------|----------|-----------|
| ----- 67 line(s) |          |            |            |          |           |
| GHA              | .dom     | 110649.073 | 106411.852 | -3.829   | -4237.221 |
| GHA              | .Land    | 7151.450   | 7272.187   | 1.688    | 120.737   |
| GHA              | .UnSkLab | 26293.582  | 26391.039  | 0.371    | 97.456    |
| GHA              | .Capital | 25277.332  | 25812.094  | 2.116    | 534.762   |
| GHA              | .NatRes  | 872.446    | 843.324    | -3.338   | -29.123   |
| ----- 34 line(s) |          |            |            |          |           |
| GHA              | .hhsld   | 74388.707  | 67748.730  | -8.926   | -6639.976 |
| GHA              | .gov     | 4787.661   | 4608.466   | -3.743   | -179.195  |
| GHA              | .inv     | 8401.429   | 7780.250   | -7.394   | -621.180  |
| GHA              | .int     | 60911.629  | 60981.382  | 0.115    | 69.753    |
| GHA              | .indtx   | 6782.183   | 6374.161   | -6.016   | -408.022  |
| GHA              | .prdtx   | 3.498      | 3.893      | 11.299   | 0.395     |
| GHA              | .imptx   | 2627.654   | 2389.674   | -9.057   | -237.980  |
| GHA              | .fcttx   | 653.654    | 655.978    | 0.356    | 2.324     |
| GHA              | .dirtx   | -9709.543  | -9245.452  | -4.780   | 464.091   |
| GHA              | .bop     | 31066.913  | 28577.953  | -8.012   | -2488.960 |
| GHA              | .facy    | 64744.509  | 65468.342  | 1.118    | 723.833   |
| GHA              | .rsav    | -4305.104  | 2594.614   | -160.268 | 6899.718  |
| GHA              | .bot     | -7030.898  | 490.000    | -106.969 | 7520.898  |
| GHA              | .X       | 24036.015  | 29067.953  | 20.935   | 5031.938  |
| GHA              | .M       | 31066.913  | 28577.953  | -8.012   | -2488.960 |
| GHA              | .outVal  | 134621.919 | 135416.640 | 0.590    | 794.721   |
| GHA              | .indTax  | 10126.754  | 9483.468   | -6.352   | -643.286  |
| GHA              | .regy    | 74871.263  | 74951.810  | 0.108    | 80.547    |

## References

MALCOLM, Gerard. Adjusting tax rates in the GTAP data base. 1998, <https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1013&context=gtaptp>



#### 4.3.11 Generalized RAS after building the SAM

After the transactions stored on the various array and vectors jointly defining the economic part of the GTAP data base have been mapped to a global SAM plus some additional arrays (see “*gtapAgg/buildSam.gms*”), a Generalized RAS coded in GAMS (see “*build/gras.gms*”) is used to check for a balanced SAM and to remove any tiny imbalances.

Before applying the iterative re-balancing, the program first makes sure that bi-lateral imports and exports are mutually compatible:

```
*
* --- render bi-lateral exports / imports consistent
*
sam0(rp,rNat,i) $ sam0(rp,rNat,i) = (sam0(rNat,i,rp) + sum(m, xMarg0(m,rNat,i,rp))) * 0.5 + sam0(rp,rNat,i) * 0.5;
sam0(rNat,i,rp) $ sam0(rNat,i,rp) = sam0(rp,rNat,i) - sum(m, xMarg0(m,rNat,i,rp));
```

Afterwards, it defines column and rows sum for all SAM accounts *isi* (alias *jsj*) and all model region *rNat* and calculates their global sum and the maximal error:

```
option kill=p_checkSam;
p_checksam(rNat,isi,"colSum") = sum(jsj, sam0(rNat,isi,jsj));
p_checksam(rNat,isi,"rowSum") = sum(jsj, sam0(rNat,jsj,isi));
p_checksam(rNat,isi,"delta") = p_checksam(rNat,isi,"colSum") - p_checksam(rNat,isi,"rowSum");
p_checksam(rNat,"max","delta") = smax(isi, abs(p_checksam(rNat,isi,"delta")));
```

The process will start only if the maximal error found is below 1 (i.e. 1 Mio USD). Otherwise, the error is considered too large and indicates some problems in the underlying data or some programming plunder in the constructing the SAM.

The RAS as an iterative procedure requires stopping criteria, here chosen as the maximal error left with 1.E-9 or if the sum of remaining errors if 1.E-11 times the number of regions and SAM accounts and the maximal errors is below 2.5E-9. Iterations are also stopped if neither the largest error nor the sum of the errors could be improved during the last iteration. The chosen thresholds consider the achievable accuracy given double precision arithmetic and the likely spread of transaction sfound in one of the rows or columns of the SAM.

The RAS does not reflect the specific structure of the matrix to balance, quite different from the filter rebalancing problem. As the balancing is here applied to sub-matrices defined by the model regions, we exclude bi-lateral transactions from corrections, i.e. trade flows are kept untouched. Equally, the code excludes accounts that are already perfectly balanced. This is achieved by defining of set of transactions to correct:

```
toCorr(rNat,isi,jsj) $ ( abs(sam0(rNat,isi,jsj)) gt 1.E-7) = yes;
toCorr(rNat,isi,jsj) $ ( abs(p_checksam(rNat,isi,"delta")) le 1.E-12) = no;
toCorr(rNat,isi,jsj) $ ( abs(p_checksam(rNat,jsj,"delta")) le 1.E-12) = no;
toCorr(rNat,isi,jsj) $ toCorr(rNat,jsj,isi) = yes;
toCorr(rNat,jsj,isi) $ toCorr(rNat,isi,jsj) = yes;
toCorr(rNat,rp,i) = no;
toCorr(rNat,i,rp) = no;
```

As seen, the last two lines exclude the bi-lateral imports and exports. Equally, transactions where both the column and row errors are below 1.E-12 are excluded. The generalized RAS is a variant of RAS designed to work with both positive and negative entries. In each iteration, first the row entries are scaled and then the column entries. To do so, the sum of positive and negative SAM entries by model regions are first defined, and the targeted column sum if defined from the current one minus half of the error:



```

p_checksam(rNat,isi,"p_colSum") $ (p_checksam(rNat,isi,"p_colSum") $ (abs(p_checksam(rNat,isi,"delta")) ge 1.E-9))
= sum(jsj $ (sam0(rNat,isi,jsj) gt 0), sam0(rNat,isi,jsj));
p_checksam(rNat,isi,"n_colSum") $ (p_checksam(rNat,isi,"n_colSum") $ (abs(p_checksam(rNat,isi,"delta")) ge 1.E-9))
= -sum(jsj $ (sam0(rNat,isi,jsj) lt 0), sam0(rNat,isi,jsj));

p_checksam(rNat,isi,"colSumD") = p_checksam(rNat,isi,"colSum") - 0.5 * p_checksam(rNat,isi,"delta");

```

Next, a correction multiplier is defined:

```

p_checksam(rNat,isi,"multC") = 0;
p_checksam(rNat,isi,"multC") $ ((abs(p_checksam(rNat,isi,"delta")) ge 1.E-9) $ p_checksam(rNat,isi,"colSumD"))
= { [ p_checksam(rNat,isi,"colSumD") - 2*p_checksam(rNat,isi,"p_colSum")
    + sqrt( p_checksam(rNat,isi,"colSumD")*p_checksam(rNat,isi,"colSumD")
    + 4*p_checksam(rNat,isi,"p_colSum")*p_checksam(rNat,isi,"n_colSum")) ]
  / (2*p_checksam(rNat,isi,"p_colSum")) } $ p_checksam(rNat,isi,"p_colSum")
+ { - p_checksam(rNat,isi,"n_colSum")/p_checksam(rNat,isi,"colSumD") -1 }
  $ (p_checksam(rNat,isi,"p_colSum") lt 1.E-8*p_checksam(rNat,isi,"n_colSum"));

```

This multiplier is applied to positive transactions and its reciprocal to negative ones:

```

sam0(toCorr(rnat,isi,jsj)) $ ((sam0(rnat,isi,jsj) gt 0) $ p_checksam(rNat,isi,"multC"))
= sam0(rnat,isi,jsj) * p_checksam(rNat,isi,"multC") + sam0(rnat,isi,jsj);

sam0(toCorr(rnat,isi,jsj)) $ ((sam0(rnat,isi,jsj) lt 0) $ p_checksam(rNat,isi,"multC"))
= sam0(rnat,isi,jsj) * (1/(1+p_checksam(rNat,isi,"multC")) -1) + sam0(rnat,isi,jsj);

```

Afterwards, the same logic is applied to the row sums. If this algorithm stalls, another variant called *dss* is switched on:

```

if ( ( (p_maxError ge p_lastMaxError) and (p_sumError ge p_lastSumError)) or (trials.pos gt card(trials)/10) and (dss eq 0),
    --- if G-RAS stalls, switch to DSS

    dss = 1;
    p_maxError = p_lastMaxError;
    p_sumError = p_lastSumError;
    sam0(rnat,isi,jsj) = p_samLast(rnat,isi,jsj);
else
    dss = 1 $ ( (p_maxError lt p_lastMaxError) or (p_sumError lt p_lastSumError)) $ dss
);

```

While RAS balances in each iteration all columns and rows, the DSS algorithm corrects the largest error, only:

```

option kill=maxis;
maxis(rnat,isi) $ ((abs(p_checksam(rNat,isi,"delta")) eq p_checkSam(rnat,"max","delta")) $ abs(p_checksam(rNat,isi,"delta"))) = YES;
maxis(rnat,isi) $ (isi.pos ne smax(jsj $ maxis(rnat,jsj), jsj.pos)) = NO;

p_checksam(maxis(rNat,isi),"scaleDSS") $ (sign(p_checksam(rNat,isi,"rowsum")) eq sign(p_checkSam(rNat,isi,"colsum")))
= sqrt( p_checksam(rNat,isi,"rowsum") / p_checksam(rNat,isi,"colsum") );

sam0(rnat,isi,jsj) $ ((not maxis(rnat,jsj)) $ toCorr(rnat,isi,jsj) $ p_checksam(rNat,isi,"scaleDSS"))
= sam0(rnat,isi,jsj) + sam0(rnat,isi,jsj) * (p_checksam(rNat,isi,"scaleDSS") -1);

sam0(rnat,jsj,isi) $ ((not maxis(rnat,jsj)) $ toCorr(rnat,jsj,isi) $ p_checksam(rNat,isi,"scaleDSS"))
= sam0(rnat,jsj,isi) + sam0(rnat,jsj,isi) * (1/p_checksam(rNat,isi,"scaleDSS")-1);

```

It is mostly not able to improve on RAS.

The RAS can consider the transactions stored the global SAM, only, not additional data such as one the Armington decomposition for each agent. Therefore, after the balancing, errors for these transactions are defined:

```

option kill=p_xerr;
p_xerr(rnat,aa,i) $ (not sameas(aa,"trdMg")) = sam0(rnat,i,aa) - xmm0(rNat,i,aa) - xdm0(rNat,i,aa);

```

And the components scaled accordingly:

```

*
* --- correct imported and domestic origin to absorb error in Amington calculation
*
xmm0(rNat,i,aa) $ (p_xerr(rnat,aa,i) $ xmm0(rNat,i,aa))
= max(0,xmm0(rnat,i,aa) + [xmm0(rnat,i,aa)/sam0(rnat,i,aa)]*p_xerr(rnat,aa,i)) $ sam0(rnat,i,aa);
xcm0(rNat,i,aa) $ (p_xerr(rnat,aa,i) $ xcm0(rNat,i,aa))
= max(0,xcm0(rnat,i,aa) + [xcm0(rnat,i,aa)/sam0(rnat,i,aa)]*p_xerr(rnat,aa,i)) $ sam0(rnat,i,aa);

xma0(rNat,i,aa) $ (p_xerr(rnat,aa,i) $ xma0(rNat,i,aa))
= max(0,xma0(rnat,i,aa) + [xma0(rnat,i,aa)/sam0(rnat,i,aa)]*p_xerr(rnat,aa,i)) $ (sam0(rnat,i,aa) $ xmm0(rNat,i,aa));
xda0(rNat,i,aa) $ (p_xerr(rnat,aa,i) $ xda0(rNat,i,aa))
= max(0,xda0(rnat,i,aa) + [xda0(rnat,i,aa)/sam0(rnat,i,aa)]*p_xerr(rnat,aa,i)) $ (sam0(rnat,i,aa) $ xcm0(rNat,i,aa));

xma0(rNat,i,aa) $ (abs(xma0(rnat,i,aa)-xmm0(rnat,i,aa)) le 1.E-10) = xmm0(rnat,i,aa);
xda0(rNat,i,aa) $ (abs(xda0(rnat,i,aa)-xcm0(rnat,i,aa)) le 1.E-10) = xcm0(rnat,i,aa);
sam0(rNat,i,aa) $ (not sameas(aa,"trdMg")) = xmm0(rNat,i,aa)+xcm0(rNat,i,aa);

sam0(rNat,"indtx",aa) = sum(i,xda0(rNat,i,aa) + xma0(rNat,i,aa) - xcm0(rNat,i,aa) - xmm0(rNat,i,aa) );

```

Resulting errors > 0.1 USD are reported:

```

p_xerr(rnat,aa,i) $ ( abs(p_xerr(rnat,aa,i)) le 1E-7) = 0;
display $ card(p_xerr) "Some deviations found after GRAS between product demand in SAM and decomposition to domestic and imported origin",p_xerr;
logfile.nr=1;

```

After these steps, the program “*gtapAgg/buildSam*” will check for a balanced SAM, discarding errors below 0.01 USD:

```

*
option kill=p_checkSam;
p_checksam(rNat,isi,"colSum") = sum(jsj, sam0(rNat,isi,jsj));
p_checksam(rNat,isi,"rowSum") = sum(jsj, sam0(rNat,jsj,isi));
p_checksam(rNat,isi,"delta") = p_checksam(rNat,isi,"colSum") - p_checksam(rNat,isi,"rowSum");
*
* --- differences below 0.01 USD are discarded
*
p_checksam(rNat,isi,"delta") $ (abs(p_checksam(rNat,isi,"delta")) le 1.E-8) = 0;
p_checksam(rNat,isi,"colSum") $ (abs(p_checksam(rNat,isi,"delta")) le 1.E-8) = 0;
p_checksam(rNat,isi,"rowSum") $ (abs(p_checksam(rNat,isi,"delta")) le 1.E-8) = 0;

```

Afterwards, relative differences are reported for errors above 0.1 Mio USD, and if any such difference is found, a warning is issued and the parameters are shown in the listing:

```

*
* --- check relative errors for differences above 0.1 Mio USD
*
p_checksam(rNat,isi,"delta %") $ (abs(p_checksam(rNat,isi,"delta")) gt 0.1) = 200*p_checksam(rNat,isi,"delta")
/ (p_checksam(rNat,isi,"colSum") + p_checksam(rNat,isi,"rowSum"));

parameter maxDiff;
maxDiff("max","delta") = smax( (rnat,isi), abs(p_checksam(rNat,isi,"delta")));
maxDiff("max","delta %") = smax( (rnat,isi), abs(p_checksam(rNat,isi,"delta %")));
maxDiff(rNat,isi) $ p_checksam(rNat,isi,"delta %") = p_checksam(rNat,isi,"delta");

display $ maxDiff("max","delta") "Warning, differences in SAM",maxDiff,p_checkSam;

```

Afterwards, all relative errors below 0.01% are removed. If errors remain, the SAM is considered imbalanced and any further processing stops:

```

scalar p_percD / 1.E-4 /;

p_checksam(rNat,isi,"colSum") $ (abs(p_checksam(rNat,isi,"delta %")) le p_percD) = 0;
p_checksam(rNat,isi,"rowSum") $ (abs(p_checksam(rNat,isi,"delta %")) le p_percD) = 0;
p_checksam(rNat,isi,"delta") $ (abs(p_checksam(rNat,isi,"delta %")) le p_percD) = 0;
p_checksam(rNat,isi,"delta %") $ (abs(p_checksam(rNat,isi,"delta %")) le p_percD) = 0;
abort $ card(p_checkSam) "Error, SAM is not balanced ",p_checkSam,SAM0;

```

#### 4.3.12 Gender Differentiated labor data

Use WorldBank GDLD gender labour data



The World Bank provides labor data matching the region and sector definition of GTAP V10 (141 regions, 65 sector, base year 2014) which are available as a GDX container with CGEBox

(*data/gdld\_2014\_v10.gdx*). The data are merged into the data set when the check-box above is operated. It can be used with the GAP Power Data base in which case proportionality assumptions are used to map the shares from the single electricity sector in the standard data set to the additional sector found in the GTAP Power Data Base.

Using the GDLD (Gender Differentiated Labour Data) will overwrite the aggregation rules from the 5 labour categories in the GTAP Data base in your aggregation file. Instead, four labour categories will be derive: female skilled, female unskilled, male skilled and male unskilled, reflecting the detail available in GDLD.

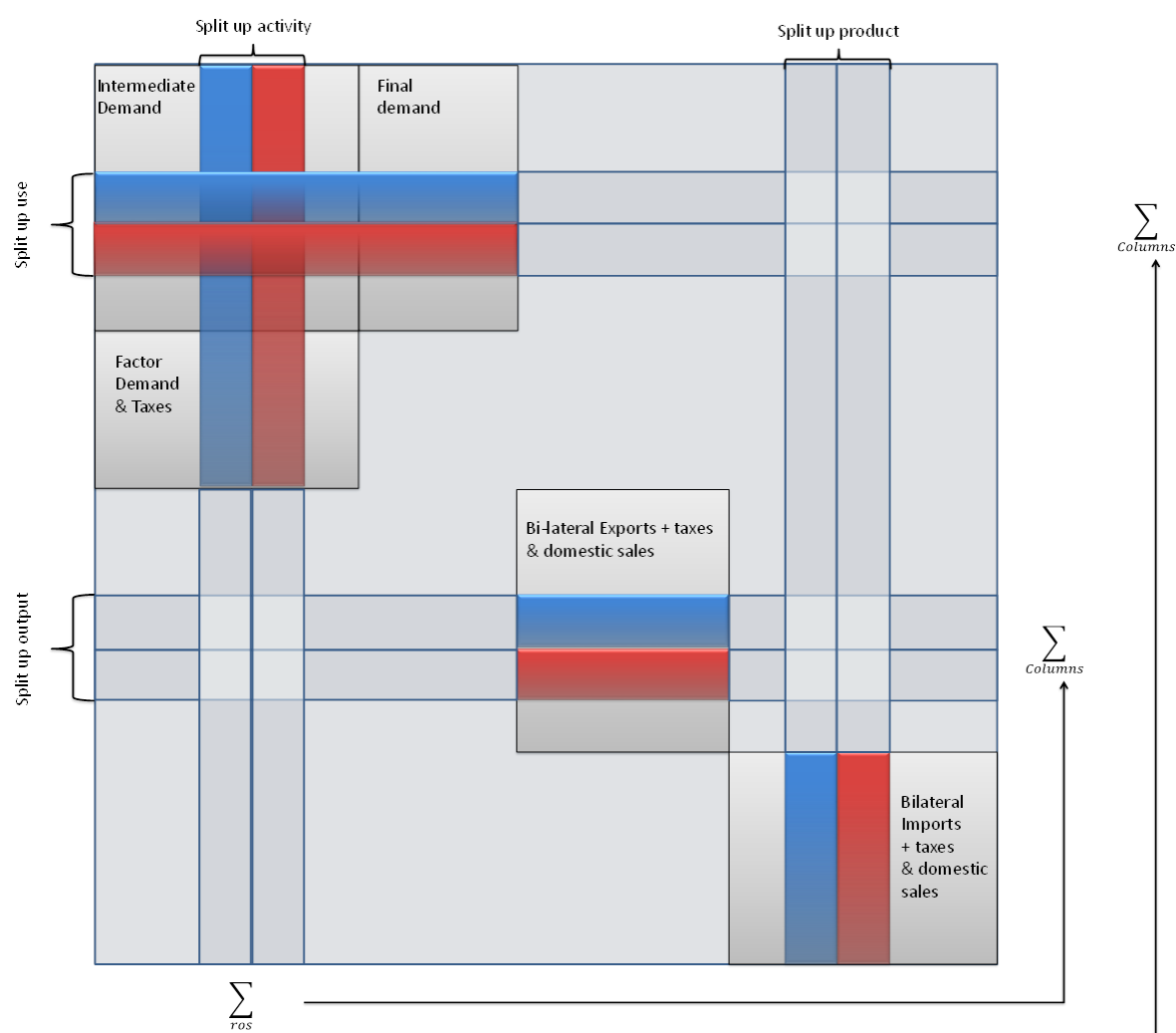
#### 4.3.13 Split utility



The split utility (*“split\split.gms”*), switched on with the check box shown above, allows to dis-aggregate the global SAM generated by the GTAPAgg data driver of CGEBox to further product and sector detail. It splits commodities and activities and rebalances the database using a mix of absolute and absolute normalized differences. Some details are discussed in Britz, W. (2021): Comparing Penalty Functions in Balancing and Dis-aggregating Social Accounting Matrices, *Journal of Global Economic Analysis* 6(1): 34-81.

The split utility requires a file with split shares in a specific format which needs to be selected on the interface. These files are stored under *“gams/build/split”* where examples can be found.

To understand the working of the split utility, consider the case where we take one activity in each region and its related output products and split it into two activities each producing one new related output. Starting from the 57 or 65 GTAP sectors, we would add two new activities and remove one which is dis-aggregated in the new data base. The same holds for the products. SAM and other data base entries not referring either to the activity to split up or its output, i.e. product, remain untouched. For any SAM cell referring to the activity or product to split, the new, more detailed SAM needs to provide a consistent dis-aggregation. At the same time, for the newly introduced columns and rows in the SAM, column and row sums must match. Additionally, auxiliary data not directly covered by the SAM such as direct and indirect tax rates differentiated by agent and by the imported and domestic origin, bi-lateral import and export taxes etc. must be defined for the split-up activities and products such as to exactly exhaust the original, more aggregated data. As detailed below, this is achieved by defining these conditions as constraints of a programming problem.



Graphic: The split problem

Mostly, the user of the split utility does not need to understand in detail the set-up and solution of the programming problem as discussed below. He feeds a-priori information on split shares (or absolute values) and tax rates into the problem, minimally, it is sufficient to provide shares on outputs for new introduced activities if these are outputting one single product. This information is provided via GAMS code written by the user.

Britz, W. (2022). "Disaggregating agro-food sectors in the GTAP data base. *Journal of Global Economic Analysis*, 7(1), 44-75 discusses the use of the split utility based on a Multi-Regional Input-Output data base for agriculture and food products, this approach is by now replaced by using FAOSTAT data. An example for splits using FAOSTAT can be found in "*gams/build/split/fao-agr.gms*". FAOSTAT data as used there are found in "FAO\_2017\_forSplits.gdx". Another example of applying the split utility is found in Escobar, N., Haddad, S., Börner, J., Britz, W. (2018). "Land use mediated GHG emissions and spillovers from increased consumption of bioplastic", *Environmental Research Letters*, who split up the chemical sector of the GTAP data base into conventional plastics, bio-based ones and the rest of the chemical sector.

The balancing problem is set up as a Linear Programming (LP) problem which renders it possible to use performant LP solvers achieving a high accuracy. It uses both equality and inequality constraints. Different from the filtering routine and the Generalized RAS discussed in sections above, the

balancing problems works globally, i.e. simultaneously for all model regions such that transactions relating to bi-lateral trade are endogenously determined with other ones. A former solution based on quadratic programming proved numerically less stable. For global SAMs with many regions while introducing at the same time many splits, the resulting data balancing problems can comprise several hundredth thousand or even millions of variables and equations. It is recommended to use CPLEX as a specialized LP solver for problems of that size even if CONOPT4 might also do the job. As the data distributed by the GTAP Center have been balanced (or at least stored) with simple precision arithmetic, the split utility will typically fail if the input data are not rebalanced based on the filter routine before as detailed above. Usually, skimming small transactions will also help to stabilize the behaviour of the split utility.

#### 4.3.14 An evolved example for a split

In the following, we discuss the input for an evolved split example to help users developing their own split input. It introduces a difference between subsistence and commercial products in the SAM where only the commercial variant is internationally traded:

```
set s_spliti / wht_sub-c "Wheat subsistence"
              wht_com-c "Wheat commercial"
```

Along with related production activities:

```
set s_splita / wht_sub-a "Wheat subsistence"
              wht_com-a "Wheat commercial"
```

These set definitions must be in a first block in the user provided split file:

```
$iftheni.mode %1==decl
↓
```

It is assumed that the subsistence activity variant uses little intermediate inputs. We first introduce assumptions on the output shares for the commercial variant and define the set of commercial activities as those where output shares are given:

```
table p_outputShare(splita,rNat)
      set.rNat
wht_com-a 0.3
pdr_com-a 0.3
gro_com-a 0.3
v_f_com-a 0.4
osd_com-a 0.3
ocr_com-a 0.5
;

set comA(splita) "Commercial variants";
comA(splita) $ sum(rNat, p_outputShare(splita,rNat)) = YES;
```

The distinction is deemed to be only relevant for low-income countries. We therefore split the regions into two: those with a per-capita factor income above 2.500 \$ and below, and set the output share for the commercial variant for the high-income countries to unity:

```

: *
: * --- Rest of the world and similar (check by counting the mapped original countries)
: * is not split up
: *
: set restOfWorld(rNat);
: restOfWorld(rNat) $ ( sum( (f,a), sam0(rNat,f,a)) / pop0(rnat) > 2500 ) = YES;
:
: set DC(rNat);
: DC(rnat) $ (not restOfWorld(rNat)) = yes;
:
: p_outPutShare(coma,restOfWorld) = 1;

```

In order to exclude international trade of the subsistence product, we introduce “eps” for a very small number as a split factor, and set the split factor for bi-lateral imports for the commercial variant to unity. Accordingly, all reported trade will be mapped to the commercial variant:

```

*
* --- no imports and exports of subsistence product
*
p_splitFactor(rNat,splitI,rNat1) $ sum(i_a(splitI,subsA),1) = eps;
p_splitFactor(rNat,splitI,rNat1) $ sum(i_a(splitI,coma),1) = 1;

```

Based on the output share, we introduce assumptions on intermediate input use:

```

*
* --- almost non intermediate use for subsistence
*
p_splitFactor(DC,subsA,i) = p_splitFactor(DC,subsA,"prod") * 0.25 * sum(as_a(subsA,a),sam0(DC,i,a));
p_splitFactor(DC,coma,i) = max(sum(as_a(comA,a),sam0(DC,i,a))*1.E-4,
sum(as_a(comA,a),sam0(DC,i,a)) - sum(subsA_coma(subsA,coma),p_splitFactor(DC,subsA,i)));

```

Similar statements introduce assumption on factor cost shares. The definition of the split factors must be introduced in a second block:

```
$elseifi.mode %1==run
```

The last block allows introducing bounds and tax rates into the balancing problems after the split-factors had been processed and a-priori values for the new SAM entries derived. In our example, we remove the fix bounds on tax rates and replace them by tiny ones, and next remove use tax rates for home production (= subsistence products):

```

*
* --- no taxes on consumption on home production
*
p_mintx(DC,splitI,aa) $ sum(i_a(splitI,subsA),1) = 0;
p_dintx(DC,splitI,aa) $ sum(i_a(splitI,subsA),1) = 0;

v_mintx.fx(DC,splitI,aa) $ sum(i_a(splitI,subsA),1) = 0;
v_dintx.fx(DC,splitI,aa) $ sum(i_a(splitI,subsA),1) = 0;

```

As the average tax for the subsistence and commercial has to exhaust the one found in the GTAP data base, the tax rates for the commercial one become endogenous variables:

```

*
* --- consumption taxes on commercial production are endogenous
*
v_mintx.lo(DC,splitI,aa) $ (sum(i_a(splitI,coma),1) $ p_mintx(DC,splitI,aa)) = -1E+10;
v_mintx.up(DC,splitI,aa) $ (sum(i_a(splitI,coma),1) $ p_mintx(DC,splitI,aa)) = 1E+10;

v_dintx.lo(DC,splitI,aa) $ (sum(i_a(splitI,coma),1) $ p_dintx(DC,splitI,aa)) = -1E+10;
v_dintx.up(DC,splitI,aa) $ (sum(i_a(splitI,coma),1) $ p_dintx(DC,splitI,aa)) = 1E+10;

```

To avoid large deviations from the desired output shares, we introduce bounds on the sum of the row entries of the subsistence variant of each split-up activity

```

*
* --- fix desired share of subsistence production
*
v_rowSum.lo(DC,SubsA) $ (p_outputShare(SubsA,DC) gt eps)
= sum((as_a(SubsA,a),i), sam0(DC,a,i) * p_outputShare(SubsA,DC)) * 0.9;
*
v_rowSum.up(DC,SubsA) $ (p_outputShare(SubsA,DC) gt eps)
= sum((as_a(SubsA,a),i), sam0(DC,a,i) * p_outputShare(SubsA,DC)) * 1.1;

```

Deviations from these bounds receive a high penalty during balancing. The program from which the example is taken can be found under “*gams\build\split\split\_subs\_fab.gms*”. In that directory, more examples can be found which can serve as the basis to develop own splits.

#### 4.3.15 Balancing equations and objective function in the split problem

The four core equations of the split program shown below ensure that the given entries in the SAM for a product or activity to split are exhausted by the newly introduced more dis-aggregated commodities or products. Here, *spliti* are the newly introduced commodities and *splita* the new activities. The cross set *is\_i* shows the link between the new commodities and the commodity to split while *as\_a* catches the relations between the activities. SAM0 is the global SAM before splitting.

The exhaustion condition of SAM cells referring to products to split found in the rows of the SAM, such as exports (excluding input demand of split-up activities, captured in the equation above), is defined as:

$$\overline{\text{sam0}}_{r,i,s} = \sum_{\text{spliti} \in i} v_{\text{sam},r,\text{spliti},s} \quad (76)$$

The exhaustion of SAM cells referring to products to split in columns, such as imports (excluding the ones referring to split-up activities), is given by the following constraint:

$$\overline{\text{sam0}}_{r,i,s,i} = \sum_{\text{spliti} \in i} v_{\text{sam},r,i,s,\text{spliti}} \quad (77)$$

The related equations in GAMS are shown below (see *build/split\_defmodel.gms*). The first dollar conditions in each equation make sure that only non-zero entries in the global data base are subject to a split. The parameter *p\_sam0* comprise starting values for the new SAM entries, whereas *v\_sam* are the endogenous new SAM entries subject to balancing. Note that the program can introduce slacks *v\_samSlack0P* and *v\_samSlack0N* attached to the original SAM entries in case of numerical difficulties:

```

*
* --- product split must exhaust original commodity entries in each column
*
e_samIJ(rNat,isSplitupI(i),notSplitupA(jsj)) $ (sam0(rNat,i,jsj)
$ sum(is_i(spliti,i) $ p_sam0(rNat,spliti,jsj), (v_sam.range(rNat,spliti,jsj) ne 0))) ..
sam0(rNat,i,jsj)
+ (v_samSlack0P(rNat,i,jsj) - v_samSlack0N(rNat,i,jsj))
=E= sum(is_i(spliti,i) $ p_sam0(rNat,spliti,jsj), v_sam(rNat,spliti,jsj));
e_samJI(rNat,notSplitupA(jsj),isSplitupI(i)) $ (sam0(rNat,jsj,i)
$ sum(is_i(spliti,i) $ p_sam0(rNat,jsj,spliti), (v_sam.range(rNat,jsj,spliti) ne 0))) ..
sam0(rNat,jsj,i)
+ (v_samSlack0P(rNat,jsj,i) - v_samSlack0N(rNat,jsj,i))
=E= sum(is_i(spliti,i) $ p_sam0(rNat,jsj,spliti), v_sam(rNat,jsj,spliti));

```

Balancing for input and factor demands as well as taxes related to activities to split *a* is ensured by the following constraint:



$$\overline{\text{sam0}}_{r, is, a} = \sum_{\text{split} \in a} v_{\text{sam}}_{r, is, \text{split}} + \sum_{\text{split} \in a, \text{split} \in is} v_{\text{sam}}_{r, \text{split}, \text{split}} \quad (78)$$

The first term on the RHS depicts SAM rows which do not refer to products subject to splitting; these are added with the second expression. The set of new and original SAM rows is indexed with *is*; *splita* and *spliti* are the newly introduced activities and products, respectively. Variables which are zero by definition (such as demand for products to disaggregate in the matrix  $v_{\text{sam}}$ ) are fixed exogenously to zero.

In the implementation in GAMS shown below, the first expression after the equal sign is the simpler case where the row *is* is not itself subject to a split. The second line refers to products which are themselves split up.

```
e_samAI(rNat, isSplitUpA(a), isi) $ (
    sum(as_a(splita, a) $ p_sam0(rNat, splita, isi), (v_sam.range(rNat, splita, isi) ne 0))
    +sum((as_a(splita, a), is_i(spliti, i)) $ (p_sam0(rNat, splita, spliti) $ sameas(isi, i)), (v_sam.range(rNat, splita, spliti) ne 0))
    $ sam0(rNat, a, isi)
) ..

sam0(rNat, a, isi)
+ (v_samSlack0P(rNat, a, isi) - v_samSlack0N(rNat, a, isi))

=E= sum(as_a(splita, a) $ p_sam0(rNat, splita, isi), v_sam(rNat, splita, isi))
+sum((as_a(splita, a), is_i(spliti, i)) $ (p_sam0(rNat, splita, spliti) $ sameas(isi, i)), v_sam(rNat, splita, spliti));
```

The exhaustion condition for the SAM columns relating to activities in the rows (outputs, production taxes) is defined as follows:

$$\overline{\text{sam0}}_{r, a, is} = \sum_{\text{split} \in a} v_{\text{sam}}_{r, \text{split}, is} + \sum_{\text{split} \in a, \text{split} \in is} v_{\text{sam}}_{r, \text{split}, \text{split}} \quad (79)$$

In GAMS:

```
e_samAI(rNat, isSplitUpA(a), isi) $ (
    sum(as_a(splita, a) $ p_sam0(rNat, splita, isi), (v_sam.range(rNat, splita, isi) ne 0))
    +sum((as_a(splita, a), is_i(spliti, i)) $ (p_sam0(rNat, splita, spliti) $ sameas(isi, i)), (v_sam.range(rNat, splita, spliti) ne 0))
    $ sam0(rNat, a, isi)
) ..

sam0(rNat, a, isi)
+ (v_samSlack0P(rNat, a, isi) - v_samSlack0N(rNat, a, isi))

=E= sum(as_a(splita, a) $ p_sam0(rNat, splita, isi), v_sam(rNat, splita, isi))
+sum((as_a(splita, a), is_i(spliti, i)) $ (p_sam0(rNat, splita, spliti) $ sameas(isi, i)), v_sam(rNat, splita, spliti));
```

The split utility allows introducing a priori-information on bi-lateral import and export taxes, taxes on consumption for the split-up products as well as production taxes, factor taxes and subsidies for the split-up activities. If these are not provided, the rates for the products resp. sectors to split-up are used.

In order to balance indirect taxes by the different Armington agents, new parameters and variables are introduced:  $p\_dm0$  and  $p\_mm0$  are starting entries for the domestic and import demands to split, and  $v\_dm$  and  $v\_mm$  are the related variables subject to balancing. For case where neither case is relevant, the original data entries as read from GTAPAgg, i.e.  $x\_dm0$  and  $x\_mm0$ , enter the RHS of the equation:



```

*
* --- indirect taxes from given tax rates and estimated domestic and import use
*      (excluding aggregate activities subject to a split)
*
e_indtx(rNat,aa) $(p_sam0(rNat,"indtx",aa) $(not sum(as_a(splita,a) $ sameas(a,aa),1))) ..
    v_sam(rNat,"indtx",aa) $ sum(sameas(aa,splita),1)
+ (sam0(rNat,"indtx",aa) + v_samSlack0P(rNat,"indtx",aa) - v_samSlack0N(rNat,"indtx",aa)) $(not sum(sameas(aa,splita),1))
*
*      (that is solved perhaps somewhat curious to keep the model linear
*      we multiply the endogenous quantity v_dm with a given tax rate dintx
*
*      --- entries for split up products or split up activities
*
=E= sum(i $(p_dm0(rNat,i,aa) $ p_dintx(rNat,i,aa)), p_dintx(rNat,i,aa) * v_dm(rNat,i,aa)
    + v_dintx(rNat,i,aa) * p_dintx(rNat,i,aa))
+ sum(i $(p_mm0(rNat,i,aa) $ p_mintx(rNat,i,aa)), p_mintx(rNat,i,aa) * v_mm(rNat,i,aa)
    + v_mintx(rNat,i,aa) * p_mintx(rNat,i,aa))
*
*      --- taxes linked to non-split up products or not split-up activities
*
+ sum(i $(not p_dm0(rNat,i,aa)), p_dintx(rNat,i,aa) * xdm0(rNat,i,aa))
+ sum(i $(not p_mm0(rNat,i,aa)), p_mintx(rNat,i,aa) * xmm0(rNat,i,aa));

```

The solution shown in the GMAS requires some commenting upon. In order to use CPLEXD as an efficient solver, the constraints are kept linear, i.e. we do not define a product of the split-up transaction such as the demand for the domestic origin  $v_{dm}$  multiplied with the related tax rate  $v_{dintx}$ . Instead, the a-priori chosen rate (either supplied by the GTAP SAM or introduced by the user)  $p_{dintx}$  is multiplied with the endogenous transaction  $v_{dm}$  and endogenous correction to the tax rate  $v_{dintx}$  is multiplied with the a priori quantity  $p_{dm0}$ . Exhaustion of the given tax income reported in the GTAP SAM is ensured by an additional equation which split up the domestic demand at agents' prices:

```

*
* --- exhaustion of domestic demand by agent at agent prices (including indirect taxes) for split-up product
*      (only required if tax rates are endogenous, otherwise e_xdmIAA will automatically also ensure consistent outcomes for xda0 )
*
e_dintx(rNat,armSplit(i,aa)) $(xda0(rNat,i,aa) $(xda0(rNat,i,aa) ne xdm0(rNat,i,aa))
    and ( sum(is_i(j,i), (v_dintx.range(rNat,j,aa) ne 0)
    + sum(as_a(splita,a) $(p_mm0(rNat,j,splita) $ sameas(a,aa)), (v_dintx.range(rNat,j,splita) ne 0)))))) ..
    xda0(rNat,i,aa) =E=
    sum(is_i(j,i),
*
*      --- these are agents not subject to a split
*
    (1+p_dintx(rNat,j,aa)) * v_dm(rNat,j,aa) $ p_dm0(rNat,j,aa)
+ v_dintx(rNat,j,aa) * p_dintx(rNat,j,aa) $ p_dm0(rNat,j,aa)
*
*      --- add the split-up activities
*
+ sum(as_a(splita,a) $(p_dm0(rNat,j,splita) $ sameas(a,aa)),
    (1+p_dintx(rNat,j,splita)) * v_dm(rNat,j,splita)
+ v_dintx(rNat,j,splita) * p_dintx(rNat,j,a));

```

An identical equation is found for the taxes on the imported origin.

To avoid that tax rates become implausible high or low, two additional inequalities restrict their value relative to the tax base:

```

e_dintx0(rNat,i,aa) $(p_dm0(rNat,i,aa) $(v_dintx.range(rNat,i,aa) ne 0) $ p_dintx(rNat,i,aa)) ..
    v_dintx(rNat,i,aa) * p_dintx(rNat,i,aa) + v_dm(rNat,i,aa) * p_dintx(rNat,i,aa)
=L= v_dm(rNat,i,aa) * max(0.75, p_dintx(rNat,i,aa));
e_dintx1(rNat,i,aa) $(p_dm0(rNat,i,aa) $(v_dintx.range(rNat,i,aa) ne 0) $ p_dintx(rNat,i,aa)) ..
    v_dintx(rNat,i,aa) * p_dintx(rNat,i,aa) + v_dm(rNat,i,aa) * p_dintx(rNat,i,aa)
=G= v_dm(rNat,i,aa) * min(-0.75, max(-0.99, p_dintx(rNat,i,aa)));

```

Similar equations imposing bounds on endogenous tax rates are present for factor taxes and subsidies, production taxes and import and export taxes.

The global SAM does not comprise information of the domestic versus imported use of the different Armington agents, see also above. Maintaining the balance here requires additional equations. The first one makes sure that the total use of product  $i$  by an Armington agent  $aa$  – new and old commodities and sectors are considered - is exhausted by domestic and imported consumption:

$$v\_sam_{r,i,aa} = v\_dm_{r,i,aa} + v\_mm_{r,i,aa} \quad (80)$$

In GAMS:

```
*
* --- total Armington demand from domestic and imported
*
e_arm(rNat,consider(i,aa)) $ ( (v_sam.range(rNat,i,aa) ne 0) $ p_sam0(rNat,i,aa)) ..
    v_sam(rNat,i,aa)
    =E=      v_dm(rNat,i,aa) $ p_dm0(rNat,i,aa)
            + v_mm(rNat,i,aa) $ p_mm0(rNat,i,aa);
```

The import and export tax equations are defined as follows, where the first equation defines all import taxes related to split-up products and the second exhausts for each bi-lateral flow the import taxes comprised in auxiliary matrices. Note that the definition during the solve of the parameter `imptxY0` for the products to split-up – i.e. the given values – refers to the total transaction, while it refers for split-up products to the tax rates. The endogenous tax rate is defined implicitly from two terms: the targeted tax rates as a constant multiplied with endogenous bi-lateral flow and a correction terms `v_imptx`.

```
*
* --- import taxes from for split-up product
*
e_imptx(rNat,spliti) $ (p_sam0(rNat,"imptx",spliti) $ (v_sam.range(rNat,"imptx",spliti) ne 0)) ..
    v_sam(rNat,"imptx",spliti)
    =E= sum(r0 $ p_sam0(rNat,r0,spliti),
            v_sam(rNat,r0,spliti) * imptxY0(r0,rNat,spliti) + v_imptx(r0,rNat,spliti) );
```

Three equations ensure that endogenous import taxes maintain plausible rates:

```
*
* --- and max and min imports taxes, required only of import tax rates are endogenous
*
e_imptxMx(rNat,rNat1,spliti) $ (sum(is_i(spliti,i),imptxY0(rNat1,rNat,i))
    $ p_sam0(rNat,rnat1,spliti) $ (v_imptx.range(rNat1,rNat,spliti) ne 0)) ..
    ( v_sam(rNat,rNat1,spliti) + 1.E-3) * imptxY0(rNat1,rNat,spliti) + v_imptx(rNat1,rNat,spliti)
    =L= ( v_sam(rNat,rNat1,spliti) + 1.E-3) * max(0.5,imptxY0(rNat1,rNat,spliti)*1.1,
        min(sum(is_i(spliti,i),(imptxY0(rNat1,rNat,i)/sam0(rnat,rnat1,i)+0.1)*5),
            sum(is_i(spliti,i),imptxY0(rNat1,rNat,i)/sam0(rnat,rnat1,i)+5)));

e_imptxCond(rNat,rNat1,spliti) $ (sum(is_i(spliti,i),imptxY0(rNat1,rNat,i))
    $ p_sam0(rNat,rnat1,spliti) $ (v_imptx.range(rNat1,rNat,spliti) ne 0)) ..
    v_sam(rNat,rNat1,spliti) * imptxY0(rNat1,rNat,spliti) + v_imptx(rNat1,rNat,spliti)
    =L= v_sam(rNat,rNat1,spliti)*100;

e_imptxMn(rNat,rNat1,spliti) $ (sum(is_i(spliti,i),imptxY0(rNat1,rNat,i))
    $ p_sam0(rNat,rnat1,spliti) $ (v_imptx.range(rNat1,rNat,spliti) ne 0)) ..
    v_sam(rNat,rNat1,spliti) * imptxY0(rNat1,rNat,spliti) + v_imptx(rNat1,rNat,spliti) =G= 0;
```

Besides the exhaustion condition for the original cells, the column and row sums for the newly introduces commodities and activities must be equal by definition:

$$v\_rowSum_{r,spliti} = \sum_{is} v\_sam_{r,is,spliti} \quad (81)$$

$$v\_rowSum_{r,spliti} = \sum_{is} v\_sam_{r,spliti,is} \quad (82)$$

$$v\_rowSum_{r,splita} = \sum_{is} v\_sam_{r,is,splita} \quad (83)$$

$$v\_rowSum_{r,splita} = \sum_{is} v\_sam_{r,splita,is} \quad (84)$$

The row sum variables for split-up products are substituted out, while they are maintained for split-up activities as they are used in other equations. One could naturally delete the `v_rowSum` variable altogether which would save equations and variables in the overall split program. It turned however useful in certain cases to control the sums, e.g. if information for the production value of a split-up activities is to be closely maintained.

```
*
* --- sum of column entries for split products must be equal to related row sum
*
e_samBalI(rNat,splitI) $ sum(js,p_sam0(rNat,js,splitI)) ..
    sum(js $ p_sam0(rNat,js,splitI), v_sam(rNat,js,splitI)) =E= sum(js $ p_sam0(rNat,splitI,js), v_sam(rNat,splitI,js));
*
* --- same, for split activities
*
e_samBalA(rNat,splitA) $ sum(js,p_sam0(rNat,js,splitA)) ..
    sum(js $ p_sam0(rNat,js,splitA), v_sam(rNat,js,splitA)) =E= sum(js $ p_sam0(rNat,splitA,js), v_sam(rNat,splitA,js));
```

Further equations ensure that f.o.b. price plus margins are equal to the c.i.f. price, assuming that the per unit margin  $xmarg0$  is the identical between the split-up commodities. Note that the order of the regional indices is reverted: the LHS has the exporter region  $rNat$  on the first index to depict the value at f.o.b. prices, the RHS has the importer index  $rp$  first to depict the value at c.i.f.:

$$v\_sam_{r,splitI,r0} = v\_sam_{r0,r,splitI} (1 + \sum_{splitI \in m} \frac{xmarg0_{m,r0,i,r}}{sam0_{r0,i,r}}) \quad (85)$$

In GAMS:

```
*
* --- Bi-lateral trade flows: fob + marg = cif
*
e_trade(rNat,splitI,rp) $ ( (p_sam0(rp,rNat,splitI) or p_sam0(rNat,splitI,rp))
    $ ( (v_sam.range(rNat,splitI,rp) ne 0) or (v_sam.range(rp,rNat,splitI) ne 0) )) ..
    v_sam(rNat,splitI,rp) * ( 1 + sum( (is_i(splitI,i),m) $ sam0(rNat,i,rp),xmarg0(m,rNat,i,rp)/sam0(rNat,i,rp)) ) )
    =E= v_sam(rp,rNat,splitI);
*
```

Equally, the given imported  $xmm0$  and domestic  $xdm0$  use for each product at market prices in the original SAM for the Armington agents needs to be exhausted:

$$x\_mm0_{r,i,aa} = \sum_{splitI \in i} v\_mm_{r,splitI,aa} + \sum_{splitA \in aa} v\_mm_{r,i,splitA} \quad (86)$$

$$x\_dm0_{r,i,aa} = \sum_{splitI \in i} v\_dm_{r,splitI,aa} + \sum_{splitA \in aa} v\_dm_{r,i,splitA} \quad (87)$$

In GAMS:

```
*
* --- split up of import and domestic demand by agent
*   exclude cases where the split up is determined by other equations
*
e_xmmIAA(rNat,armSplit(i,aa)) $ (xmm0(rNat,i,aa) $ (not ((not SplitI(i)) and sum(sameas(splitA,aa),1) )))) ..
    xmm0(rNat,i,aa)
    =E= sum(is_i(j,i), v_mm(rNat,j,aa) $ p_mm0(rNat,j,aa)
        + sum(as_a(splitA,a) $ (p_mm0(rNat,j,splitA) $ sameas(a,aa)), v_mm(rNat,j,splitA)));
e_xdmIAA(rNat,armSplit(i,aa)) $ (xdm0(rNat,i,aa) $ (not ((not SplitI(i)) and sum(sameas(splitA,aa),1) )))) ..
    xdm0(rNat,i,aa)
    =E= sum(is_i(j,i), v_dm(rNat,j,aa) $ p_dm0(rNat,j,aa)
        + sum(as_a(splitA,a) $ (p_dm0(rNat,j,splitA) $ sameas(a,aa)), v_dm(rNat,j,splitA)));
```

Similar conditions hold for the demand of the domestic and imported origin at market prices and which take indirect taxes into account:

$$xma0_{r,i,aa} = \sum_{j \in i} v_{mmr,j,aa} (1 + p_{mintx_{r,j,aa}}) + v_{mintx_{r,j,aa}} + \sum_{splita \in aa} v_{mmr,j,splita} (1 + p_{mintx_{r,j,splita}}) + v_{mintx_{r,j,splita}} \quad (88)$$

$$xda0_{r,i,aa} = \sum_{j \in i} v_{dmr,j,aa} (1 + p_{dintx_{r,j,aa}}) + v_{dintx_{r,j,aa}} + \sum_{splita \in aa} v_{dmr,j,splita} (1 + p_{dintx_{r,j,splita}}) + v_{dintx_{r,j,splita}} \quad (89)$$

Based on this, indirect taxes revenues for each agent are defined:

$$sam0_{r,"indtx",aa} = \sum_{j \in i} v_{mmr,j,aa} (1 + p_{mintx_{r,j,aa}}) + v_{mintx_{r,j,aa}} + \sum_{j \in i} v_{dmr,j,aa} (1 + p_{dintx_{r,j,aa}}) + v_{dintx_{r,j,aa}} \quad (90)$$

The imported use for the newly introduced commodities, added up over the Armington agents, must be equal to the bi-lateral imports at c.i.f. prices plus import taxes:

$$\sum_{aa} v_{mmr,spliti,aa} = \sum_{r0} v_{samr,r0,spliti} (1 + imptxyY0_{r0,r,spliti} + v_{imptx_{r0,r,spliti}}) \quad (91)$$

Note there that  $v_{imptx}$  is an endogenous correction of the tariff revenues. In GAMS:

```
*
*      - - - - -
*      -- total imported use from imports
*
e_imp(rNat,spliti) ..
*
*      -- imported use
*
sum(aa $ p_mm0(rNat,spliti,aa), v_mm(rNat,splitI,aa))
=E=
sum(r0 $ p_sam0(rNat,r0,spliti),
v_sam(rNat,r0,splitI) * (1 + imptxyY0(r0,rNat,spliti)) + v_imptx(r0,rNat,spliti));
.
```

Finally, the market balance for the new commodities must hold:

$$\sum_{aa} v_{dmr,spliti,aa} + \sum_{r0} v_{samr,spliti,r0} - v_{samr,"exptx",spliti} = \sum_{splita \in aa} v_{samr,splita,spliti} \quad (92)$$

In GAMS:

```
e_mkt(rNat,spliti) ..
*
*      -- domestic use
*
[ sum(aa $ p_dm0(rNat,spliti,aa), v_dm(rNat,splitI,aa))
*
*      -- exports
*
+ sum(r0 $ p_sam0(rNat,spliti,r0), v_sam(rNat,splitI,r0))
*
*      -- exports taxes
*
- v_sam(rNat,"exptx",spliti) $ p_sam0(rNat,"exptx",spliti)
]
*
*      -- output
*
=E= sum(splita $ p_sam0(rNat,splita,spliti), v_sam(rNat,splita,spliti));
```

An inequality keeps at least a 25% share of targeted total domestic use related to targeted production (but with a limit of a 75% target):

```

*
*
e_domUse(rNat,spliti) $ (sum(aa,p_dm0(rNat,spliti,aa)) and sum(splita,p_sam0(rNat,splita,spliti))) ..
*
sum(aa $ p_dm0(rNat,spliti,aa), v_dm(rNat,splitI,aa))

=G= sum(splita $ p_sam0(rNat,splita,spliti), v_sam(rNat,splita,spliti))
* 0.25 * min(0.75,sum(aa,p_dm0(rNat,spliti,aa))/sum(splita,p_sam0(rNat,splita,spliti)));

```

It turned out useful to introduce bounds for factor use cost shares:

```

*
* --- plausibility bounds on cost shares
*
e_ShareLo(rNat,consider(isi,splita)) $ (p_sam0(rNat,isi,splita)
$ (v_sam.range(rNat,isi,splita) ne 0)
$ (p_ShareLo(rNat,isi,splita) ne 0)) ..

p_shareLo(rNat,isi,splita) * v_rowSum(rnat,splita) =L= v_sam(rNat,isi,splita);

e_ShareFlo(rNat,consider(f,splita)) $ (p_sam0(rNat,f,splita)
$ (v_sam.range(rNat,f,splita) ne 0)
$ (p_ShareFlo(rNat,f,splita) ne 0)) ..

p_shareFlo(rNat,f,splita) * sum(fAl,v_sam(rnat,fAl,splita)) =L= v_sam(rNat,f,splita);

e_ShareUp(rNat,consider(isi,splita)) $ ((v_sam.range(rNat,isi,splita) ne 0) $ (p_ShareUP(rNat,isi,splita) ne 0)
$ (not propAssmpt(rnat,isi,splita))
$ (p_ShareUP(rNat,isi,splita) ne 1) $ p_sam0(rNat,isi,splita)) ..

p_shareUP(rNat,isi,splita) * v_rowSum(rnat,splita) =G= v_sam(rNat,isi,splita);

e_minPropAssmpt(rnat,splitA) $ sum(js,p_sam0(rNat,js,splitA)) ..

v_propAssmpt(rnat,splita) =G= v_rowSum(rnat,splita) / sum((isi,as_a(splita,a)),sam0(rNat,isi,a))*0.2;

```

The objective function minimizes absolute (normalized) deviation between the desired new SAM entries  $p\_sam0$  and the endogenous entries  $v\_sam$  which fit into the equation system defined above. In order to do so, we need to define separate variables for the positive and negative deviations:

```

e_samSlack(rNat,considerHPD(isi,jsj)) $ ((not sam0(rnat,isi,jsj)) $ p_samCor0(rnat,isi,jsj) $ (not propAssmpt(rNat,isi,jsj))
$ (mapVal(p_wgts(rnat,isi,jsj)) ne 8)
$ (v_sam.range(rnat,isi,jsj) ne 0) $ p_sam0(rNat,isi,jsj)) ..

v_sam(rNat,isi,jsj) + v_samSlackP(rNat,isi,jsj) - v_samSlackN(rNat,isi,jsj) =E= p_samCor0(rnat,isi,jsj);

```

These slack variables enter the objective function with weights:

```

--- SAM entries
+ sum( (rNat,considerHPD(isi,jsj)) $ (p_samCor0(rNat,isi,jsj) $ (mapval(p_wgts(rnat,isi,jsj)) ne 8) $ (not propAssmpt(rNat,isi,jsj))),
(v_samSlackP(rNat,isi,jsj)+v_samSlackN(rNat,isi,jsj))

--- we decrease the relative weight for trade to avoid that countries with overall tiny weight drive
the trade results for large ones

/(abs(p_sam0(rNat,isi,jsj))+p_relDamp(rNat)
+ (p_relDamp("wor")-p_relDamp(rNat)) $ (isNation(isi) or isNation(jsj))) * (1+p_wgts(rnat,isi,jsj)))

```

Along with similar entries for the tax rates. To avoid that tiny imbalances left in the global SAM to split up prevent finding a solution, tiny slacks for the production activity related SAM entries to split up are introduced.

As detailed below,  $p\_sam0$  and  $p\_samCor0$  are identical for larger starting values of the split-up SAM entries, but changed to favor sparsity for very small ones. The objective function minimizes squared relative differences, in order to prevent that deviations for quite tiny entries are too important, the normalization adds one dollar to the absolute value of the starting value  $p\_sam0$ . The second term in the LHS only scales the objective function to improve overall scaling, it counts the non-zero entries and divides the sum by factor ten.

#### 4.3.16 Proportionality assumptions for cost structures during splitting

For many dis-aggregation exercises of SAMs, a-priori information is available for selected intermediate input and factor demands, only. Split factors for the remaining cost items are then derived

based on proportionality assumptions. These do not carry additional information. It would therefore make sense to set their weight in the objective function to zero. The split factors where a priori information is available should then via the balancing conditions adjust the non-controlled ones. However, this likely gives the balancing framework unwanted degrees of freedom. Many different combinations of only indirectly controlled input demands might balance the problem at unchanged penalties. Technically, the problem will have a degenerate solution.

Using small penalties for costs shares derived from proportionality assumptions might help to avoid degeneracy, but especially for small absolute estimates, might provoke numerical challenges. Additionally, penalizing deviation for costs share derived from proportionality assumptions is likely to lead to some violation of proportionality conditions. This reflects that is typically impossible to derive completely consistent a-priori information. If this would be possible, the whole balancing exercise would not be necessary.

Therefore, proportionality assumptions are explicitly controlled in here by two constraints. Let *propAssmpt* define a three-dimensional indicator set of region, input or factor demand and activity combinations derived from to proportionality assumptions. For these cases, the following conditions are enforced:

$$v\_sam_{r,isi,splita} = v\_propAssmpt_{r,splita} \sum_{splita \in a} sam0_{r,isi,a} \bigwedge propAssmpt_{rnat,isi,splita} \quad (93)$$

$$v\_sam_{r,spliti,a} = v\_propAssmpt_{r,spliti} \sum_{spliti \in ai} sam0_{r,i,a} \bigwedge propAssmpt_{rnat,spliti,a} \quad (94)$$

Where *v\_propAssmpt* are endogenously determined distribution factors which ensure that all cost shares included in *propAssmpt* are receive the same share on the total. These factors are not entering the objective function. Equally, entries relating to input demands controlled by proportionality assumptions are removed from the objective.

As an example, consider the crop activities where FAOSTAT delivers a-priori information on production, land demands and seed use. The difference for a disaggregated crop activity between its share on production and its share on land and seed use of the aggregate activity should be its share on all other costs. These other cost items shall be distributed on proportionality assumptions. The equation above will ensure that the relations between these other cost items do not change. For instance, while the relation between capital and land will change for a dis-aggregated crop – reflecting a different cost share of land based on yield information from FAOSTAT–, the relation between capital and a labour category will be identical to the aggregate activity for an example), reflecting the constraint above. A lower yield, such as in case of rye in the example below, will imply a higher cost share of land and lower cost shares of other factors (and intermediate demand, no shown in the table).

The user can therefore define a set *propAssmpt(rnat,isi,splita)*. All elements on *isi* that are not primary factors are than subject to the same split factor on the related total. For primary factors, a separate common split factor is defined. For the largest flow, slacks are introduced to avoid infeasibilities stemming on proportionality assumptions.

```
*
* --- Same splitfactor for pre-defined lists of input demands (there is slack for the activities with the largest
*      a-priori demand)
*
e_propAssmptA(propAssmpt(rNat,isi,splita)) $ (p_sam0(rnat,isi,splita) $ (v_sam.range(rnat,isi,splita) ne 0)) ..
      v_sam(rNat,isi,splita) =E= ( v_propAssmpt(rnat,splita) $ (not f(isi))
      + v_propAssmptF(rnat,splita) $ f(isi) ) *sum(as_a(splita,a),sam0(rnat,isi,a))
      + (v_samSlackP(rNat,isi,splita) - v_samSlackN(rNat,isi,splita)) $ largestPropAssmpt(rnat,isi,splita);
```

A further example might help to understand the logic of the equations. Assume that you don't have information on the how exactly the service shares differ across split-up activities. Let's assume that the list of service products is called *iServ*. You activate the proportionality assumptions by settings *propAssmpt(rnat,iServ,splita) = yes*. Let's assume that the first split-up activity receives 10% of the intermediate demand of the related original activity to split up. The equations above will then also make sure that it receives 10% of any other products from the *iServ* set. The cost-share relations between the list of products will be kept. The activity might still receive in total a higher or lower cost share of services as the related aggregate, say, by being more capital intensive. This is possible as the equation only refers to the composition of the total service demand implicitly defined by the elements on *iServ*. As seen, the list of individual elements where cost-share relations are kept are activity specific. Clearly, other constraints in the system will enforce that the reported cost position for the aggregated activity will be distributed to the newly introduced aggregate ones. These constraints interact with the ones depicting proportionality assumptions.

In a similar manner, it is possible to enforce equal distribution of split-up products over non-split up activities. Assume, for instance, that you split-up chemicals into fossil fuels and biomass based on some output statistics, but you have no information about the use shares of the two types of chemicals in the different sectors. The following equations allow to enforce equal shares across a list of sectors:

```
*
* --- Same splitfactor for pre-defined lists of split-up input demands of not-split up activities
*
e_propAssmptI(propAssmpt(rNat,spliti,a)) $ (p_sam0(rnat,spliti,a) $ (v_sam.range(rnat,spliti,a) ne 0) $ (not splita(a))) ..
v_sam(rNat,spliti,a) =E= v_propAssmpt(rnat,spliti) * sum(is_i(spliti,i),sam0(rnat,i,a))
+ (v_samSlackP(rNat,spliti,a) - v_samSlackN(rNat,spliti,a)) $ largestPropAssmpt(rnat,spliti,a);
```

If proportionality assumptions are used, we also keep the relation between imported and exported use across the cases:

```
e_propAssmptID(propAssmpt(rNat,spliti,a)) $ (p_sam0(rnat,spliti,a) $ p_dm0(rnat,spliti,a) $ p_mm0(rnat,spliti,a)
$ (v_dm.range(rnat,spliti,a) ne 0)
$ (v_sam.range(rnat,spliti,a) ne 0) $ (not splita(a))) ..
v_dm(rNat,spliti,a) =E= v_propAssmpt(rnat,spliti) * sum(is_i(spliti,i),xdm0(rnat,i,a))
+ (v_dmSlackP(rNat,spliti,a) - v_dmSlackN(rNat,spliti,a)) $ largestPropAssmpt(rnat,spliti,a);
```

In cases where we split an activity, we have usually no information on differences in domestic and imported use shares of intermediate demand. We might know that a split-up crop activity has a higher fertilizer use compared to another, but it is unlikely that we also know if this implies less or more domestic versus imported fertilizers are used in the total fertilizer use of the two variants. Two equations therefore fix the composition of the imported and domestic use for non-split up products:

```
*
* --- keep import / domestic shares for not-split up input
*
* (1) Only the equation for the larger of the domestic or imported origin will be active
* to avoid linear dependencies with the equation for the sum of the two
*
* (2) For the same reason, if only one of the two positions is reported, drop both equations
*
e_sameDomShare(rNat,i,splita) $ (p_sam0(rnat,i,splita) $ (p_dm0(rnat,i,splita) gt p_mm0(rnat,i,splita))
$ p_mm0(rnat,i,splita) $ p_dm0(rnat,i,splita) $ (v_dm.range(rnat,i,splita) ne 0)
$ (v_sam.range(rnat,i,splita) ne 0) $ (not spliti(i))) ..
v_dm(rNat,i,splita) =E= v_sam(rnat,i,splita)
* sum(as_a(splita,a) $ sam0(rnat,i,a),xdm0(rnat,i,a)/sam0(rnat,i,a))
+ (v_dmSlackP(rNat,i,splita) - v_dmSlackN(rNat,i,splita)) $ largestPropAssmpt(rnat,i,splita);
e_sameImpShare(rNat,i,splita) $ (p_sam0(rnat,i,splita) $ (p_mm0(rnat,i,splita) ge p_dm0(rnat,i,splita))
$ p_mm0(rnat,i,splita) $ p_dm0(rnat,i,splita) $ (v_mm.range(rnat,i,splita) ne 0)
$ (v_sam.range(rnat,i,splita) ne 0) $ (not spliti(i))) ..
v_mm(rNat,i,splita) =E= v_sam(rnat,i,splita)
* sum(as_a(splita,a) $ sam0(rnat,i,a),xmm0(rnat,i,a)/sam0(rnat,i,a))
+ (v_mmSlackP(rNat,i,splita) - v_mmSlackN(rNat,i,splita)) $ largestPropAssmpt(rnat,i,splita);
```



#### 4.3.17 Preparation of the split factors

The minimum information required to perform a split are split factors for the production values. These factors do not need to add up to the unity, one could also enter totals as the GAMS code will scale the sum of the split factors for each SAM cell to unity.

If no split factors for split-up products the four following demand categories are given:

```
set dem(is) / hhsld,gov,inv,int /;
```

The program will use the given split factor for the output the activity producing the split-up product instead:

```
*
* --- if split information of final demand category is missing, use output factor instead
*
  p_splitFactor(rNat,spliti,dem) $ (not sum(is_i(spliti1,i) $ is_i(spliti,i), p_splitFactor(rNat,spliti1,dem)))
    = sum(i_a(spliti,splitA), p_splitFactor(rNat,splitA,"prod"));
*
* --- if splitfactor is still missing, assume equal shares
*
  p_splitFactor(rNat,spliti,dem) $ (not sum(is_i(spliti1,i) $ is_i(spliti,i), p_splitFactor(rNat,spliti1,dem))) = 1;
```

Split factors below 1.E-5 (i.e. a share < 0.001%) will be deleted; however, “eps” entries are maintained that imply that the split factor should be zero in the final solution:

```
p_splitFactor(rNat,splitI,is) $ ((abs(p_splitFactor(rNat,splitI,is)) le 1.E-5)
                                $ (p_splitFactor(rNat,splitI,is) ne eps)) = 0;
```

Afterwards, the non-zero split factors for the demand categories are scaled to unity:

```
p_splitFactor(rNat,splitI,is) $ ( (p_splitFactor(rNat,splitI,is) ne eps) ^ p_splitFactor(rNat,splitI,is))
    = p_splitFactor(rNat,splitI,is) / sum(is_i(spliti1,i) $ is_i(spliti,i), p_splitFactor(rNat,spliti1,is));
```

These two statements are repeated to decrease the chance of very tiny split factors. That logic is also applied for split-factors on output “*prod*”, and for intermediate input use and output of commodities. Attention: the order of column and rows in the split factors are reverted compared to the SAM to ease their handling in calling programs, i.e. `p_splitFactor(rNat,splitA,is)` refers to the use of row *is* in the activity *splitA*.

Split factors are used as follows. Explicit split factors for a specific SAM cell (intermediate input or factor demand by an activity, output coefficients, demand positions, export flows) are given precedence. If no split-factor for an individual intermediate input demand is found, a split factor for general intermediate demand (“*int*”) for the newly introduced product is used. For intermediate input demands of split-up products by split-up activities, the product of the “*int*” and “*prod*” factors is used if no explicit factor is present. If a split factor for an intermediate input demand is missing, the split factor for production (“*prod*”) of the activity demanding the input is used instead. For individual outputs of a new activity, if no explicit split factor is given, the split factor for production “*prod*” is used as well. The same holds if factor demands of new activities are missing. During these assignments, split factors representing original zeros in the underlying database are not overwritten. As discussed below in next section, split-factors constructed for cost shares where no original a-priori information is available have no impact on the results.

After these assignments, a check tests for each SAM cell to disaggregate if related split factors are present. Missing split factors can result from removing small shares or from “true” zeros in the original data. If no split-factors are found for a SAM cell to be disaggregated, the aggregated SAM



value is distributed in equal shares to the new SAM cells. Afterwards, split factors for any SAM cell to disaggregate are guaranteed to be present. A-priori estimates of new SAM cells to introduce are generated by multiplying their split factor with the related aggregate SAM cell. All new SAM cells are afterwards scaled to exhaust the related aggregate ones.

A challenge provides the decomposition of the individual demand positions to the domestic and imported origin. Using the shares for the related aggregate product for newly introduced demands can lead to inconsistent a-priori information which does not match the a-priori data on disaggregated bi-lateral trade and domestic production. For instance, import demand of products in a region might be created where no imports are reported, or domestic demand where no output is produced. To address this, the domestic and import shares for all Armington demands for the aggregate products are calculated first. This step produces import and domestic demand shares in sum of all Armington agents  $aa$  for each national market  $r$  for the aggregated products to split  $i$ :

$$\text{impShare}_{r,i} = \frac{\sum_{aa} \text{xm0}_{r,i,aa}}{\sum_{aa} \text{sam}_{r,i,aa}} \quad (95)$$

$$\text{domShare}_{r,i} = 1 - \text{impShare}_{r,i} \quad (96)$$

Symbols with a  $0$  at the end, such as  $\text{sam0}$ , refer to the benchmark or a priori values.  $\text{xm0}$  refers to import and  $\text{xdm0}$  to domestic demand. Endogenous variables under control of the split balancing framework start with a  $v\_$ . Accordingly,  $v\_sam$  refers to the newly introduced, disaggregated SAM cells.

Next, imports for disaggregated products  $\text{spliti}$  into region  $r$  from the exporter regions  $r0$  are set in relation to their targeted total demands (domestic and imported) to arrive at import share estimates for the disaggregated products  $\text{spliti}$  in this region  $r$ :

$$\text{impShare}_{r,\text{spliti}} = \frac{\sum_{r0} v\_sam_{r,r0,\text{spliti}} (1 + p\_imptx_{r0,r,\text{spliti}})}{\sum_{aa} v\_sam_{r,\text{spliti},aa}} \quad (97)$$

where  $\text{spliti}$  are the newly, more disaggregated products and  $p\_imptx$  the bilateral import tax rates applied on the import flows at c.i.f. value.

If domestic output is observed, the domestic share is determined as the residual, with a security threshold of 1%. This is important for cases where the a-priori estimates for imports exceed the total targeted demand:

$$\text{domShare}_{r,\text{spliti}} = \max(0 + 0,01 \exists \text{output}_{r,\text{spliti}}, 1 - \text{impShare}_{r,\text{spliti}}) \quad (98)$$

Due to the maximum operator and the fact that domestic shares must become zero if no domestic output is reported, the import share has to be re-calculated afterwards:

$$\text{impShare}_{r,\text{spliti}} = 1 - \text{domShare}_{r,\text{spliti}} \quad (99)$$

Based on these shares, relations between the aggregate and disaggregated import and domestic use shares are defined for the disaggregated products:

$$\text{impRel}_{r,\text{spliti}} = \frac{\text{impShare}_{r,\text{spliti}}}{\sum_{i \in \text{spliti}} \text{impShare}_{r,i}} \quad (100)$$

$$\text{domRel}_{r,\text{spliti}} = \frac{\text{domShare}_{r,\text{spliti}}}{\sum_{i \in \text{spliti}} \text{domShare}_{r,i}} \quad (101)$$

This allows the calculation of a scaling factor  $\text{scaleArm}$  which ensures that the SAM entries to disaggregate will be exhausted when the distribution factors for the domestic and imported origin calculated above are applied:

$$\text{scaleArm}_{r,\text{spliti},aa} = \sum_{i \in \text{spliti}} \frac{\text{xdm0}(\text{rnat}, i, aa) \text{domRel}_{r,\text{spliti}} + \text{xmm0}(\text{rnat}, i, aa) \text{impRel}_{r,\text{spliti}}}{\text{sam0}_{r,i,aa}} \quad (102)$$

Afterwards, the reported aggregated import and domestic demand shares for a product and an Armington agent are multiplied with the demand for the split-up product, as derived from the split factors ( $\text{v\_sam}_{r,\text{spliti},aa}$ ), and with the relation between the disaggregated and aggregated import respectively domestic shares, as well as with the scaling factor:

$$\text{v}_{mm,r,\text{spliti},aa} = \text{v}_{sam,r,\text{spliti},aa} \sum_{i \in \text{spliti}} \frac{\text{xmm0}_{r,i,aa}}{\text{sam0}_{r,i,aa}} \text{impRel}_{r,\text{spliti}} \text{scaleArm}_{r,\text{spliti},aa} \quad (103)$$

$$\text{v\_dm}_{r,\text{spliti},aa} = \text{v\_sam}_{r,\text{spliti},aa} \sum_{i \in \text{spliti}} \frac{\text{xdm0}_{r,i,aa}}{\text{sam0}_{r,i,aa}} \text{domRel}_{r,\text{spliti}} \text{scaleArm}_{r,\text{spliti},aa} \quad (104)$$

$\text{v\_mm}$  refers to the disaggregated import demands by Armington agent and  $\text{v\_dm}$  to the disaggregated domestic ones, while  $\text{xdm0}$  and  $\text{xmm0}$  depict the Armington demands at market prices from the filtered GTAP Data Base for the related non-split up products.

Afterwards, further checks and potential changes to the a-priori data are introduced:

- If no domestic or import demand for an Armington demand to disaggregate is found, the given total is assigned and the so-constructed disaggregated values are scaled to exhaust the given total.
- In case where exports are reported, but no import demand (from aggregated import demand over all Armington agents), the export flow is deleted if there are other export flows of disaggregated products relating to the same GTAP product on the same trade link. If export flows cannot be deleted as they are the only ones reported, each Armington demander showing import demand in each import country where such missing export flows are observed gets assigned a very tiny import demand.
- In case of missing production, exports flows and domestic demands are deleted.

Afterwards, indirect, export and factor tax rates are defined from the aggregate transactions and assigned to the disaggregated ones. Bilateral import tariffs are taken from TASTE as long as the necessary scaling factor to align the detailed tariff rate to given tariff revenues is not larger than five or smaller than one fifth. Individual ad-valorem tariff rates which exceed the average rates for a product on a bilateral trade flow by a very large factor are deleted.

After such corrections, the disaggregated data are re-scaled to exhaust the given totals. The resulting data provide the a-priori information for the linear-loss estimator framework (for a technical discussion see Britz 2021 for details) discussed above. During the solution of the split problem, solely weights for the “secure” data (TASTE, production, intermediate demand and land use from FABIO in the Britz 2021) are used, which higher weights for production, bi-lateral trade and land use. The framework also introduces penalties for very small entries for new SAM cells to favour sparsity and comprises constraints to control for reasonable factor cost shares and tax rates.

#### 4.3.18 Split solution process

As seen from the equation structure above, the constraints are kept linear to ease solving the problem. As the objective function is linear, besides CONOPT as the default solver user in CGEBox, also highly performant LP solvers such as CPLEX can be used. That has shown to speed up dramatically splitting up SAMs with many sectors and regions where the rebalancing problem comprises several 10 thousand of variables and equations, and could successfully work on split problems with >2 Mio variables. The QP solver (CONOPT4 or CPLEX) can be chosen on the interface and will be used as the LP solver for the split problem. In order to speed up the solution process, the split factors are used to define starting values for the new SAM entries  $v\_sam$ , for instance:

```
v_sam.l(rNat,splitI,a) = sum(is_i(splitI,i), sam0(rNat,i,a) * p_splitFactor(rNat,splitI,"int"));
```

The same holds for other elements such as domestic and imported use of the Armington agents, for instance:

```
v_mm.l(rNat,splitI,aa) $ sum(is_i(splitI,i), xmm0(rNat,i,aa))
= v_sam.l(rNat,splitI,aa) * sum(is_i(splitI,i), xmm0(rNat,i,aa)/sam0(rNat,i,aa));
v_dm.l(rNat,splitI,aa) $ sum(is_i(splitI,i), xdm0(rNat,i,aa))
= v_sam.l(rNat,splitI,aa) * sum(is_i(splitI,i), xdm0(rNat,i,aa)/sam0(rNat,i,aa));
```

In order to avoid that, for instance, tiny split shares in production lead to extremely small SAM entries in intermediate consumption, entries with a values < 1 USD cent are pulled to zero. That has shown to help with sparsity issues:

```
p_samCor0(rnat,is,js) = p_sam0(rnat,is,js);
*
* --- pull tiny SAM entries to zero (sparsity), units USD - all < 1 cent is pulled towards zero
*
p_samCor0(rNat,consider(is,js)) $ ( (p_sam0(rNat,is,js) gt 0) $ (p_sam0(rNat,is,js) lt 1.E-2)) = -1;
p_samCor0(rNat,consider(is,js)) $ ( (p_sam0(rNat,is,js) lt 0) $ p_sam0(rNat,is,js) $ (p_sam0(rNat,is,js) gt -1.E-2)) = +1;
*
p_samCor0(rNat,consider(is,js)) $ ( (p_sam0(rNat,is,js) gt 0) $ (p_sam0(rNat,is,js) lt 1.E-3)) = -10;
p_samCor0(rNat,consider(is,js)) $ ( (p_sam0(rNat,is,js) lt 0) $ p_sam0(rNat,is,js) $ (p_sam0(rNat,is,js) gt -1.E-3)) = +10;
*
p_samCor0(rNat,consider(is,js)) $ ( (p_sam0(rNat,is,js) gt 0) $ (p_sam0(rNat,is,js) lt 1.E-4)) = -100;
p_samCor0(rNat,consider(is,js)) $ ( (p_sam0(rNat,is,js) lt 0) $ p_sam0(rNat,is,js) $ (p_sam0(rNat,is,js) gt -1.E-4)) = +100;
```

The global SAM is solved several times, changing the option files used by the solver in case of numerical problems:

```
set tries / try1*try10/;
scalar nTries / 0/;
while ( (nTries < card(tries))
and ( ( (m_split.sumInfes > m_split.tolInfeas*m_split.numEqu/1000) and (m_split.sumInfes > 1.E-5))
or ( (m_split.modelstat > 2) and (m_split.modelstat ne 7))
)
)
$batinclude 'util/title.gms' "'Search feasible solution for split, nTry '" nTries:0:0 "' , sumInfes '" m_split.sumInfes
nTries = nTries + 1;
```

Several steps are taken to help with infeasibilities:

1. The option file is changed, option files with a higher number use less accurate feasibility requirement, used if previous solution attempts failed.

```

*
*   --- if the sum of infeasibilities is too high, relax accuracy of solver
*
IF ( (m_split.suminfes > m_split.tolinfeas*m_split.numEqu/1000),
    if ( nTries > 2, m_split.optFile = 6);
    if ( nTries > 4, m_split.optFile = 7);
    if ( nTries > 6, m_split.optFile = 8; m_split.tolinfeas = 1.E-7);
    if ( nTries > 7, m_split.optfile = 9; m_split.tolinfeas = 1.E-6);
    if ( nTries > 8, option LP=CONOPT4;m_split.optFile = 8);
);
*
*   --- same if the solver reports some error
*
if ( (nTries > 2) $ (m_split.modelStat > 2), m_split.optfile = 6);
if ( (nTries > 4) $ (m_split.modelStat > 2), m_split.optfile = 7);
if ( (nTries > 6) $ (m_split.modelStat > 2), m_split.optfile = 8);
if ( (nTries > 7) $ (m_split.modelStat > 2), m_split.optfile = 9);
if ( (nTries > 8) $ (m_split.modelStat > 2), m_split.optfile = 8);

```

2. Requirements that cost shares must fall in a certain range are step-wise relaxed:

```

--- relax lower /upper share constraints if binding

p_shareLo(rNat,is,splita) $ ( ( e_shareLo.l(rNat,is,splita) gt m_split.tolinfeas)
                             or (abs(e_shareLo.m(rNat,is,splita)) gt 1.E-12))
                             and (p_shareLo(rNat,is,splita) ne p_shareup(rNat,is,splita))
                             and (p_shareLo(rNat,is,splita) gt 0))
= p_shareLo(rNat,is,splita) * 0.75;

p_shareLo(rNat,is,splita) $ ( ( p_shareLo(rNat,is,splita) le 1.E-5) $ (p_shareLo(rNat,is,splita) gt 0) and (nTries > 2)) = 0;
p_shareLo(rNat,is,splita) $ ( ( p_shareLo(rNat,is,splita) le 1.E-4) $ (p_shareLo(rNat,is,splita) gt 0) and (nTries > 4)) = 0;
p_shareLo(rNat,is,splita) $ ( ( p_shareLo(rNat,is,splita) gt 0) $ (nTries > 8)) = 0;

p_shareUp(rNat,is,splita) $ ( ( e_shareUp.l(rNat,is,splita) lt -m_split.tolinfeas)
                             or (abs(e_shareUp.m(rNat,is,splita)) gt 1.E-12))
                             and (p_shareLo(rNat,is,splita) ne p_shareup(rNat,is,splita))
                             and (p_shareUp(rNat,is,splita) gt 0))
= max(1,p_shareUp(rNat,is,splita) * 1.25);

p_shareUp(rNat,is,splita) $ ( nTries > 6 ) = 1;

```

3. Corrections to the desired tax rates are step-wise relaxed:

```

--- give some freedom in tax bounds (ar pulled towards zero anyhow in objective)

if ( nTries > 2,
    $$batinclude 'build\split_set_taxBounds.gms' 1.E-3
);
if ( nTries > 4,
    $$batinclude 'build\split_set_taxBounds.gms' 1.E-2
);
if ( nTries > 6,
    $$batinclude 'build\split_set_taxBounds.gms' 1.E-1
);
if ( nTries > 8,
    $$batinclude 'build\split_set_taxBounds.gms' 1
);

```

If the final solution has still a sum of infeasibilities above a certain threshold, the program will abort with an error message:

```

IF ( (m_split.suminfes > m_split.tolinfeas*m_split.numInfes*10),
    execute_unload "%scdir%\split_%dataset%.gdx" v_sam.l,v_dm.l,v_mm.l,v_fShare.l;
    option QCP=conopt;
    solve m_split using QCP minimizing v_hpd;
    abort "Split program could not balance SAM";
);

```

#### 4.3.19 Processing the solution from the split problem

After successful balancing, the information is copied to the SAM (see “*build/split\_reportAndMap.gms*”):

```
--- assing results from dis-aggregated SAM entries to global SAM
SAM0(rNat,consider(is,js)) = v_sam.l(rNat,is,js);
```

The same is done with the domestic and import use vectors, afterwards, the information for split up commodities is deleted:

```
*
* --- assgin imports and domestic sales by agent
*
xnm0(rNat,consider(i,aa)) = v_mm.l(rNat,i,aa);
xdm0(rNat,consider(i,aa)) = v_dm.l(rNat,i,aa);

xnm0(rNat,i,aa) $ sum(is_i(spliti,i), 1) = 0;
xdm0(rNat,i,aa) $ sum(is_i(spliti,i), 1) = 0;

xnm0(rNat,i,a) $ sum(as_a(splita,a), 1) = 0;
xdm0(rNat,i,a) $ sum(as_a(splita,a), 1) = 0;

xma0(rNat,consider(i,aa)) = xnm0(rNat,i,aa) * (1+mintx(rNat,i,aa));
xda0(rNat,consider(i,aa)) = xdm0(rNat,i,aa) * (1+dintx(rNat,i,aa));
```

There are further statements not shown which refer to the CO2 and Non-Co2 emissions factors, assumed currently to be equal across the split-up commodities respectively activities.

Furthermore, behavioral parameters are copied over, as well as taxes, and entries for the aggregate ones deleted.

```
*
* --- assing behavioral parameters and tax rates (and other elements not comprised in SAM, but used by model)
*
sigmaw0(rNat,Spliti) = sum(is_i(spliti,i), sigmaw0(rNat,i));
sigmam0(rNat,Spliti) = sum(is_i(spliti,i), sigmam0(rNat,i));

eh0(rNat,Spliti) = sum(is_i(spliti,i), eh0(rNat,i));
bh0(rNat,Spliti) = sum(is_i(spliti,i), bh0(rNat,i));

sigmap0(rNat,Splita) = sum(as_a(splita,a), sigmap0(rNat,a));
sigmav0(rNat,Splita) = sum(as_a(splita,a), sigmav0(rNat,a));

*
* --- assign factor taxes
*
fcttxY0(rNat,f,splitA) $ sam0(rNat,f,splitA) = sam0(rNat,f,splitA) * sum(as_a(splita,a), fcttxY0(rNat,f,a)/sam0(rNat,f,a));
fcttsY0(rNat,f,splitA) $ sam0(rNat,f,splitA) = sam0(rNat,f,splitA) * sum(as_a(splita,a), fcttsY0(rNat,f,a)/sam0(rNat,f,a));

fcttxY0(rNat,f,A) $ sum(as_a(splita,a),1) = 0;
fcttsY0(rNat,f,A) $ sum(as_a(splita,a),1) = 0;

*
* --- assign import and export taxes
*
imptxY0(rp,rNat,splitI) $ sam0(rNat,rp,splitI) = sam0(rNat,rp,splitI) * sum(is_i(splitI,i), imptxY0(rp,rNat,i)/sam0(rNat,rp,i));
exptxY0(rNat,rp,splitI) $ sam0(rNat,splitI,rp) = sam0(rNat,splitI,rp) * sum(is_i(splitI,i), exptxY0(rNat,rp,i)/sam0(rNat,i,rp));

imptXY0(rp,rNat,i) $ sum(is_i(splitI,i),1) = 0;
exptXY0(rNat,rp,i) $ sum(is_i(splitI,i),1) = 0;
```

Equally, trade margins are copied (not shown). Afterwards, the SAM entries for the split-up commodities and activities are removed and some SAM entries re-calculated to avoid problems from rounding errors in the solution process.

```
*
* --- remove SAM entries referring to aggregated rows and columns
*
sam0(rNat,i,js) $ sum(is_i(spliti,i),1) = 0;
sam0(rNat,is,j) $ sum(is_i(spliti,j),1) = 0;

sam0(rNat,a,js) $ sum(as_a(splita,a),1) = 0;
sam0(rNat,is,a) $ sum(as_a(splita,a),1) = 0;

* --- Government account Taxes

sam0(rNat,gov,"indtx") = sum(js, sam0(rNat,"indtx", js)) ;
sam0(rNat,gov,"prdtx") = sum(a, sam0(rNat,"prdtx", a)) ;
sam0(rNat,gov,"imptx") = sum(i, sam0(rNat,"imptx", i)) ;
sam0(rNat,gov,"exptx") = sum(i, sam0(rNat,"exptx", i)) ;
sam0(rNat,gov,"fcttx") = sum(a, sam0(rNat,"fcttx",a)) ;
sam0(rNat,gov,"fctts") = sum(a, sam0(rNat,"fctts",a)) ;

* --- bilateral bops, bops and trade margins, and bop = foreign saving

sam0(rNat,rp,"bop") = sum(i, sam0(rNat,i,rp)) ;
sam0(rNat,"bop",rp) = sum(i, sam0(rNat,rp,i)) ;
sam0(rNat,"trdmg","bop") = sum(i, sam0(rNat,i,"trdmg")) ;
sam0(rNat,inv,"bop") = -(sum(rp, sam0(rNat,rp,"bop")) - sam0(rNat,"bop",rp)) + sam0(rNat,"trdmg","bop")) ;
```

There is a further section not detailed in here which handles the case of a non-diagonal make being introduced with the split.

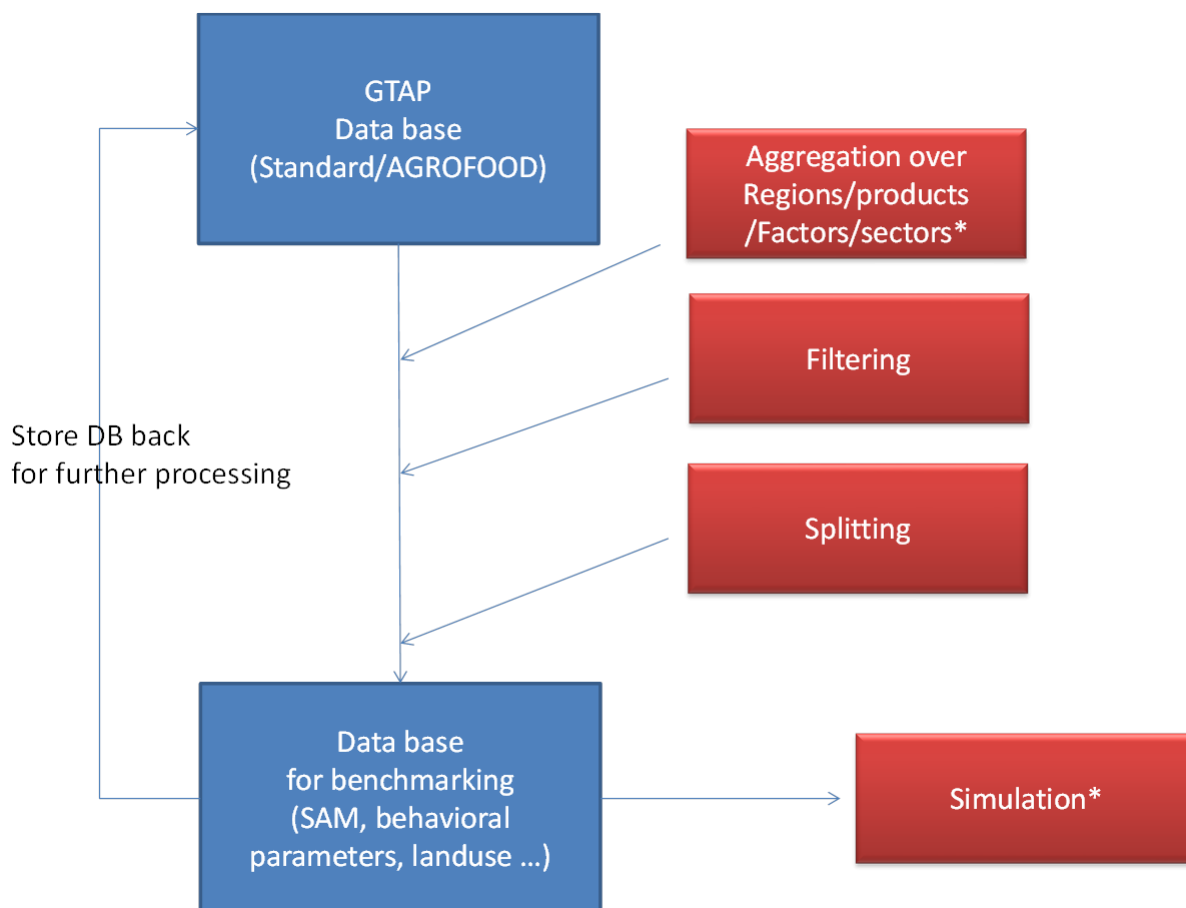
After the SAM is build, it is checked for balancing error and an error is thrown if larger relative errors are found. Afterwards, a generalized RAS routine is run to remove any remaining tiny balancing errors.

#### 4.3.20 MRIO split

The introduction of MRIO split values and the possibility to dis-aggregate commodities in the global SAM is discussed in some detail in section at the end of that documentation.

#### 4.3.21 Storing results from GTAPAgg as a new Data base for further processing

While the split utility has been successfully tested on extremely large split problem, such as introduced 50 additional sectors and products in the GTAP V10 Data base with full regional and sector detail, there are cases where it is useful to perform the split in steps. One might need, for instance, cost and expenditure shares of split-up products or activities to derive split factors for additional steps. The data base driver of CGEBox therefore opens the possibility to store the outcome of the split process back as a data base in the same format as the data downloaded from the GTAP center. A typical processing flow is depicted below.



Graph: Data base generation for further processing

This option is activated by the selection box “Generate new DB” as shown below. The user has to fill in this case the name of the sub-folder under the data directory where the data base should be stored:

The screenshot shows the 'Input data' dialog box with the following settings:

- Data base version: GTAPV10\_2014
- Data base variant (GTAP,POWER): gtap\_d
- Aggregation file: full
- Postfix for data set name: (empty)
- Generate MRIO split factors: ☐
- Apply SAM split: ☒ Split information file: fabio\_agr
- Split model: m\_split
- Regionalisation non-NUTS2: ☐
- Use manual re-aggregation: ☐
- Output all symbols to GDX: ☐
- Generate new DB: ☒
- Name for folder for new DB: FABIO



The outcome of the process is a folder which can comprise the following data files, where the presence of AEZ, air pollution and non-Co2 data base depends on their availability in the input data set.

```
aggStore
IDE aez_d.gdx
IDE airpolut.gdx
IDE cgebox.gdx
IDE gsdfnco2.gdx
IDE gtap_d.gdx
IDE gtap_m.gdx
IDE gtap_p.gdx
IDE gtap_s.gdx
```

#### 4.4 Model size and solution behavior

The model template including the GTAP-AEZ and GTAP-AGR modules was intensively tested under a suite of test shocks, applied simultaneously to all regions and sectors: 50 % reductions in tax rates (direct, consumption, factor, tariffs/export subsidies), 20% endowment changes and 10 % tfp shocks for all primary factors. The model was tested in three configurations: (1) the GTAP standard model with fixed allocation of foreign savings and (2) alternatively the global bank mechanism as a default in the standard GTAP model, and (3) the global bank mechanism in combination with GTAP-AGR and GTAP-AEZ. The tests used a full sector dis-aggregation (57 sectors) and varyingly sized regional aggregations (10,24,36,45,56,68). Additionally, a 10x10 model was solved for which the GTAP-AGR and GTAP-AEZ modules cannot be used due to missing sectoral detail. The parameterization was kept at defaults. Most of the tested models are large with regard to the number of sectors and countries compared to applications of GTAP reported in publications. All tests were also run solving the model as MCP in PATH and as a CNS in CONOPT.

GAMS version 25.5.4 was used, on a computer server, in combination with a beta-release of CONOPT4 which executes certain part of the NLP algorithm in parallel. We would like to acknowledge the continued support of Arne Drud, especially to let use CONOPT4 for the tests. Results with CONOPT3 are somewhat slower. The pre-solve algorithm with five pre-solve steps was applied. The aim of these tests was not only to ensure a stable numerical implementation of the model, but also to gain experiences with solution behavior on larger dis-aggregations.<sup>116</sup> The times reported in the table below are for a full model run without post-model processing switch on, but including the time needed to store all symbols in a GDX container, i.e. they cover loading the data base, model

---

<sup>116</sup> The aggregation definition files (\*.agg) used are found in the repository in the data directory. The test shocks can be found in the “scen” directory.



calibration, a benchmark solve and solving the shock.<sup>117</sup> The test show the expected more than linear increases in solution time if model size and complexity expands. Solution times with a fast multi-core laptop will be about double the times reported below.

**Table 2: Data bases used in the standard tests**

| Model size | Filtering thresholds (relative minimum # of transactions) | Resulting non-zeros in global SAM, including trade flows | Model size (GTAP Standard, maximal number of variables substituted out) |
|------------|-----------------------------------------------------------|----------------------------------------------------------|-------------------------------------------------------------------------|
| 10x10      | None                                                      | ~4.700                                                   | ~8.000                                                                  |
| 57x10      | None                                                      | ~35.000                                                  | ~92.000                                                                 |
| 57x24      | None                                                      | ~157.000                                                 | ~225.000                                                                |
| 57x36      | 1%, 160.000                                               | ~120.000                                                 | ~96.000                                                                 |
| 57x45      | 1%, 160.000                                               | ~130.000                                                 | ~99.000                                                                 |
| 57x56      | 1%, 160.000                                               | ~133.000                                                 | ~98.000                                                                 |
| 57x68      | 1%, 160.000                                               | ~135.000                                                 | ~99.000                                                                 |

Note: The number of transactions accounted for during filtering is somewhat higher than the resulting non-zero SAM entries

Many of the test solves will actually run even somewhat faster without the pre-solves switched on. However, if the solver fails to solve the shock without any pre-solves quickly, it can often spent several minutes until all infeasibilities are removed. As the potential further gains in solution speed by switching off the pre-solves are limited, we opted to show in here results obtained with the default settings. For small models and special applications such as sensitivity analysis, it might however pay off to check if the overhead of using the pre-solves can be avoided.

---

<sup>117</sup> The distribution which includes the GUI also comprises the described above test suite for the model, realized as an input file for the batch utility (gui\testbatch.txt). It is recommended to run the test suite after changes to the model code.

**Table 2: Solution time with different data bases and model configurations, CNS**

| Data base | GTAP standard, Fixed allocation of global savings | GTAP Standard, Global bank mechanism | GTAP Standard + GTAP-AGR+GTAP-AEZ<br>Global bank mechanism |
|-----------|---------------------------------------------------|--------------------------------------|------------------------------------------------------------|
| 10x10     | 10 sec                                            | 10 sec                               | (not possible)                                             |
| 57x10     | 30 sec                                            | 30 sec                               | 35 – 45 sec                                                |
| 57x24     | 1. min 30 sec                                     | 1 min 30 sec – 2 min                 | 1 min 40 sec – 2 min 30 sec                                |
| 57x36     | 1min                                              | 1 min – 1 min 30 sec                 | 1 min – 1min 30 sec                                        |
| 57x45     | 1 min 10 sec                                      | 1 min 10 sec – 2 min                 | 1 min 15 sec – 3 min                                       |
| 57x56     | 1.min 20 sec                                      | 1 min 20 sec – 3 min 20 sec          | 1 min 20 sec – 3 min 20 sec                                |
| 57x68     | 1 min 30 sec                                      | 1 min 30 sec – 3 min                 | 1 min 30 sec – 11 min <sup>118</sup>                       |

Even if additional tests showed that rather large models such as a 57x82 variant can be solved even without any filtering at least on some larger shocks in around ten minutes, some of the shocks tested failed with a model using the full resolution of the data base (57x134), but also with certain larger shocks on the 57x82 variant in any reasonable time. The tests were only possible with the beta version of CONOPT4 which is distributed with newer GAMS versions, as CONOPT3 will exceeds its internal memory limit of 8GB on very large models (around 600.000 equations which implies a SAM with about the same number of entries). That implies that users might run into very long solution times under a combination of larger shocks and very detailed data bases or might even not be able to load the model into the solver.

Analysis of changes in simulated welfare changes under a multi-lateral trade liberalization scenario which dismantled 50% of all import tariffs and export subsidies indicate that filtering can influence substantially results obtained. Unfortunately, the same holds for changing the regional and sectoral resolution of the model. We can therefore generally only recommend running highly dis-aggregated models with limited filtering. For further detail on solution times and aggregation bias due to sectoral and regional aggregation or the use of filtering, refer to Britz et al. 2015 who provide a larger sensitivity analysis with the standard GTAP model in GAMS.

The tests have shown that in combination with the pre-solves, solving the model as a CNS with CONOPT is usually faster if the model is larger compared of using PATH and solving as a MCP. Very small models are typically faster solved with PATH. Having both solvers available for tests might

---

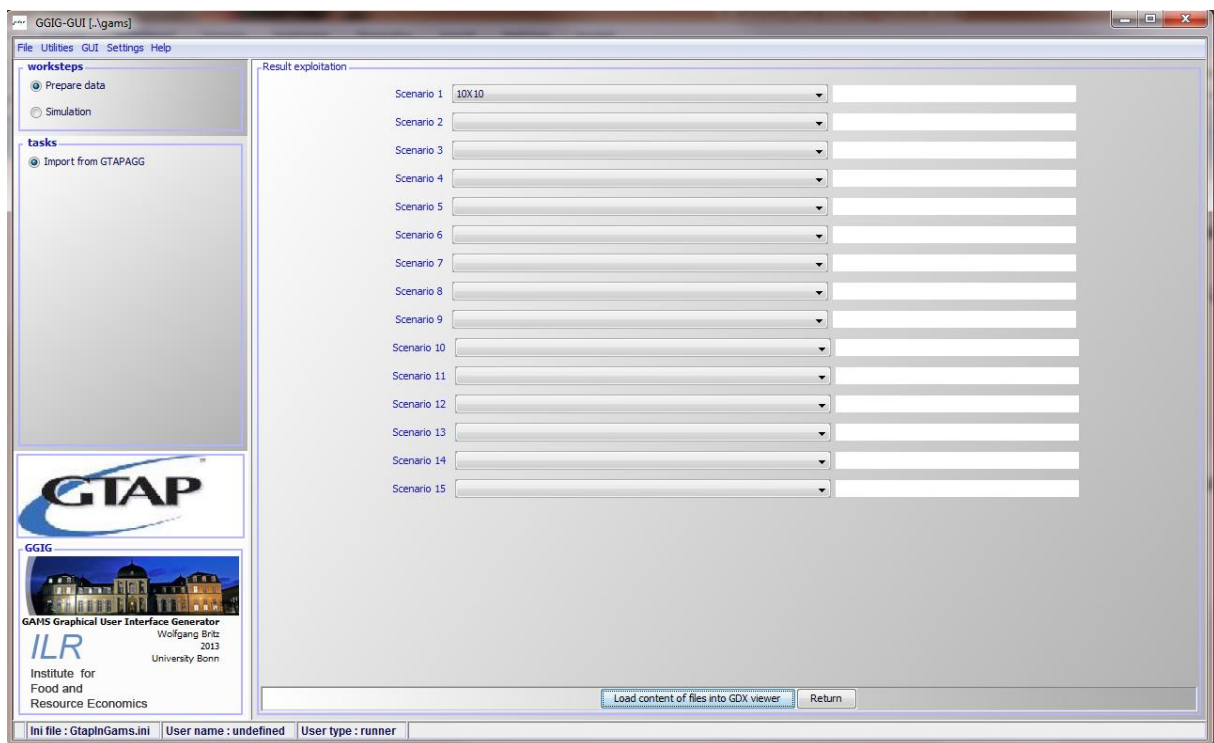
<sup>118</sup> The long solution times occurred under the direct tax cuts. As the standard GTAP model assumes that FDI is taxed at the same rate as domestic one, the direct tax cut impacts fully the expected returns to capital by foreign investors which drive the distribution of foreign savings under the global bank mechanism. A 50% reduction as simulated in the test shock can provoke very large differences in expected returns depending on the regional resolution of the data base. Thus, a new equilibrium can require massive changes in investments to decrease or increase the regional capital stock in order to drive marginal returns to capital up and down towards the global average. For some regions, that scenario might not feasible without fixing some variables at their bounds. Solving the model as a MCP should yield automatically that solution. However, solving as a CNS will yield infeasibilities in that case. Therefore, an algorithm is comprised in the GAMS code which tries to find, based on equations becoming infeasibility, which variables to fixed, which however requires to solve the model potentially several times. That algorithm can also fix similar problems which might not be related to the global bank mechanism.

hence pay off if solving large models is part of a project. And clearly, certain policy instruments such as production quotas are best captured by a KTT-condition embedded in a MCP model. Note that the GUI allows letting the model first be solved with CONOPT (as a CNS), and if that fails, to try a test with PATH (as MCP):

Checks for medium sized models have shown no differences in the results produced by CONOPT or PATH. That seems to indicate that once a model is declared feasible, the results can be used. Care should be clearly given when using PATH as a MCP solver after structural changes to equations: MCP solvers might e.g. accept a fixed variable paired to an equation which results in a non-square system, whereas solution as a CNS will throw an error in that case.

#### 4.5 Inspecting the resulting data base

You can check if loading and filtering worked by pressing on “Exploit results”, selected as shown as output file generated, a GDX comprising different parameter and set definitions:



Once you press “Load content of files into GDX viewer”, you can inspect the individual symbols used in the model:

GDVViewer

File Views Selection

☐ Use inlets in equation output  
☐ Use small font for non-selected var terms

Variable selection

Equation selection

Symbol from GDV

Dim2 (hhdd.gov,...)

GrainsCrops-c Meat.stk-c Extraction-c ProcFood-c TextWapp-c LightMnt-c HeavyMnt-c Util\_Cons-c

| user input | name      | type | dims  | records | long text                                   |
|------------|-----------|------|-------|---------|---------------------------------------------|
| is         | set       | 1    | 48    |         |                                             |
| aa         | set       | 1    | 14    |         | Armstrong agents                            |
| a          | set       | 1    | 10    |         |                                             |
| i          | set       | 1    | 10    |         |                                             |
| r          | set       | 1    | 10    |         |                                             |
| f          | set       | 1    | 5     |         |                                             |
| t          | set       | 1    | 2     |         |                                             |
| cbd        | set       | 1    | 1     |         |                                             |
| ind        | set       | 1    | 1     |         |                                             |
| fr         | set       | 1    | 1     |         |                                             |
| fd         | set       | 1    | 4     |         | Final demand                                |
| h          | set       | 1    | 1     |         | Households                                  |
| gov        | set       | 1    | 1     |         | Government                                  |
| inv        | set       | 1    | 1     |         | Investment                                  |
| spn0       | parameter | 3    | 4632  |         |                                             |
| nda0       | parameter | 3    | 1292  |         | Domestic purchases at agents' price         |
| ima0       | parameter | 3    | 1272  |         | Import purchases at agents' price           |
| imn0       | parameter | 3    | 1272  |         | Import purchases at market price            |
| imarg0     | parameter | 4    | 68377 |         | flows for mode m across all transport nodes |
| eh0        | parameter | 2    | 100   |         | CDE expansion parameter                     |

Double clicking on any on the symbols in the table on the RHS will open a new window as shown above.

The effect of using filtering and the rebalancing can be checked with the symbol “itrlog” which shows totals ,aggregated over commodities or sectors, for each rebalancing step (per country and “trace” at global level), the number of nonzero at global level and related changes. The example below shows the output from filtering a 32x34 case. Filtering removes about 76% of the original almost 135.000 non-zero items (see row “Total”). The line “curRelTol” shows the applied threshold in the iteration, which is increased stepwise to the desired maximum of 0.25%. Inspecting the relation between the dropped items and the cutoff used in that preliminary iteration can help to find a good compromise between model size and a too aggressive filtering where also more important transaction are wiped out.

[illegible]

## 4.6 Pre- or post-model aggregation

The standard case in GTAP applications is a project specific pre-model aggregation. Given the total size of the data base – there are already about 1 Million non-zero trade flows reported – some pre-model aggregation can be hardly avoided, it yields smaller models which are faster to solve, easier to debug and typically also numerically more stable with less small values being present. However, aggregation comes clearly at a price: not only might detail of interest for the application be lost, but

also peaks of policy instruments etc. wiped out. That can have important effects, Ko and Britz 2013<sup>119</sup> show in example applications that one might increase considerably simulated welfare from FTAs by simply increasing sectoral and regional detail.

CGEBox tries to soften the decision on what regional and sectoral aggregation level to use based on three features. Firstly, filtering and rebalancing of a data base can help to solve models with rather high number of sectors and regions (but clearly, the filtering process will also introduce aggregation bias). Secondly, the equations of the model had been carefully arranged; additionally scales and bounds for variables are introduced to stabilize the solution behavior of the solvers and to speed up solution time. And thirdly, post-model processing allows for a second aggregation step for reporting purposes. To do so, the user has to provide a second aggregation definition file (\*.agg, can be produced by GTAPagg). See also Britz and Van der Mensbrugghe 2016<sup>120</sup> for discussion of solution times and aggregation bias.

---

<sup>119</sup> Ko, J.-H. and Britz, W. (2013) Does Regional and Sectoral Aggregation Matter? Sensitivity Analysis in the Context of an EU-South Korea FTA", <https://www.gtap.agecon.purdue.edu/resources/download/6313.pdf>

<sup>120</sup> Britz, W., van der Mensbrugghe, D. (2016): Reducing unwanted consequences of aggregation in large-scale economic models - A systematic empirical evaluation with the GTAP model, *Economic Modelling* 59: 462-473

## 5. Scenario runs

### 5.1 Defining the scenario

#### 5.1.1 Implementation of shock files in GAMS

Running shocks with a CGE (or more generally an equilibrium model) is strongly related to the concept of a partition which defines what is endogenous and what exogenous, also called a closure. Different closures reflect that such a model comprises many variables or parameters, much more than equations. All these variables and parameters are possible candidates to become endogenous which implies that they will adjust to the shock while the exogenous ones will stay unchanged. Solving the model requires fixing exactly so many variables or parameters that the remaining number of endogenous variables is exactly equal to the number of equations. Fixing more than the number of equations will typically lead to infeasibilities, fixing less generally implies that the solution is not unique. Besides offering some defaults closures on the interface, the user can also provide its own closure files.

In both cases, shocks are only possible for non-endogenous elements of the model which can be technically either GAMS parameters or fixed variables where the lower is equal to upper bound. A symbol defined in GAMS as a parameter can however never become an endogenous variable. A shock file redefines the level of parameters in the model or fixes exogenous variables to a different value, away from the benchmark level. Consequently, the benchmark results of the endogenous variables do not longer provide a solution to the equations. The solver needs hence to search for a new combination of values for the endogenous variables which provide a fix-point for the new exogenous values.

The installation comprises a set of example shock files of which we discuss here two. We start with a shock which reduces all bi-lateral import tariffs and export subsidies, a partial multi-lateral trade liberalization (see *gams\scen\policy\_shocks\tarifs.gms*). Such a shock file is always structured in two parts: a section which declare symbols, called before the solution loop, and a section, executing during the simulation loop, where parameters or fixed variable levels are updated.

The declaration section starts with the pre-processor command “\$ifthen “%includeMode%”==”decl”, that statement is necessary and must be found in all shock files. In our example below, we declare two parameters, one for shocking the import tariffs (*p\_taxShockTariffs*) and one for shocking the export subsidies (*p\_taxShockExpSubs*). As these parameters are only used in the shock file, their names can be freely chosen by the user as long as they do not clash with other symbol names. In our simple example below, we shock all exporters (first dimension), importers (third) dimension and products (second dimension), reducing the taxes and subsidies by 50%. In the declaration part, we only define the size of the shock and what to shock, it is not yet applied to parameters or variables in the model. The declaration code alone would hence not force the model away from the benchmark.

```
$iftheni "%inclMode%"=="decl"
*
*
*   --- define which bi-lateral combinations and commodities to shock
*       that can be achieved in multiple ways, we define here a multiplier
*
parameter p_taxShockTariffs(r,i,rp)
           p_taxShockExpSubs(r,i,rp);
*
*   --- here, we could e.g. also insert a table, read from
*       EXCEL or from a GDX container
*
*       0.9 would mean reduce by 90%
*
p_taxShockTariffs(r,i,rp) = 0.50;
p_taxShockExpSubs(r,i,rp) = 0.50;

$else
```

The second part after the \$else statement applies the defined levels of the shock to re-define new fixed levels of the two variables *imptx* and *exptx* on the LHS. This will lower as a first order effect the import prices faced by demanders and the tax income of the governments with subsequent second and third round effects towards a new equilibrium.

The RHS of the assignments shown below reduces the benchmark level (“%t0%” is either the year “t00” in a recursive-dynamic run or the benchmark point “bench” in a comparative static one) of the exogenous variables according to the desired shock, implementing it in equal steps over the simulation horizon. How that stepwise application of the shock is programmed in GAMS is discussed next. The set *tSim* in a comparative static simulation only comprises two elements, “bench” and “sim”. The shock file is only called for “sim” to leave the benchmark point unchanged; the position of the element “sim” in the set *tSim* is second. Hence, *tSim.pos*-1 yields unity. The divisor is the number of elements in the set, i.e. *card(tSim)*, minus one, which is also equal to unity. Accordingly, the code will for a comparative-static run multiply the shock as declared in the parameters above with one, in other words, the shock is fully applied in the one and only simulation run with the model.

Now, we look at the case with a recursive-dynamic run over 10 years, the elements in the set *tSim* would be called t00, t01, .... t10. If the driver calls the code for the element *t01*, which is the second element in the set, the numerator is  $(2-1) = 1$  as in the example above while the denominator is  $(11-1)=10$ , such that 10% of the shock will be applied. In each year, the numerator will increase by one, such that in the next year 20% of the shock will be present and so forth.

```

*
* r = importer regions, i=products, rp= exporters, tsim= simulation time point
* imptx: Bilateral import taxes

imptx.fx(r,i,rp,tsim) $ (imptx.l(r,i,rp,tsim) $ p_taxShockTariffs(r,i,rp))
= imptx.l(r,i,rp,"%T0%")
*
* --- this distributes the reduction in equal steps over the
* simulation horizon (or in step-wise shocks to ease solving
* if mimic comp-static is switched on)
*
* * (1-p_taxShockTariffs(r,i,rp) * (tsim.pos-1)/(card(tsim)-1));
*
* --- note that we only cut export subsidies, i.e. the cases where exptx < 0
*
exptx.fx(r,i,rp,tsim) $ ((exptx.l(r,i,rp,tsim) lt 0 ) $ p_taxShockExpSubs(r,i,rp))
= exptx.l(r,i,rp,"%t0%") * (1-p_taxShockExpSubs(r,i,rp) * (tsim.pos-1)/(card(tsim)-1));

```

The mechanism with the step-wise changes can also be used in a pseudo recursive-dynamic run without capital accumulation or other features relating to dynamics. In that case, the sole purpose is to apply not the full shock in one go to the model. That might ease solving strong shocks on complex models. Similar files are available which define shocks consumption, factor and direct taxes.

Another type of often applied shocks changes productivity levels. The related variables are mostly called lambda + postfix, where the postfix is the name of the variable to shock, as shown in the table below:

| Shifter variable  | Meaning                                          |
|-------------------|--------------------------------------------------|
| axp(r,a,t)        | Production function frontier                     |
| Lambdand(r,a,t)   | Demand for intermediate demand composite         |
| lambdava(r,a,t)   | Demand for value added composite                 |
| lamdaf(r,f,a,t)   | Factor demand, separate for each factor          |
| lambdaio(r,i,a,t) | Intermediate demand, separate for each commodity |

The code below shows the content of the example file “*gams\scen\end\_and\_tp\_shocks\factor\_specific\_tp.gms*”. It increases the factor productivity for all regions, activities and factors by 10%.



```
$iftheni "%inclMode%"=="decl"
*
    parameter p_tpf(r,f) "Shifter for factor demand";
*
*   --- here, we could e.g. also insert a table, read from
*       EXCEL or from a GDX container
*
    p_tpf(r,f) = 1.1;
*
$else
*
*   r= region, f=factors, a=activities, t=time
*   If Lambdaf is not equal across factors, implies Hicks non-neutral TP
*
    lambdaf.fx(r,f,a,tSim) $ p_tpf(r,f)
*
*   --- this distributes the reduction in equal steps over the
*       simulation horizon (or in step-wise shocks to ease solving
*       if mimic comp-static is switched on)
*
    = ( 1 + (p_tpf(r,f)-1) * (tsim.pos-1)/(card(tsim)-1)) ;
$endif
```

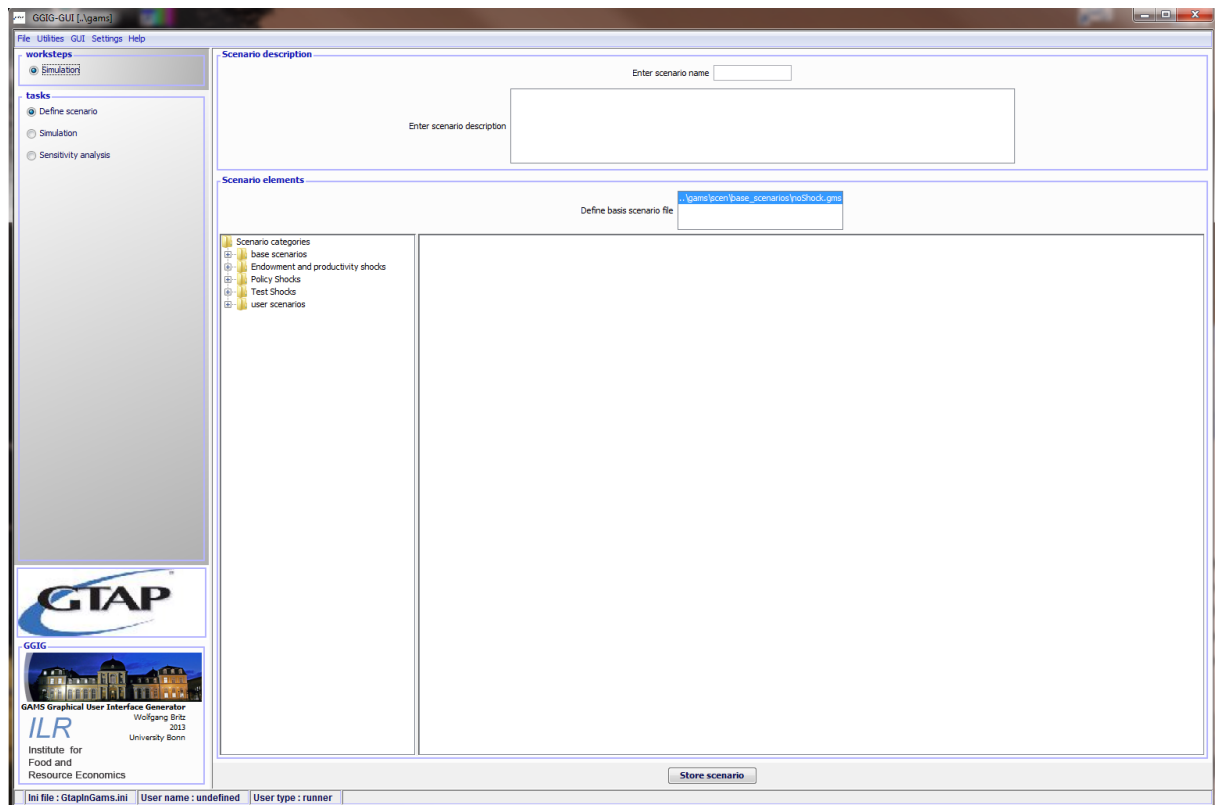
Note again the dollar operation “\$ p\_tpf(r,f) on the LHS. That ensures that only non-zero changes are used in the shock. If, for instance, values for *p\_tpf* are only found for one region, the productivity shifters for all other regions will stay at their benchmark point. The usual way for the experienced coder is to define shock files in text editor such as GAMIDE, alternatively, the simpler build in scenario editor can be used.

The most important step after a shock was run is to make sure that the desired changes show up in the model. One way to do so is to let the GAMS code store all symbols into a GDX file (see also next section). Opening a symbol such as *lambdaf* in the example above should then show the desired change from the benchmark to simulated point. Productivity shifters and tax rates can also be inspected by pre-define views in the exploitation tools.

### 5.1.2 Scenario editor

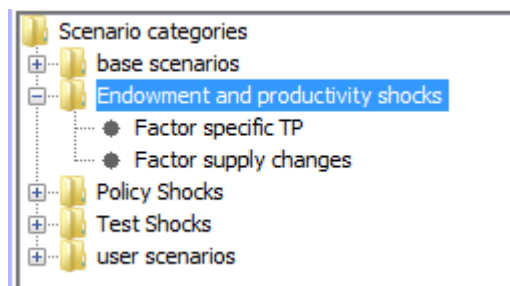
The interface helps with the definition of scenarios – it has a set of example files with shocks which can be combined and edited directly in the GUI. Alternatively, you can use any text editor and define a shock in GAMS and store the file somewhere under “\gams\scen”. That is the recommend way to work in order to define complex scenarios and for users familiar with GAMS coding. In order to use the in-built scenario editor instead, click on “Define scenario” and the interface should look like below.

## Defining the scenario

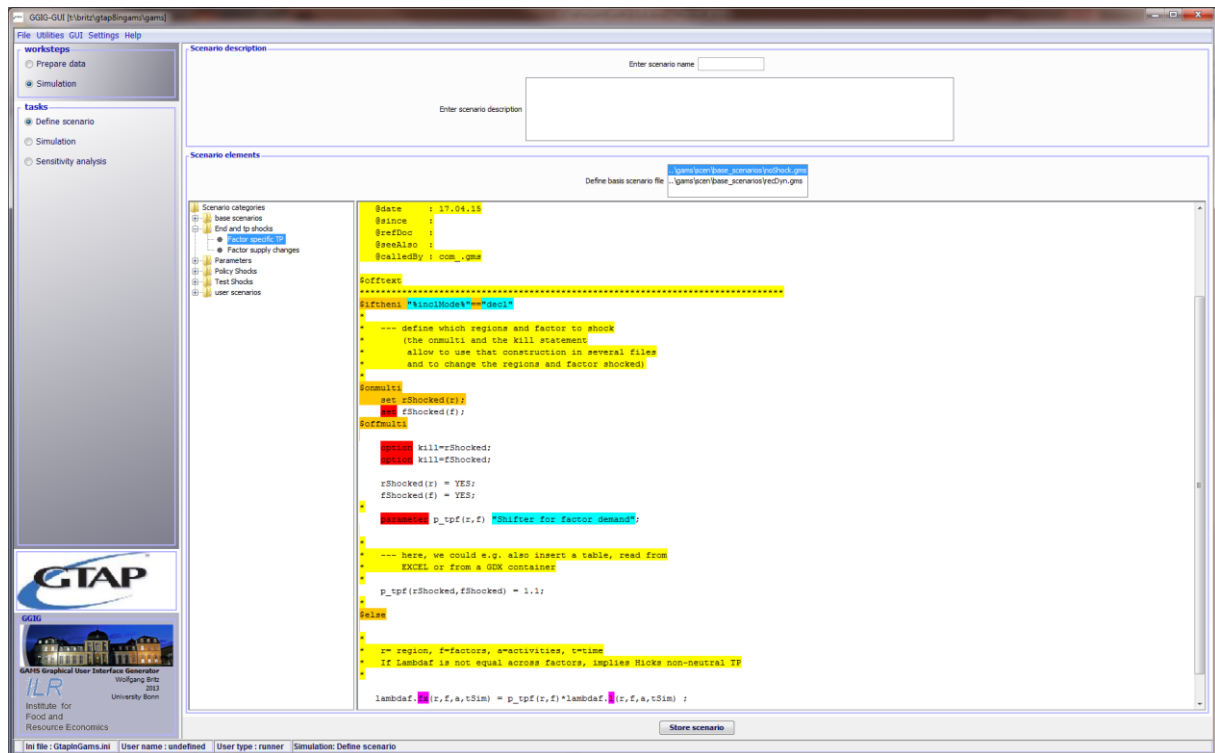


Currently, there are two base scenarios:

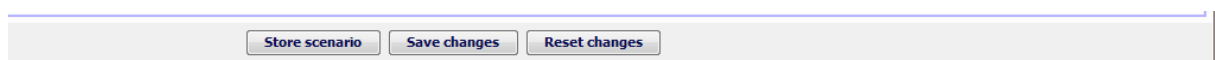
- **No shock** – that is simply an empty file
- **RecDyn** – a test implementation for a driver for a recursive-dynamic baseline.
- In order to add the content of pre-defined shock files to your base scenarios, click on a tree element, e.g. “Endowment and productivity shocks”. You should be able to see a list of files as below:



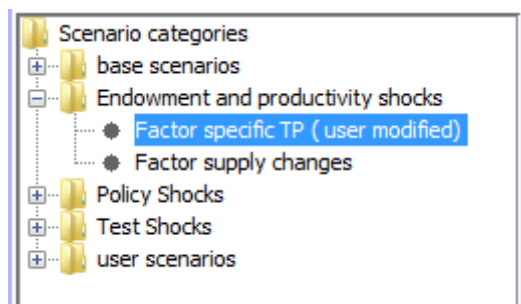
Double-click for example on “Factor specific TP”, and the content of that file is loaded in the editor on the right as seen below:



You can now modify the content with the editor by typing directly in the right hand window. Once you have edited the code, additional buttons will appear:

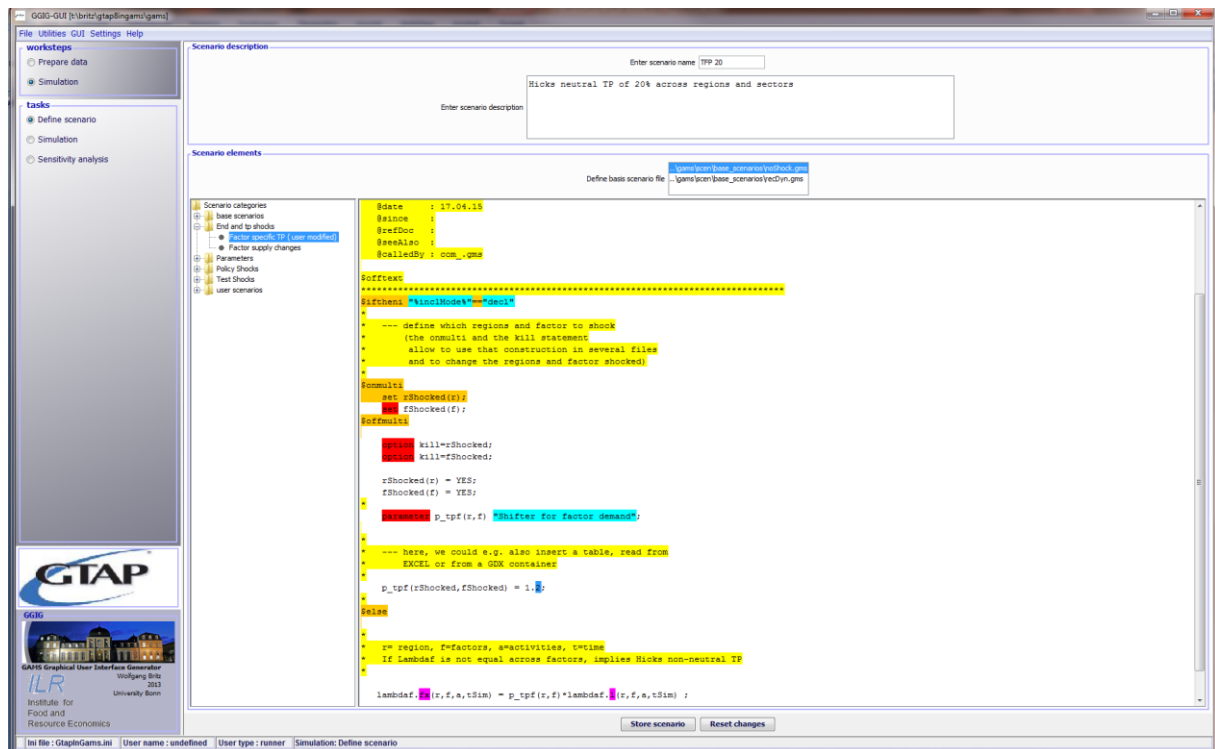


The list with the scenario groups and shock files will now indicate that you have made changes to the shock file.

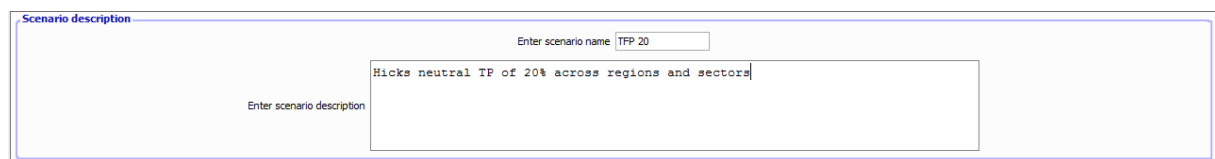


In order to save your edits, press “Save changes”. If you do not press “Save changes”, your edits won’t make it in final shock file generated. You can now double click on further files to add them to your shock file. The files you have currently combined are shown in blue; you can also deselect some of these files again. Files which were edited by you (only a temporarily copy is changed) have a “(user modified)” appended to their name. Edits in such files are shown in blue (in the example below, the factor 1.1 was changed to 1.2).

## Defining the scenario

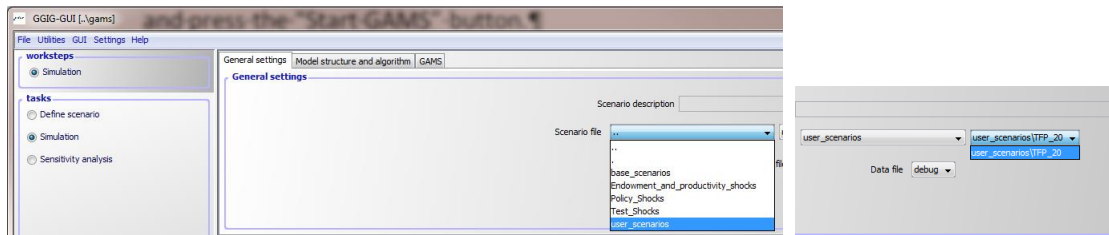


Once you have chosen the files you want to combine and potentially edited (and saved!) your changes, you should now enter name for the shock file to create, e.g. TFP20 and a description of the scenario as shown below:



Afterwards, press “Store scenario”:



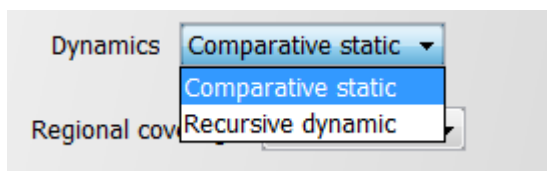


and press the “Start GAMS” button.

That should run your counterfactual against the reference and generate results in a GDX file.

The tab “General settings” gives you the following additional input possibilities:

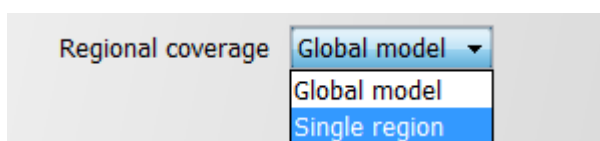
- **Entering a scenario description:** Normally, the output files generated are named after the scenario file used. This name can be overwritten with a non-empty content of this text field.
- **Post fix for result files:** The post fix is appended to the names of the output files. It can e.g. used to differentiate versions using the same shock, e.g. results generated based on different closures or by using different model configuration.
- Choice between the Comparative static or Recursive dynamic version of the model:



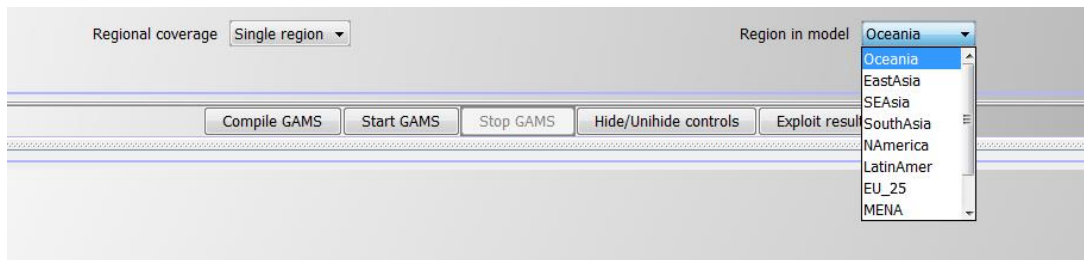
If the “Recursive dynamic” option is switched on, the following options are shown:



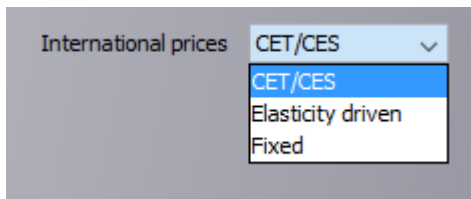
1. The number of years defines the simulation horizon. The actual years simulated depend on the benchmark year of the data set chosen (e.g. 2014 as the default with GTAP Version 10).
  2. The time resolution defines which years of the horizon are simulated (time steps). The first and last year are always included.
  3. The report frequency defines for which of the simulation years the results are saved. It must be multiple of the time resolution or equal to the time resolution. If a zero is used, the user can manually edit the years for which results should be generated.
  4. “Mimic comp-static” allows switching off capital accumulation such that a comparative-static model is solved in several steps. That is useful for step-wise introduction of shocks, also called “shock slicing”.
- Choice between the Global Model or Single Region version of the model:



If the single region mode is switched on, a drop-down box appears where the region can be chosen:



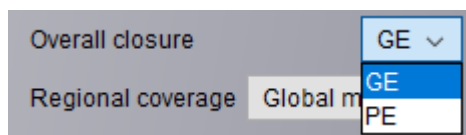
Furthermore, the way international prices are related to that, how import and export flows are treated can be chosen in case of a single region model:



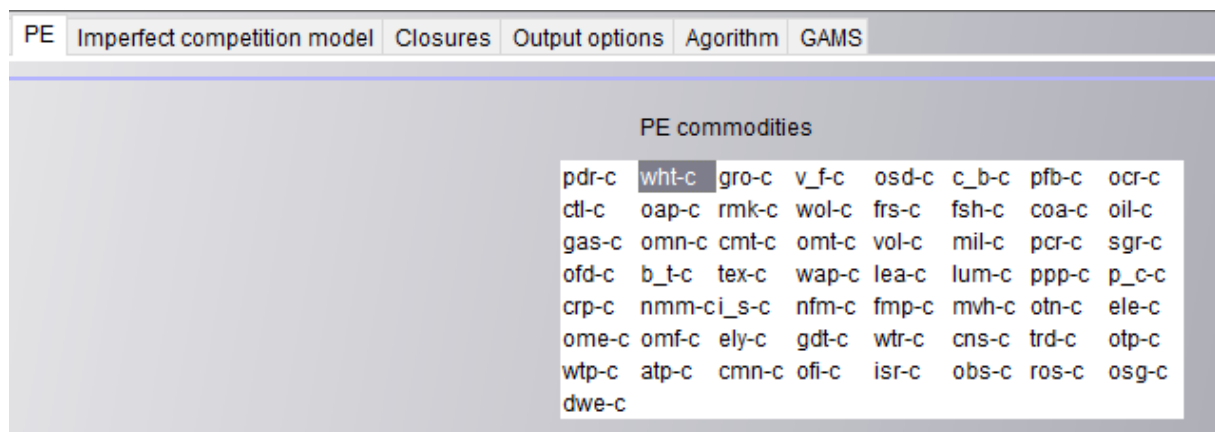
- CET/CES: Imports to other regions react to changes in the region's cif prices according to the share equations of the lower Armington nest at fixed total import quantities and aggregate import prices. If CET elasticities for exporters to the region are not infinite, the lower CET nest at the export destination are active and fob prices react accordingly. Otherwise, they are fixed.
- Elasticity driven: in that case, aggregate import and export prices of the region are driven by iso-elastic functions of the region's total import and export, respectively.
- Fixed: import and export prices of the region are fixed. That will typically only make sense if a CET between domestic and export supply is active.

The reader should note that the alternative to a single region model is a layout with a two-region aggregation of a GTAP Data Base: the single region against all other regions aggregated into one Rest-of-the-World model region. This layout will automatically reflect cases where the small country assumption makes limited sense, for instance, cases where the region singled out has a large export or import share in selected market.

- Choice between the General Equilibrium or Partial Equilibrium closure of the model:



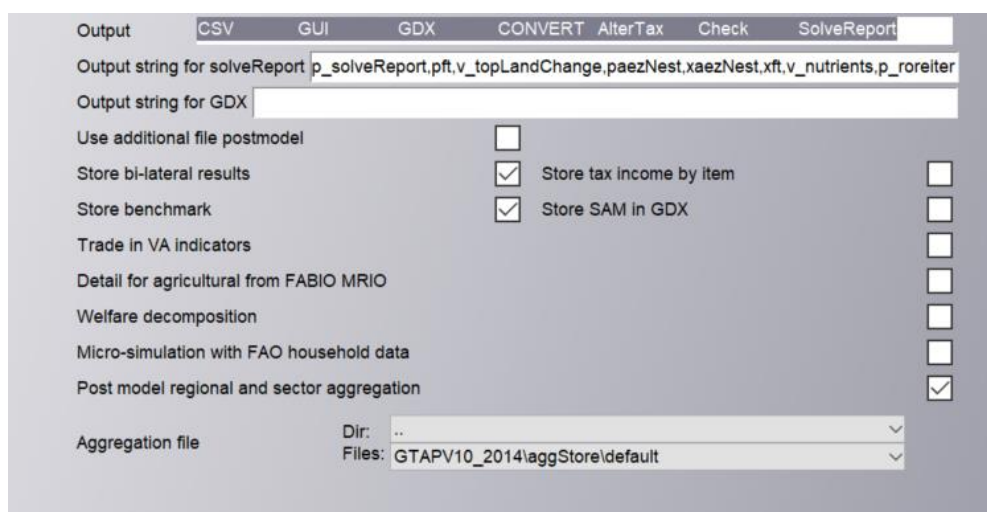
The GE closure is the usual set-up for a General Equilibrium model. The PE closure allows solving for only one; several or all commodity markets while clearing factor markets at fixed income. The product markets which clear can be chosen on the tab "PE":



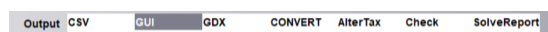
Please note that the usual welfare interpretation of a CGE is no longer valid if income is exogenous and/or only some markets clear.

### 5.3 Output options

The output options give some flexibility how to inspect the final model solution, which typically will relate to the levels (and derived from there, changes) in endogenous variables. Related to this, it is useful to also check changes in exogenous elements to make sure that the shock was correctly implemented. The output options panel looks like this if only GUI (the default) is selected:



Also, make sure that output for the GUI is selected:



The different output options are sketched as:

- CSV will generate a SAM and some other core results and store it in CSV format under “results/run”
- GUI will generate a GDX container with a parameter for use with the exploitation tools (see below)
- GDX will store all other GAMS symbols (parameters, variables, equations, sets etc.) in a GDX container, including the output parameter for the GUI if GUI selected. The user can overwrite this behavior by inputting a list of symbols to export in the text file “Output string for GDX”.



- CONVERT will generate output for the “Equation and Variable viewer” which is similar to AnalyseGE (see section below).
- AlterTax will store the generated SAM and some vectors back in a format which can be used as input for the model and thus allows constructing a new benchmark based on a model run.
- SolveReport will store during simulation a parameter *p\_solveReport* in a GDX file in the scratch directory, for instance, to track the behavior of the model in recursive dynamic run.

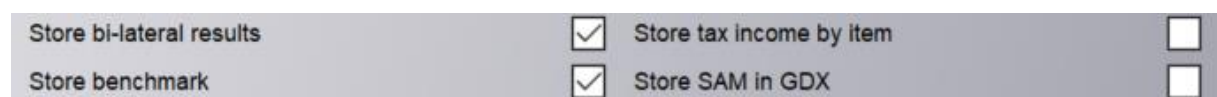
The screenshot shows a text input field with the text "p\_solveReport" inside. Above the field, the label "Output string for solveReport" is visible.

The user can add, separated by comma, further symbols to be stored in the GDX container.

The user can add its own file for post-model reporting, for instance, to calculate summary indicators by switching “*Use additional file postmodel*” on. In this case, a user provided file can be selected. Such files are stored in the directory “/gams/scen/post\_model”. This is only useful if also “GDX” output is switched on and the “Output string for GDX” is left empty or comprises the generated report parameter.

The screenshot shows a horizontal control bar. On the left, the text "Use additional file postmodel" is followed by a checked checkbox. To the right of the checkbox is the text "Postmodel file". Further right is a dropdown menu currently displaying the word "empty" with a small downward arrow.

Most other options are relevant for GUI output only:

The screenshot shows a horizontal control bar with four options, each with a checked checkbox. The options are: "Store bi-lateral results", "Store benchmark", "Store tax income by item", and "Store SAM in GDX".

The option “*Store-bi-lateral results*” will put bi-lateral trade flow information into the result set. For data sets with many regions and products, switching this off can dramatically reduce the size of the data set generated. Similarly, “*Store tax income by item*” will reduce the number of items stored for each transaction. Normally, for each transaction, the quantity index, the price index, the value (quantity times price), the tax rate and also the value of quantity times (price plus tax) will be stored.

The same holds for the post-model aggregation. The screenshot below shows what post model aggregation contributes: in most views, results are not only available in the regional and sector disaggregation of the underlying data base, but additionally also aggregated results as shown below.

| Demand, by product [0]      |                    |                    |                   |                     |                    |                   |                                |
|-----------------------------|--------------------|--------------------|-------------------|---------------------|--------------------|-------------------|--------------------------------|
|                             | Regions            | Items              |                   |                     | Origins            |                   |                                |
|                             | World              | Quantity           | Total             |                     |                    |                   |                                |
|                             | TARIFFS            |                    |                   |                     |                    |                   |                                |
|                             | Total              | Household demand   | Government demand | Intermediate demand | Investment demand  | Export demand     | International transport demand |
| Total                       | 137582,97<br>0,41% | 31212,10<br>0,30%  | 9534,91<br>0,02%  | 54631,16<br>0,01%   | 12416,49<br>-0,69% | 23091,78<br>2,03% | 591,84<br>4,30%                |
| Agriculture                 | 2058,32<br>0,43%   | 800,55<br>0,29%    | 1,87<br>12,82%    | 504,73<br>0,58%     | 0,57<br>-0,73%     | 413,50<br>0,78%   |                                |
| Mining and Extraction       | 5474,23<br>2,91%   | 58,66<br>-1,99%    |                   | 2484,97<br>0,57%    | 1,05<br>0,20%      | 1542,53<br>6,36%  |                                |
| Light Manufacturing         | 17319,36<br>1,02%  | 3950,65<br>1,51%   | 18,43<br>1,91%    | 4854,85<br>0,74%    | 985,41<br>-0,50%   | 5116,43<br>1,23%  |                                |
| Heavy Manufacturing         | 32869,18<br>0,68%  | 2062,65<br>0,66%   | 53,34<br>0,27%    | 15055,72<br>-0,24%  | 2527,85<br>-0,77%  | 11201,62<br>2,35% |                                |
| Utilities and Construction  | 11765,99<br>-0,44% | 509,02<br>0,28%    | 9,54<br>-0,57%    | 3485,60<br>-0,59%   | 7415,18<br>-0,64%  | 292,50<br>5,26%   |                                |
| Transport and Communication | 22034,38<br>0,20%  | 9171,68<br>0,21%   | 67,26<br>0,88%    | 10126,40<br>0,09%   | 593,31<br>-0,63%   | 1725,47<br>1,05%  | 591,84<br>4,30%                |
| Other Services              | 46061,50<br>0,01%  | 14658,88<br>-0,01% | 9384,46<br>0,01%  | 18118,90<br>0,01%   | 893,11<br>-1,15%   | 2799,73<br>0,40%  |                                |
| Total                       | 137582,97<br>0,41% | 31212,10<br>0,30%  | 9534,91<br>0,02%  | 54631,16<br>0,01%   | 12416,49<br>-0,69% | 23091,78<br>2,03% | 591,84<br>4,30%                |
| Agriculture                 | 2058,32<br>0,43%   | 800,55<br>0,29%    | 1,87<br>12,82%    | 504,73<br>0,58%     | 0,57<br>-0,73%     | 413,50<br>0,78%   |                                |
| Paddy rice                  | 21,36<br>-1,22%    | 6,28<br>-0,03%     | 0,19<br>-1,34%    | 11,79<br>-1,76%     |                    | 1,55<br>-1,52%    |                                |
| Wheat                       | 99,08<br>-0,12%    | 12,88<br>0,28%     |                   | 26,67<br>-0,05%     |                    | 29,76<br>-0,24%   |                                |
| Cereal grains nec           | 124,97<br>0,45%    | 48,10<br>0,76%     | 0,42<br>22,49%    | 17,66<br>-1,59%     |                    | 34,18<br>0,75%    |                                |
| Vegetables, fruit, nuts     | 910,70<br>0,70%    | 477,79<br>0,29%    |                   | 252,22<br>1,14%     |                    | 114,20<br>2,43%   |                                |
| Oil seeds                   | 133,49             | 8,28               |                   | 58,67               |                    | 33,27             |                                |

## 5.4 Model structure, parameterization and factor markets

You might also select from different model setups:

| General settings: Model structure Parameters and indices AEZ NUTS2 Tariff lines myGTAP MRIO PE G-REDM Imperfect competition model Closures Output options Algorithm settings GAMS |                     |                       |  |               |  |                             |                           |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|-----------------------|--|---------------|--|-----------------------------|---------------------------|
| Model structure                                                                                                                                                                   |                     |                       |  |               |  |                             |                           |
| Modules and parameters                                                                                                                                                            |                     |                       |  | Demand system |  | Factors                     |                           |
| myGTAP                                                                                                                                                                            | GTAP_AGR            | GTAP_AEZ              |  | CDE           |  | Mobile and sluggish factors | Factors with fixed prices |
| GTAP_E                                                                                                                                                                            | Labor_nest          | capSkLab_nest         |  | LES           |  |                             |                           |
| Water_nest                                                                                                                                                                        | CO2_Emissions       | Non_CO2_Emissions     |  | AIDADS        |  | Land                        | Capital                   |
| Aggregate_GovToFinalDemand                                                                                                                                                        | Aggregate_Armington | Aggregate_firm_demand |  | CD            |  | UnSkLab                     | UnSkLab                   |
| GTAP_Melitz                                                                                                                                                                       | CapVintages         | GTAP_NUTS2            |  |               |  | SkLab                       | SkLab                     |
| GTAP_MRIO                                                                                                                                                                         | TariffLines         | Parameters            |  |               |  | NatRes                      | NatRes                    |
| G-REDM                                                                                                                                                                            |                     |                       |  |               |  |                             |                           |

### 5.4.1 Modules and parameters

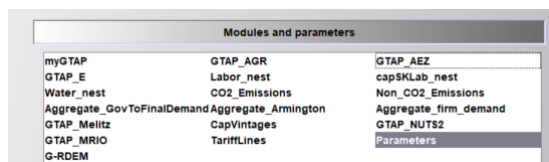
Currently, different modules and extensions as shown above can be added to the GTAP standard model. Please note that these modules are not a full replica of the GEMPACK implementations. The labor nest – which is also part of GTAP-E – depicts substitution between skilled and unskilled labor. The “CapSkLab nest” depicts substitution between capital and skilled labor.

The “Aggregate intermediate demand” extension reduces the number of Armington agents to four: final, government and investment demand and aggregate intermediate demand. The latter replaces the sector specific nests in the GTAP standard model. “Aggregate Armington” makes a step further by using the same shares for domestic and import demand for all agents.

GTAP\_NUTS2 (see section “Sub-national dis-aggregation of production and factor markets in CGEBox”, page 146ff) provides a dis-aggregation of the production function and factor markets to

sub-regions, assuming homogenous output products from the regions in a nation. Given current data availability, that module works for European countries, only.

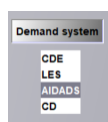
In order to introduce non-default parameters (such as a CET or non-zero substitution between ND and VA), switch “Parameters”



and selected the desired file on the “Parameter and Indices” tab, see section 5.6.1, page 391.

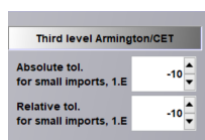
### 5.4.2 Demand system

The modular approach covers also the choice of the demand system for private households. Specifically, the standard CDE demand system found in GTAP can be replaced by a Linear Expenditure System or a Cobb-Douglas representation.



The LES system will determine the marginal budget shares based on income elasticities derived from the parameterization of the CDE demand system. The commitment terms are derived by assuming an 80% share of non-committed income, ensuring that commitments are non-negative and do not exceed 50% of observed demand quantities. The CD system uses the observed budget shares to determine the share parameters.

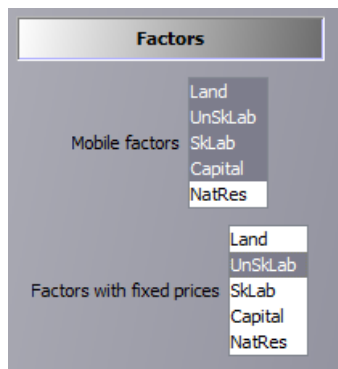
The tolerance for small imports steers the third level of the Armington system and additionally the CET, if switched on. The option can only be activated as developer and otherwise not visible.



That features found in GLOBE uses a Leontief relation between total imports respectively exports and small trade flows. That can overcome numerical stability issues with tiny trade flows and decreases overall model complexity. The user should note that it is not possible to use simultaneously a Leontief relation on the import and the export side for the same flow. Therefore, the code will not use the Leontief on the export side if already active on the import side.

## 5.5 Factors

The first selection possibility allows defining which factors are mobile – potentially combined with a CET approach to model sluggish factor mobility – and which are sector specific and thus immobile.

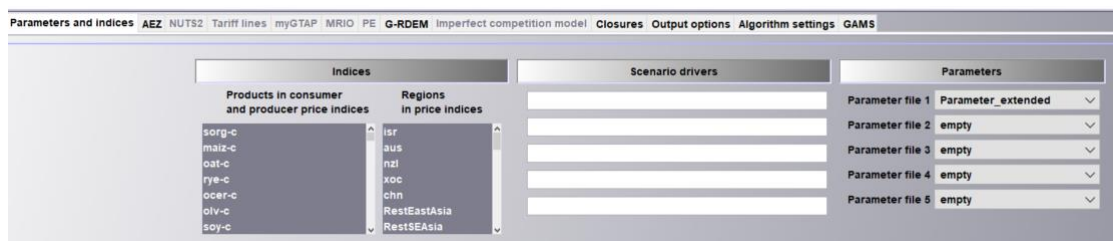


The second selection allows fixing factor prices - either economy wide or sector specific, depending on the choice above – such that their stock is endogenously adjusting. That is an often found solution to model unemployment based on a fixed wage rate. The reader should note that the model can also be parameterized with factor supply elasticities to model e.g. a wage curve approach or to introduce a land supply curve.

## 5.6 Parameters and indices

What is shown on the panel depends on whether the “Parameters” module is switched on and if the user is registered as a developer or not.

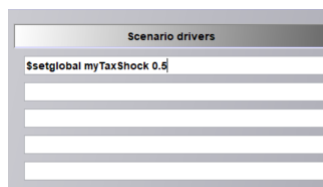
### 5.6.1 Indices



The indices are only visible for a developer; otherwise, all products and regions are used as a default. The selection of parameter files was discussed above.

### 5.6.2 Scenario drivers

The scenario drivers allow to write text directly in the include file. Consider the following example:



The input from the interface is directly inputted in the include file “com\_inc.gms”:

```
* ---- Scenario drivers
$setglobal myTaxShock 0.5
```

That allows to e.g. use that global variables in a shock file to change the size of the shock without the need to modify the GAMS code. That for instance the following extension of the example shock file “scen/policy/tariffs.gms”:

```

imptx.fx(r,i,rp,tsim) $ (imptx.l(r,i,rp,tsim) $ p_taxShockTariffs(r,i,rp))
    = imptx.l(r,i,rp,"%T0%")
*
* --- this distributes the reduction in equal steps over the
*      simulation horizon (or in step-wise shocks to ease solving
*      if mimic comp-static is switched on)
*
*      * (1-%myTaxShock%*p_taxShockTariffs(r,i,rp) * (tsim.pos-1)/(card(tsim)-1));

```

The global variable is here introduced to change the size of the shock directly from the interface. That option is helpful for batch files to run multiple variants of the same basic shock as shown below. Note that the names after the \$setglobal can be chosen by the user as preferred.

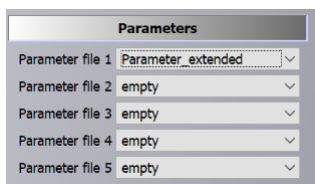
```

for trade = 0 \ 1
for migr  = 0 \ 1
for savf  = 0 \ 1
    ScenDriver1 = $setglobal scenDriver1 %trade%
    ScenDriver2 = $setglobal scenDriver2 %migr%
    ScenDriver3 = $setglobal scenDriver3 %savf%
    ScenDriver4 = $setglobal scenDriver4 %trade%
    batchName = brexit_t%trade%_m%migr%_s%savf%_n%trade%
    Scenario description = brexit_t%trade%_m%migr%_s%savf%_n%trade%
    execute=Gamsrun
endfor
endfor
endfor

```

### 5.6.1 Parameters

Non-Default parameter files can be used with the model if the module “Parameters” is switched on. In that case, the tab parameters and indices allows to selected up to 5 different files which comprises changes to the default parameterization:

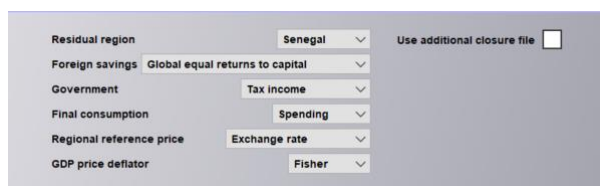


These files are stored in the directory “gams\scen\Parameters”. The user can hence change existing files or add new ones which then can be selected from the interface.

The possibility to use non-default parameters is important as certain model extensions are driven by parameter settings. To give an example: in the standard GTAP model, the transformation elasticity between domestic output and exports is infinite, such that producer prices between the two destinations are equal and physical balancing is used. The GAMS version allows alternatively using a CET structure where prices differs and consequently, a non-linear quantity aggregator is used. That feature can be switched on by providing a parameter file in GAMS format where a CET elasticity different from infinity is set.

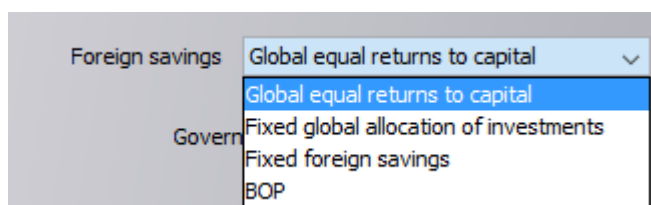
**Note:** Modules and extensions and the use of non-default parameters are disabled if pre-defined configurations are used.

## 5.7 Closures

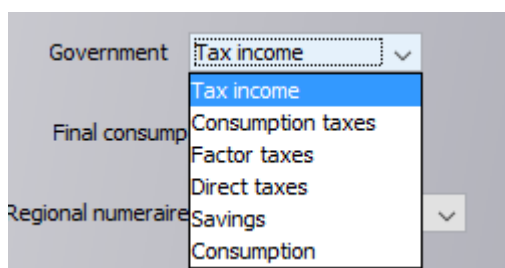


As shown above, the code supports alternative closures different from the ones used in the standard GTAP model:

- For **foreign savings**, besides the default global bank mechanism which leads to identical expected returns to capital across all regions, the model can run with a fixed global allocation of investments and with fixed foreign savings. These closures are only available in the global model set-up. The balance of trade (BOP) solves for the exchange rate, it requires that the factor price index is used as the regional numeraire.

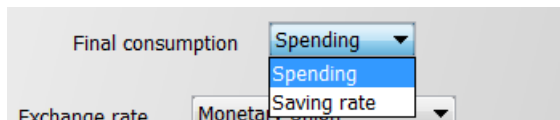


- The standard closure for the **government** account is to calculate tax income at given tax rates, and to let the top level utility function allocate a certain share of regional income to government consumption. The reader should note that there is no immediate relation between changes in tax income and government consumption. Changes in government consumption relative to private consumption and regional saving depend on the one hand on changes of the related price indices and on the other on the elasticity of the private consumption share to changes in overall utility. As such, there is no closed government account in the regional household approach, one could argue that implicitly government saving adjust. The alternative closures endogenize tax shifters either on all or selected products in consumption or for all or selected factors while keeping the real tax income for government consumption fixed, using the government price index to define real tax income. These two closures come closer to a representation of separate government account.

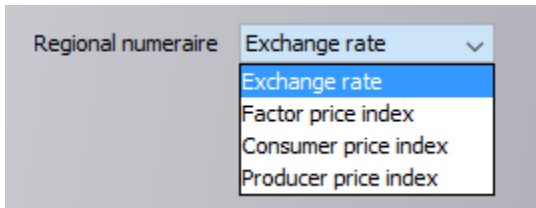


For the closures modifying taxes can be further detailed by applying the shifters. The closures where “savings” and “consumption” adjust are only available with the myGTAP extensions where a separate government account is present.

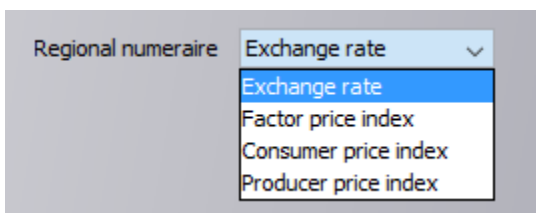
- The standard closure for **private consumptions and saving** is to have flexible budget shares for private and government consumption and regional savings. The alternative closures allow to fix private consumption spending and to let the saving rate adjust.



- The user can also use a regional reference price. That does not influence the results (quantities, welfare impacts), but simulated price levels will differ.



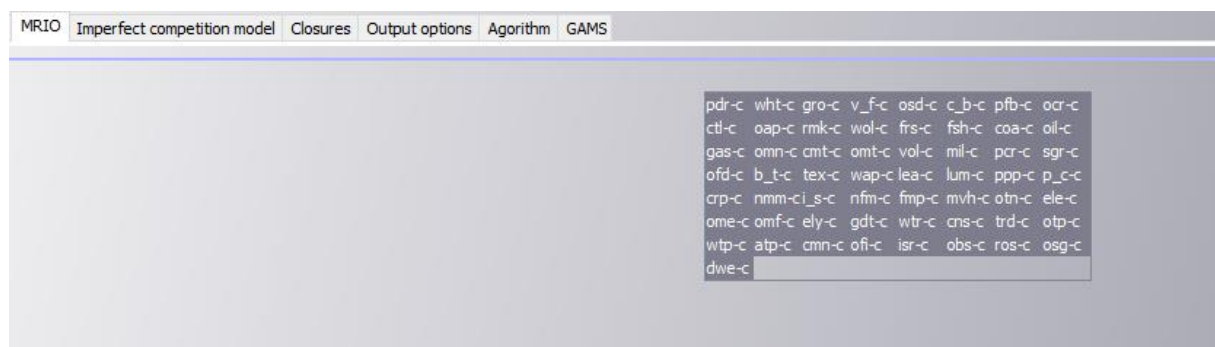
- For the single region layout, two closures for the current account balances are offered: either fixed foreign saving with a flexible exchange rate or the reverse combination.



**Note:** The pre-defined configurations also define specific closures.

The user can also provide its own closure file to introduce alternative and additional closures.

Currently, the MRIO extension lets the user solely depict the commodities for which bi-lateral import demand is dis-aggregated to total intermediate demand and each final demand agent:



The list of products shown adjusts dynamically with the data base chosen. Hence, check it when switching the data base.

Note that the MRIO extensions requires that MRIO split factors were generated during data base generation:



Generate MRIO split factors ☒

## 5.8 Imperfect competition model

Activating the Melitz module will generate a new tab with the following input possibilities:

The screenshot shows the 'Imperfect competition model' interface with two tabs: 'Commodities' and 'Model setup'. The 'Commodities' tab displays a list of commodity codes (e.g., pdr-c, wht-c, gro-c, v\_f-c, osd-c, c\_b-c, pfb-c, ocr-c, etc.) in a grid. The 'Model setup' tab shows a dropdown menu for 'Melitz' with options: 'Only aggregated demand (no heterogenous firms)', 'Krugmann', and 'Melitz'. Below the dropdown, there are input fields for 'Pareto distribution lower limit' (0.5), 'Shape parameter' (4.60), and 'Max cost share for HET own domestic interm. demand' (0.00). A checkbox for 'Initialize Melitz model for shock' is checked.

Under Melitz commodities, the commodities / sectors can be chosen the parameter of the Pareto-distribution of the productivities defined.

The model allows for three different base configurations:

1. “Only aggregated demand, not heterogenous firms” changes the demand structure of the standard GTAP such that only one Armington nest shared by all agents is present. The substitution elasticity is defined as in the case when the full Melitz model is used. The remaining equations of the Melitz model are absent and replaced by the equations of the standard GTAP model. That allows dis-entangling impacts of the full Melitz model from the structural changes on the demand side.
2. “Krugman”: Monopolistic competition at industry level, the number of firms in each sector and region is endogenous and defines the number of varieties.
3. “Melitz”: Melitz model with fix costs and number of firms operating differentiated by trade link, additionally to the industry detail from Krugmann.

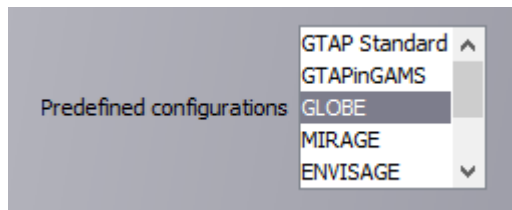
The checkbox “Fix cost nest” introduces a differentiation between variable and fix costs input composites where the latter only comprises primary factors, at least as long as there is some minimal primary factor cost share left in the variable input composite. The spinner termed “Max cost share of HET domestic interm. demand” allows switching off love of variety for domestic intermediate demand by same industry depending on the costs share. Only sector-regional combinations with a share above the threshold will be excluded. Setting the threshold to unity will hence leave the love of variety effect switch on for all domestic intermediate demand, while zero will switch it off on all cases.

Note that with the MRIO extension switch one, bi-lateral demand by agent is not based on equal shares, but reflects the shares used in the MRIO-extension.

## 5.9 Pre-configurations

In order to replicate as close as possible some existing well-known global CGEs, pre-configurations are defined which use the modular structure of CGEBox to mimic these models.





### 5.9.1 Standard GTAP model

The details of the implementation of standard GTAP model in GAMS are not discussed in here, see Van der Mensbrugghe 2015. It is however worthwhile to note that using the pre-configuration does not substitute prices out such that the full equation system can be seen. This option will however slow down solution for large models.

### 5.9.2 GTAPinGAMS

The GTAPinGAMS by Tom Rutherford is quite similar to the standard GTAP model with however the following differences which are mostly simplification which render that model ideal as a didactic model:

- There is no global bank mechanism, foreign savings are fixed
- Regional savings and investment demand are fixed as well

### 5.9.3 GLOBE

The GLOBE pre-configuration uses the following attributes. The page numbers below relate to the following GLOBE documentation: McDonald S., Thierfelder, K. (2014): Globe v2: A SAM Based Global CGE Model using GTAP Data.

- The agents share both Armington nests (see page 42), i.e. the Armington aggregation module is switched on
- The regional household is switched off by using a dummy implementation of myGTAP with one household. No remittances etc. are considered (see page 52).
- The government is closed by savings which implies that real consumption is fixed (see page 58). The distribution of government demand to commodities is based on Leontief coefficients.
- Household demand uses a fixed saving rate and adjusts its total spent on consumption to exhaust household income.
- Foreign savings are fixed
- The capital account balance for each country is closed based on flexible exchange rates
- A LES demand system is used
- The VA nest comprises a sub-nest for different labor qualities
- VA and the intermediate composite can be substituted
- There is a CET matching the Armington

The following specifics of GLOBE can currently not be captured by CGEBOX:

- The GLOBE region which differentiates bi-lateral margins trades and thus allows to define bi-lateral trade balances
- The rather flexible tax rate adjustment closures
- The Armington prices are gross of taxes, i.e. consumption taxes are equalized across agents (see page 53 for the definition in GLOBE).

The reader should note that alternative closures are possible, but then, the pre-configuration cannot be used and the different options must be set manually.

#### 5.9.4 MIRAGE

The MIRAGE configuration is defined as follows:

- The GTAP-AGR module is switched on
- There is sub-nest under the VA nest substituting between labor and capital
- All industry sectors are depicted a la Krugmann
- The agents share the Armington specification in the Krugmann model (which endogenous preference shifters depicting love for varieties)
- A LES demand system is used
- There is separate government account closed by government savings
- A specific parameter set is used

#### 5.9.5 ENVISAGE

- The GTAP-AGR and GTAP-E modules are switched on
- A LES demand system is used
- Capital vintages are differentiated in comparative static mode, considering 5 years of depreciation
- A separate government account is closed under fixed savings adjusting direct taxes
- A specific parameter files is used which e.g. introduced a CET on the supply side

#### 5.9.6 CGEBOX

- The GTAP-AGR, GTAP-AEZ and GTAP-E modules are switched on
- Capital vintages are differentiated in comparative static mode, considering 5 years of depreciation

#### 5.9.7 CGEBOX+

As above, but additional, the Melitz module is used for industry sectors

For further details on these pre-defined configurations, consult directly the GAMS code

### 5.10 Options for solving the model: Algorithm

The algorithm panel is available in two layouts. The default one is for normal user levels:

| CC Damages                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Imperfect competition model | SPE | Closures | Output options | Solve options |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|-----|----------|----------------|---------------|
| <div> <div>Algorithm and solve</div> <div> <div>Solution algorithm</div> <div>Algorithm for CNS</div> <div>Algorithm for NLP</div> </div> <div> <div>CNS</div> <div>CONOPT4</div> <div>CONOPT4</div> </div> </div> <div> <div>Solve</div> <div> <div>Maximal seconds for each solve</div> <div>for first solve</div> <div>Number of repeated solves</div> <div>Stop after solving year</div> <div>Presolve rounds</div> <div>SAM considered imbalanced at 1.E-</div> </div> <div> <div>1.000</div> <div>30</div> <div>3</div> <div>-1</div> <div>0</div> <div>6</div> </div> </div> |                             |     |          |                |               |

Allowing to solve the model multiple times (number of repeated solve) is useful for the following reasons with CONOPT4 as a solver. The first solve is done as a CNS (Constrained System of Equations) which is usually faster than solving the problem as an NLP (Non-Linear Programming) problem in CONOPT4. If this CNS problem is not solved to full feasibility, a second attempt to solve the same model as an NLP is made. This implies no structural change to the model, solely a dummy objective is added. This can help, for instance, in cases where the model comprises some highly non-linear equations. Equally, with each repeated solve, the feasibility tolerances are slightly relaxed.

If “Stop after solving year” is different from -1, the model will be solely generated for the chosen point as the iteration limit is set to zero while solution printing is switched on. This can be helpful to check if the shock is correctly set-up by inspecting which equations are infeasible. Note that the pre-solves are performed with this option. This option can hence also be used to check how successful the pre-solves are in reducing infeasibilities. For information on the pre-solves, please refer to the next section. The thresholds for SAM balance tests can be increased for badly defined input data set.

The steering panel is available in a second version for developers and debuggers which gives additional control on the solution and benchmarking process:

The model can be solved

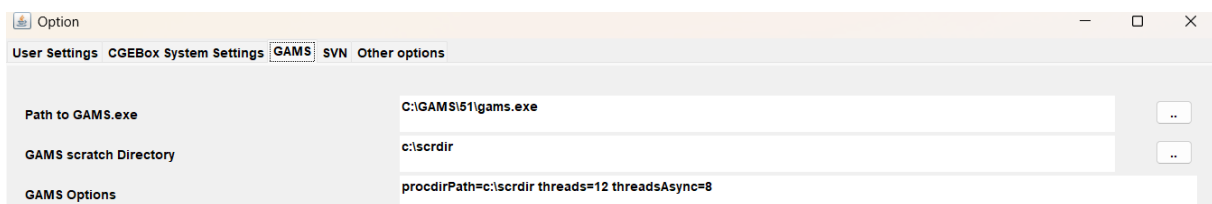
1. as a constrained system of equations (CNS, formally on demand solved with a dummy objective to yield an NLP which might help when solving with CONPT4)
2. an MCP (Mixed Complementary Problem). The latter option is interesting if tax rates are to be endogenized e.g. under emissions ceilings, trade or production quotas, features not comprised in the standard model.
3. EMP (Extended Mathematical Programming). This was explored to endogenous policy instrument and letting government play against each other. Requires a specific set-up of the model's equation with an additional steering file. Not intended for production runs.

The settings shown above are recommended defaults:

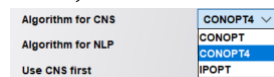
Solving the model as a constrained system of equations CNS - especially in connection with pre-solves, see below – has proven to be generally faster for larger models (> 250.000 transactions in SAM) compared to MCP. The CNS solution will also automatically check if the model is square. MCP gives additionally flexibility such as introducing Tariff Rate or production quotas. Equally, the MCP

version will automatically check for unpaired equations and variables. If during development of extensions or shock files, GAMS throws the error of a non-square model, switching to MCP might help to find unpaired, but not free variables. It is possible to first try to solve the model as CNS, and if that fails, switch to MCP.

The standard option file is 1. For CONOPT4, it uses a feasibility tolerance of 1-E7 and introduces definitional equations also for a square NLP model (flg\_forDefc=1). The log output is set to each iteration and multiple starts are not allowed (num\_rounds=1). The maximal allowed variable limit is 1+E20. The pre-processor reports infeasible structures up to 100 linked variables/equations. The option file 2 relaxes the feasibility tolerance to 5.E-7 with otherwise identical settings. It is automatically chosen if the model cannot be solved in a first attempt. Option file 3 relaxes the convergence tolerance further, and is not intended for production runs, but for debugging purpose. Both with MCP and CNS, the above-mentioned combinations of solver settings have proven to work best. The user can also choose which solver to use for NLP and CNS. CONOPT4 which parallelizes e.g. matrix factorizations has shown considerable speedups for very large models and is now the recommended NLP solver. The current default set by GAMS is that CONOPT4 can use up to four threads. Users can use the settings dialog to allow solvers to use more thread (such as threads=12). Note that during grid solves, the “threadsAsync=8” allows GAMS to use up to 8 threads in parallel during grid solution (pre-solves, filtering and some other applications in CGEBox). The default of -1 would leave solely one thread available for other processes on the machine.

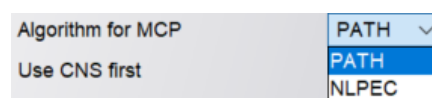


For older GAMS version, CONOPT4 was not available or only as a beta version, the user can turn



back to CONOPT: . IPOPT comes for free with a GAMS license and will generally not be able to solve larger models.

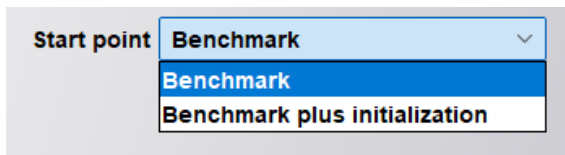
If one of the MCP options is chosen, the MCP algorithm can be chosen:



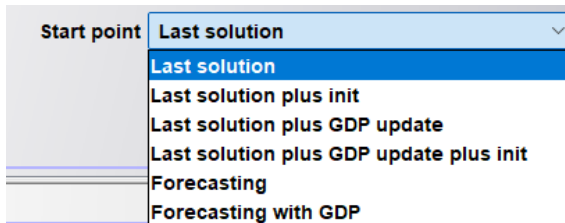
NLPEC is not a solver, but a model re-formulation tool. It will convert the MCP as defined from the pairing to an NLP which is then delegated for solution to CONOPT4. This has proven to solve also very large MCP problems. It allows using MCP without a license for PATH as a dedicated MCP solver to solve GE problem in CGEBox comprising complementary condition, such as Tariff-Rate Quotas.

If the model is solved as an CNS CONOPT4 will use a somewhat simpler and faster algorithm compared to NLP which might however sometimes fail. It is generally recommended to try “Use CNS first”. Less experienced users are generally discouraged to switch to developer level and change the settings e.g. under “Options”, especially relating to variable substitutions and scaling.

The code supports different options to define a start point for the next solve<sup>121</sup>. The available options differ for comparative-static runs:



and recursive-dynamic ones:



For shock slicing (mimic-comp static) the options with “plus GDP update” are taken out.

These options can be combined with pre-solves. The forecasting options defined in the next section are available with recursive-dynamic model runs, only. The standard option is to use the last solution (see *solve/iterloop.gms*). For a comparative-static run, this implies that the benchmark values will be used as the starting point. The GDP update requires the either “G-RDEM” is used or the shock file defines the GTAP growth on a parameter *p\_growthRate*. Otherwise, it is simply neglected. The meaning of adding “init(ialization)” are discussed below.

```

$iftheni.rgdpUpdate "%rgdpUpdate%"=="on"
*
*   --- forecasting based on GDP changes: copy national GDP and population growth to sub-regional units
*
p_growthRate(subr,"GDP|PPP",tSim) = sum(r_r(subr,disr), p_growthRate(disr,"GDP|PPP",tSim));
p_growthRate(subr,"population",tSim) = sum(r_r(subr,disr), p_growthRate(disr,"Population",tSim));
p_growthRate(rs,"GDPperCap",tSim) = p_growthRate(rs,"GDP|PPP",tSim) - p_growthRate(rs,"Population",tSim);

$endif.rgdpUpdate

```

Generally, when GDP updates are used, quantity variables (outputs, intermediate and final demands etc.) are shifted with the real GDP growth rate. For factor prices, their change between the two previous solves relative to the GDP growth rate is used to define an update factor. For other prices, the change in the GDP price index between the two previous solves relative to the GDP growth rate is used to define an update factor.

### 5.10.1 First-order Taylor approximation to define starting points

In order to speed up processing in recursive-dynamic runs or when shock slicing is used, the code can use a first-order Taylor approximation to estimate a starting point for the next solve (see *gams/solve/forecast.gms*), called “forecasting” for selected variables *x*:

<sup>121</sup> For a more detailed discussion of the different solving options, please refer to Britz (2025): How to solve large recursive-dynamic Computable General Equilibrium models, [https://ageconsearch.umn.edu/record/352149/files/Dispap\\_1\\_2025\\_Britz.pdf](https://ageconsearch.umn.edu/record/352149/files/Dispap_1_2025_Britz.pdf).

$$x_t = x_{t-1} + (years_t - years_{t-1}) \frac{x_{t-1} - x_{t-2}}{years_{t-1} - years_{t-2}}$$

The approximation is restricted to a maximum relative change defined by the argument *maxR* between two consecutive years. If the current time step *p\_timeStep(tSim)* is larger, a multiple of *maxR* is used.

```
*
* --- linear forecast, where p_drv is either time or real GDP
*
*      x(t) = x(t-1) + [x(t-1)-x(t-2)]/[p_drvForecast(t-1)-p_drvForecast(t-2)] * alpha * [p_drvForecast(t)-p_drvForecast(t-1)]
*
* Where alpha is variable specific, smaller alpha lead to more conservative changes
* and p_timeStepCor reflects the number of years from t-1 to t and from t-2 to t-1
*
* The forecast is then put in bounds around x(t-1) of (1 * beta))
* where beta is variable specific, smaller beta lead to more conservative changes
* and p_timeStep is reflects the number of years from t-1 to t
*
*      x(t) = min( max[ x(t-1) * (1 + beta,x(t-1) * (1 - beta)],
*                  max[min(x(t-1) * (1 + beta,x(t-1) * (1 - beta))],x(t)]);
*
* $macro m_extrL1(var,t,maxR)          $ ((var.range(t)          ne 0) $ &var(t-2)) \
= min( max(&var(t-1)*(1+maxR*p_timeStep(tSim)),&var(t-1)*(1-maxR*p_timeStep(tSim))), \
max(min(&var(t-1)*(1+maxR*p_timeStep(tSim)),&var(t-1)*(1-maxR*p_timeStep(tSim))), \
var.lo(t)*2,&var(t-1)          + (&var(t-1) - &var(t-2))*p_drvForecast("wor",t)));
*
* $macro m_extrL2(var,d1,t,maxR)       $ ((var.range(d1,t)       ne 0) $ &var(d1,t-2)) \
= min( max(&var(d1,t-1)*(1+maxR*p_timeStep(tSim)),&var(d1,t-1)*(1-maxR*p_timeStep(tSim))), \
max(min(&var(d1,t-1)*(1+maxR*p_timeStep(tSim)), \
&var(d1,t-1)*(1-maxR*p_timeStep(tSim))), \
var.lo(d1,t)*2,&var(d1,t-1)          + (&var(d1,t-1) - &var(d1,t-2))*p_drvForecast(d1,t)));
```

Starting values for variables which drive the variable initialization (see next section) are defined from the previous solution plus a share of the absolute changes from between the two proceeding solves. The share is defined by the parameter *p\_drvForecast*. Two options can be used. The first is the based on time as shown in the formula above. It uses the current time step divided by the last time step. This considers the case where, for instance, the current time step is equal to one, but the previous was a step over two years. In the case, solely half of the changes between the two last solutions would be used.

```
p_drvForecast(rs,t)      = years(t);
if ( tSim.pos gt 2,
    p_drvForecast(rs,tSim) = (p_drvForecast(rs,tSim)-p_drvForecast(rs,tSim-1)) / (p_drvForecast(rs,tSim-1)-p_drvForecast(rs,tSim-2));
);
```

The second option is to use changes in real GDP per capita in case of growth. Tests have shown that it tends to perform worse. These tests also revealed that using absolute changes is more stable compared to relative ones, it also accommodates cases such as the BOT which can switch their sign. Tests to use also a second-order approximation were not successful, they provoked infeasibilities in pre-solves in some cases by producing implausible starting points for specific regions and markets.

Afterwards, the variables to project are updated using these macros. As seen below, the first argument of the macro is the variable name, followed by the driving sets (indices). Generally, price changes are restricted to +/-25% and quantity changes are (virtually) unrestricted, such as for the case of *xs* below. For specific quantity variables such as total imports *xmt* and exports *xet*, more cautious maximal relative changes are used.

```
xet(rsNat,i,tSim)      m_extrL3(xet,rsNat,i,tSim,1.)
xmt(rsNat,i,tSim)      m_extrL3(xmt,rsNat,i,tSim,1.)
xs(xsFlag(rs,i),tSim)  m_extrL3(xs,rs,i,tSim,10000.)
xds(rsNat,i,tSim)      m_extrL3(xds,rsNat,i,tSim,1.)
ps(xsFlag(rs,i),tSim)  m_extrL3(ps,rs,i,tSim,0.25)
psd(xsFlag(rs,i),tSim) m_extrL3(psd,rs,i,tSim,0.25)
px(xpFlag(rs,a),tSim)  m_extrL3(px,rs,a,tSim,0.25)
pmt(rsNat,i,tSim)      m_extrL3(pmt,rsNat,i,tSim,0.25)
```

Note that forecasting can be also with shock-slicing. In this mode, the starting point for the next slice is the previous slice plus the changes between last two slices solved.

### 5.10.2 Variable initialization before solving

Compared of using the last solution for updating block with variables, potentially with the same multiplier derived from targeted real GDP growth, the Taylor-approximation will lead to different quantity and price changes for each forecasted variable, provoking potentially a high number of larger infeasibilities at the resulting start point. In order to provide instead a starting point to the model where many equations are feasible, the projected variable values from the Taylor-approximation combined with exogenous changes from the shock (tech shifters, changes in tax rates, population etc.) are used to recursively define starting values for most variables in the model (see “*gams/iniVars/iniVars.gms*”). To give an example: composite prices for technology nests, the value-added nests and the intermediate demand nests are defined from given product and factor prices; from there, input demands are defined. As a consequence, a basis which usually small infeasibilities in some equations, only, is passed on to the solver.

If “Benchmark plus initialization” or “Last solution plus GDP update plus init” is chosen, the code for the initialization is also active. In case of a comparative-static run, starting prices reflect changes in tax rates from the shock and further calculation take additional changes from the shock into account, such as productivity shifters. This might provide a better start point to the solve.

### 5.10.3 Pre-solves

For large models, pre-solves can be used during which single region CGEs for each region in the model are solved independently. That process can be repeated several times to inform the individual country CGE models about changes in f.o.b. price and import demand in others regions. The single region models are only introduced as an intermediate step towards the solution of the full model. The pre-solves aim at providing a good starting point for a solve of the full global model. Therefore, the structure of the intermediate single-country CGE models needs to reflect the bi-lateral trade relations and other linkages between regions in the final multi-regional model to solve. Accordingly, from the global model, solely the bi-lateral import and export demand functions of trading partners at fixed total import demands and export supply quantities and related price indices are kept. These single country CGEs thus differ from regular single country CGEs which are usually solved either at fixed international prices or by rendering import and export prices depending on import and export quantities, not considering different trading partners. The layout of the single country models therefore differs somewhat from the usual single country CGE structure and solely uses equations and variables already comprised in the model at unchanged parameterization.

In the standard GTAP model, there is no CET on the supply side. We therefore simply assume that a change in demand for bilateral exports of a single country has a negligible impact on the supply price of each importer. Accordingly, we drive the single country models with fixed fob prices on the import side. On the export side, however, we exploit the Armington structure: we feed changes in the country’s supply prices in the Armington share equations of the importers at fixed total import and prices. The reader should note that solving the model also with fixed prices for exports would not help much in providing a good starting point for the model. Without a CET, the export price is by definition equal to the supply price in the domestic market. If one would hence fix the export prices, the domestic prices in the current solve would also be fixed. There would hence be no updated price information be passed along from solving the single country models to the next round of solves.

Further links in the full models between regions are based on the global factor price index which is fixed during a global model solve to reflect that there is no money solution, i.e. the behavioral equations are of degree zero in prices. However, during single model solves, fob prices are fixed such

that it is impossible to fix at the same time the factor price index. In order to still drive the single models during iterations towards a global factor price index of unity, the fob prices are divided before each solve by the currently calculated global factor price index as are the saving price index and the factor prices of all regions which enter fixed the single model solves.

Finally, we reflect the global bank mechanism by calculating iterations the average expected return at global level, based on weighting with regional net investments. A heuristic estimates a change in foreign savings which should drive the country's expected rate towards the global average.

There are a number of options available for the pre-solves:

1. **The number of pre-solves.** For larger shocks and a higher number of regions, three iterations are recommended as the pre-solves are relatively cheap. In recursive dynamic runs, one round of pre-solves is often giving the fastest results.
2. **Use of grid solves.** This option solves the single country models in parallel. The grid solve is the recommended option. Solving the single countries one after one implies hard to control changes in import and export prices depending on the order in which the regions are defined. With the grid solve, all regions use the same fixed import and export prices of trading partners during one solve round. Equally, even cheap modern laptop are multi-core machines which benefit from the option to solve models in parallel.
3. **Stopping the pre-solve loop if all single models are feasible.** Tests have shown that especially in recursive-dynamic run or with shock slicing, the first-order approximation combined with variable initialization provide typically excellent starting points such that the sum of infeasibilities after a first round of solves when combining the solutions into the global model is typically already quite small.

CNS models based with CONOPT have proven to benefit much more from the pre-solves than the MCP version. The paper by Britz and Van der Mensbrugghe (2016)<sup>122</sup> documents how the pre-solves impacted on solution behavior at that time. Since then, improvements in solvers, processor speed and the number of cores available even in laptops, as well as code improvements in CGEBox mean that using pre-solves can slow the overall model solve for smaller shocks and medium sized models. For medium sized models (less than 20 regions at full product detail of the GTAP Data Base), usually, pre-solves might not gain much. The working paper by Britz (2025)<sup>123</sup> reports on different options discussed above for solving larger models.

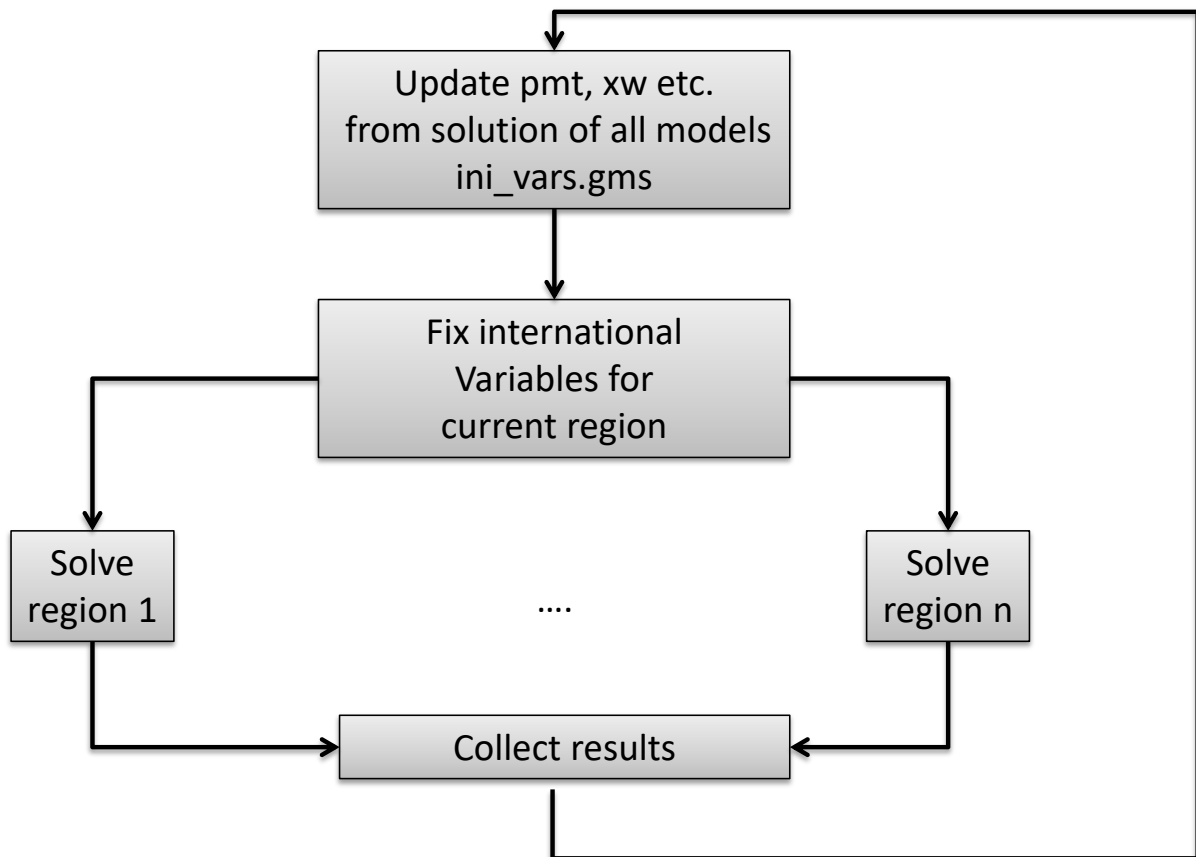
The basic process of the pre-solves is depicted in graphic below. In between iterations, variables in each regional model which depend on other regions' endogenous variables such as their c.i.f. prices driven by f.o.b. exporter prices are updated. Next, variables not related to the current region but comprised in the single region model are fixed and the model is sent to the grid for solution. While the solve process for this model starts in the background, the next regional model is prepared and sent to grid. The code then waits until all model solutions are returned and collects their solutions to start the next iteration.

---

<sup>122</sup> Britz, W., & van der Mensbrugghe, D. (2016). Reducing unwanted consequences of aggregation in large-scale economic models-A systematic empirical evaluation with the GTAP model. *Economic Modelling*, 59, 463-472

<sup>123</sup> Britz, W. (2025). How to solve large recursive-dynamic Computable General Equilibrium models, [https://ageconsearch.umn.edu/record/352149/files/Dispap\\_1\\_2025\\_Britz.pdf](https://ageconsearch.umn.edu/record/352149/files/Dispap_1_2025_Britz.pdf)





To solve the model for one region, only, all model equations therefore carry a dynamic set *rs*, indicating the region(s) in current solve. The pre-solve code can be found in “*gams\solve\presolve.gms*”. The number of pre-solves is defined via the interface and drives a loop:

```
loop( presolve,
```

In that loop, several steps are performed:

1. A heuristic determines the foreign savings for that region, later fixed during the solve, using an interpolation over previous solves:

```

-----
p_roreIter(rsNat,"savfIter",preSolve) $ p_roreIter(rsNat,"distMin",preSolve)
= p_roreIter(rsNat,"savfMin",preSolve)
+ [ ( p_roreIter(rsNat,"savf",preSolve-1)-p_roreIter(rsNat,"savf",preSolve) )
    / (p_roreIter(rsNat,"rorg",preSolve-1)-p_roreIter(rsNat,"rorg",preSolve)) )
    * (p_roreIter("wor","rorg",preSolve)-p_roreIter(rsNat,"rorg",preSolve)) ]
    $ (p_roreIter(rsNat,"rorg",preSolve-1)-p_roreIter(rsNat,"rorg",preSolve));

savf.l(rsNat,tSim) $ p_roreIter(rsNat,"savfIter",preSolve) = p_roreIter(rsNat,"savfIter",preSolve);

```

2. A loop over the regions is performed and dynamic sets are updated which indicate the nations in the current model to generate *rsNat* and sub-regions therein *rs*, including the nation itself. The set *rrComb* comprises all bi-lateral combinations where the region is either an importer or an exporter.

```

loop( rr $ rNat(rr),

    --- include only current region in model

    option kill=rs;
    option kill=rsNat;
    rs(rr)          = YES;
    rsNat(rr)       = YES;
    rs(subr) $ r_r(subR,rr) = YES;

    --- bilateral combinations of trade flows to be included:
    all import and export flows of current regions

    option kill=rrComb;
    rrComb(rNat,rNat1) $ (rsNat(rNat) or rsNat(rNat1)) = YES;

```

3. The variables for the current region are initialized from the given results so far in “*solveIniVars.gms*” and negative respectively very small variables reset in “*solveResetNegative.gms*”. That step is only included in the loop itself if the models are solved sequentially. If the models are solved in parallel on a grid, that step is performed outside of the loop over the regions at the beginning of the iteration.
4. Nest, the trade related variables for all regions are fixed and next, bounds introduced for the region in the current solve, e.g.

```

xmt.fx(rNat,i,tSim) = xmt.l(rNat,i,tSim);

xmt.lo(rsNat,i,tSim) $ xmtFlag(rsNat,i) = 0;
xmt.up(rsNat,i,tSim) $ xmtFlag(rsNat,i) = +inf;

```

5. The model for that region is solved:

```

$$iftheni.mt "%firstModeltype%"=="MCP"

    --- MCP solution
    solve gtap using mcp;
$$else.mt

    --- presolve in CNS mode, switch to NLP for current region if infeasible in last presolve
    (NLP might come up with something useful)

    if (useCnsFirst and (p_rorIter(rr,"numinfes",presolve-1) eq 0),
        solve gtap using cns;
    else
        solve gtap using nlp minimizing dummy;
    );

$$endif.mt

```

The advantages of the pre-solve process compared to solving the full model directly on the shock depend on the size of the model and its complexity, the nature of the shock, as well as the hardware and solver used. For small to medium sized models under smaller shocks, using the pre-solves will typically drive up the overall computing time, especially when CONOPT4 is used that parallelizes part of the solution process. There are also cases where the pre-solves might render it harder to solve the full model and can provoke infeasibilities. It is generally worth a try to use the pre-solves if the solution time of the full model exceeds 30 seconds and a multi-processor computer is used.

The *rrComb* set will ensure that only the bi-lateral flows in which the current region is involved are introduced in the single region model, e.g.:

```

xw(rNat1,i,rNat,t)/xw.scale(rNat1,i,rNat,t) =e= amw(rNat1,i,rNat,t)*xmt(rNat,i,t)/xw.scale(rNat1,i,rNat,t)

* [ { (pmt(rNat,i,t)/[(1 + imptx(rNat1,i,rNat,t) + mtax(rNat,i,t))
*
* --- Note WB: cif price is defined via macro %pmcif%, normally not a variable in model
*
* %pmcif%(rNat1,i,rNat,t)/chipm(rNat1,i,rNat) ])**sigmaw(rNat,i)
* m_lambdam(rNat1,i,rNat,t)**(sigmaw(rNat,i) - 1)}
*
* $iftheni.smallShare "%hasSmallShare%"=="on"
* $ (not isSmallImpShare(rNat1,i,rNat))
* + m_lambdam(rNat1,i,rNat,t) $ isSmallImpShare(rNat1,i,rNat)
* $endif.smallShare
* ];

```

In the above equation and the case of the current region acting as the exporter *rNat1*, the total price of imports *pmt* and total imports *xmt* of the trade partner *rnat* are fixed, but changes in the current regions f.o.b. price will change the c.i.f. price – along e.g. with shocks to the importers tariffs – and change the simulated export flows *xw*. Thus, at least as long as the exporter's share in the total imports of *rNat* is not very large, the major impact of the Armington structure is captured. Furthermore, as *pmt* and *xmt* will be updated after each pre-solve, the individually regional model solves will iteratively approximate the solution of the global model.

Similarly, if a CET on the export side is present, for the case where *rnat1* acts as an importer, the impact of changing import demand on the f.o.b. price faced is approximated:

```

*
* --- This equation substitutes out the equilibrium condition for bilateral trade, equation (56) in VDM 2018
* (makes it possible to drop equation (68) in VDM 2018)
*
* peeq(rNat,iIn(i),rNat1,ts(t)) $ (m_xwFlag(rNat,i,rNat1) $ psFlag(rNat,i,"pe") $ rrComb(rNat,rNat1)) ..
*
* 0 =e= ( m_xws(rNat,i,rNat1,t)/xw.scale(rNat,i,rNat1,t)
*         - gw(rNat,i,rNat1,t)*xet(rNat,i,t)/xw.scale(rNat,i,rNat1,t)
*         * [ 1
*             + { (pe(rNat,i,rNat1,t)/pet(rNat,i,t))**omegaw(rNat,i) } $ isSmallExpShare(rNat,i,rNat1)
*             ] $ (not isSmallExpShare(rNat,i,rNat1)) ]
*         ) $ (omegaw(rNat,i) ne inf)
*

```

The foreign savings are fixed during the pre-solves such that the related equation is not introduced:

```

*
* --- Determination of foreign savings, equation (80) in VDM 2019
*
* savfeq(rsNat(rNat),ts(t)) $ ( ( (RoRFlag eq equalReturnToInv)
*                                or ((RoRFlag eq fixedAllocationOfInv) $ (not rres(rNat))))
*                                $ (savf.range(rNat,t) ne 0))..
*
* 0 =e=

```

The same holds for the capital account equation which in the full model determines the foreign savings for the residual region, see the first dollar condition which requires that at least two regions with unfixed foreign savings need to be present in the model:

```

*
* capAccteq(ts(t)) $ ( ( (sum(rsNat $ (savf.range(rsNat,t) ne 0),1) gt 1)

```

The parameter *p\_solveReport* reports various statistics such as the time needed or potential infeasibilities:

```

p_solveReport(tSim,"Presolves, maxSumInfes") = sMax( (preSolves,rNat), p_roreIter(rnat,"sumInfes",preSolves));
p_solveReport(tSim,"Presolves, infes, step") $ p_solveReport(tSim,"Presolves, maxSumInfes")
= sMax( (preSolves,rNat) $ p_roreIter(rnat,"sumInfes",preSolves),preSolves.pos);
p_solveReport(tSim,"Presolves, maxNumInfes") = sMax( (preSolves,rNat), p_roreIter(rnat,"numInfes",preSolves));
p_solveReport(tSim,"Presolves, seconds") = timeElapsed - p_solveReport(tSim,"Presolves, seconds");

```

That allows also tracking the solution behaviour in recursive-dynamic solves:

Output

Output string for solveReport

It is stored after each year in the scratch directory. The GDX viewer from GAMS Studio allows keeping a file open for read while other processes update the file in the background. The user can hence open that file in the viewer and check there the solution behaviour while the simulation continues:

| Entry | Symbol        | Type | Dim | Nr Elem |
|-------|---------------|------|-----|---------|
| 1     | p_solveReport | Par  | 2   | 29      |

|                       |               |               |               |
|-----------------------|---------------|---------------|---------------|
| p_solveReport(*, *)   |               |               |               |
|                       | t00           | t01           | t02           |
| timeElapsed           | 21.959000     | 151.863000    | 185.751000    |
| prep,seconds          | 0.163000      | 98.649000     | 1.188000      |
| Presolves, seconds    |               | 16.959000     | 17.639000     |
| full Solve, seconds   | 8.610000      | 12.319000     | 12.644000     |
| Full solve, sumVar    | 133090.000000 | 134275.000000 | 134275.000000 |
| Full solve, modelStat | 1.000000      | 16.000000     | 16.000000     |
| Full solve, solveStat | 11.000000     | 1.000000      | 1.000000      |
| Full solve, nTries    | 1.000000      | 1.000000      | 1.000000      |
| Full solve, resUsd    | 0.359000      | 4.843000      | 5.125000      |
| seconds               | 21.219000     | 128.586000    | 32.165000     |

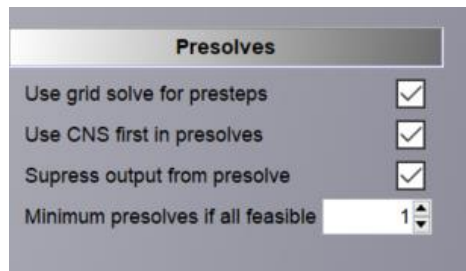
The report comprises information on the seconds for certain steps, model statistics (number of variables and equations, total and non-linear non-zero Jacobian entries, non-zero Hessian entries), on memory use.

The use of the pre-solves is switched off by settings the number of pre-solves to zero on the tab “Algorithm settings”:

LES\_CES PE G-REDM Imperfect competition model SPE Closures Output options Algorithm settings GAI

| Solve                                                                  | Options                                                                                                                                                                       |
|------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Maximal seconds for one solve <input type="text" value="240"/>         | Substitute out import/domestic prices <input checked="" type="checkbox"/> Armington prices <input checked="" type="checkbox"/> quantities <input checked="" type="checkbox"/> |
| Number of repeated solves <input type="text" value="3.0"/>             | Global scaling factor, 1.E- <input type="text" value="3"/>                                                                                                                    |
| Zero iterations benchmark <input type="checkbox"/>                     | Scaling exponent <input type="text" value="1.0"/>                                                                                                                             |
| Zero iterations shock <input type="checkbox"/>                         | Minimum scaling factor, 1.E- <input type="text" value="0"/>                                                                                                                   |
| Kill bounds <input checked="" type="checkbox"/>                        | Threshold for transactions, 1.E- <input type="text" value="10"/>                                                                                                              |
| Presolve steps with fixed import prices <input type="text" value="0"/> | Threshold for SAM imbalance tests, 1.E- <input type="text" value="6"/>                                                                                                        |

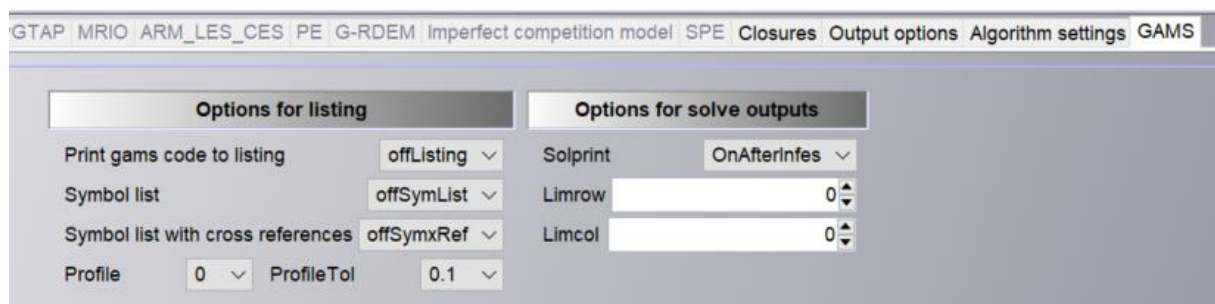
If a non-zero value is chosen, additional controls for the pre-solve process become available if the user level is set to developer:



Switching the grid solve off makes sense only to debug the process. Switching off the CNS in the first solve already might help with CONOPT4 in cases of (1) highly non-linear relations or (2) if CONOPT4 bounces against a bound of a variable which stops the search for a smaller sum of infeasibilities in CNS mode. If the single region models are solved as an NLP with a dummy objective instead, CONOPT4 might find a solution or at least a solution with a small sum of infeasibilities, but will require more time in the standard case of a feasible solve. The pre-solve script will automatically solve the next iteration in NLP mode with the previous one was infeasible.

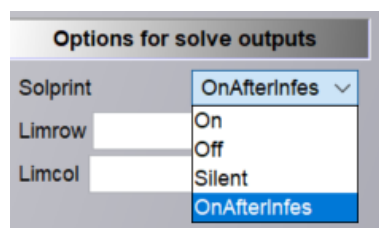
If solution listings are to be checked for the pre-solves, the output can be activated. Finally, it is possible to leave the loop of pre-solves early if all single model region were feasible. This is especially useful if the solver might not find a feasible solution in the early pre-solves in recursive-dynamics runs. Allowing, for instance, for up to three pre-solve rounds and leaving the minimum pre-solves at one will in many cases mean that just one round of pre-solves is performed.

## 5.11 GAMS



The left-hand side controls what is outputted to the listing. For details, refer to the GAMS documentation.

The recommend setting for “solprint” is “OnAfterInfes” as it makes mostly limited sense to inspect the typically quite long model listing if the model is solved sucessfully. For complex simulation, “OnAfterInfes” offers the possibility to generate a listing if the model could not be solved without infeasibilities in the last try in each solved period.



Note that all symbols (variables, equations) can be stored in a GDX container with GDX output switched on which is often more efficient to inspect individually variables compared of using the listing file.

A daunting exercise provides debugging infeasible model solutions provoked by bounds on variables. Non-zero lower bounds for many variables are introduced by default to prevent math trap errors such as from division by zero or logarithms of zero values during the solution process, which typically will mean that the solver will fail. Finding variables where the bounds might have chosen too tight in a large model listing is almost impossible. Therefore, if the last model solve fails with some larger infeasibilities, embedded python code will automatically generate log output on infeasible equations and variables which are at their bounds to ease the debugging process.

## 6. Post-model processing

### 6.1 Decomposing welfare changes in CGEBox

- Wolfgang Britz, December 2018 -

#### 6.1.1 Background

When analyzing impacts of policy changes or other shocks, the simultaneous adjustments of various elements impacting welfare such as factor and commodity prices, tax income or foreign savings render it challenging to understand what drives the simulated welfare change. CGEBox offers two approaches to decompose welfare changes which stem from different “schools”. In the world of partial equilibrium modelers, income is fixed, but products are often close substitutes such that cross-price effects in demand matter. Therefore, that community often decomposes the welfare change with regard to individual product price changes. The first approach detailed next draws on that concept and adds also the perspective of changes in individual factor prices and quantities.

The general equilibrium community has another focus and typically decomposes welfare with regard to terms-of-trade effects and allocative effects from removing tax distortions. That approach is, for instance, implemented in the GEMPACK version of the GTAP Standard model and provided in CGEBox as well. Both approaches have some overlap, but clearly complement each other and are useful to understand where welfare changes result from.

#### 6.1.2 Decomposition with regard to overall price indices and income components

With regard to overall price indices, we measure the impact from updating the price indices from the benchmark to the simulation points, e.g. for savings:

```
u.1(rsNat,%2) = au.1(rsNat,"%t0%") * uh.1(rsNat,"hhsld","%t0%")**betap.1(rsNat,"%t0%")
               * ug.1(rsNat,"%t0%")**betag.1(rsNat,"%t0%")
               * [us.1(rsNat,"%t0%")*pSave.1(rsNat,"%t0%")/pSave.1(rsNat,%2)]**betas.1(rsNat,"%t0%");
```

Next, we use an EV calculation to express that utility changes in terms of required utility to reach the new income level at base year prices (in file *postModel\calcExp.gms*).

```
p_results(rsNat,"V",%4,%3,"tot",%1,"%version%")
$ (p_results(rsNat,"V","ra","u","tot",%1,"%version%") gt eps)
= (u.1(rsNat,%2)-u.1(rsNat,"%t0%")) * p_results(rsNat,"V","au","mm","tot",%1,"%version%") *
  (p_pConsHlp(rsNat,%2)/betap.1(rsNat,"%t0%")**betap.1(rsNat,"%t0%"))
* [ {( pg.1(rsNat,"%t0%")/betag.1(rsNat,"%t0%"))**betag.1(rsNat,"%t0%")
    }$ axg(rsNat,"%t0%") + 1 $ (not axg(rsNat,"%t0%")) ]
* [ {( pi.1(rsNat,"%t0%")/betas.1(rsNat,"%t0%"))**betas.1(rsNat,"%t0%") }
    $ (p_results(rsNat,"V","ra","betas","tot",%1,"%version%") ge 0)
    + 1 $ (p_results(rsNat,"V","ra","betas","tot",%1,"%version%") lt 0) ];
```

The argument %4 is the code for the change measured, for the saving price index it is “ps”:

```
$batinclude 'postModel/calcExp.gms' '%2' '%2' 'mm' 'ps'
```

Similar statements are found for the private household and government price indices.

To measure impacts of changes in factor income, depreciation and indirect taxes, we only update these components from benchmark to shock, e.g. for factor income excluding changes in depreciation:



```

*
* --- update only factor income
*
    u.l(rsNat,%2) = u.l(rsNat,"%t0%") * (facty.l(rsNat,%2)
        + fdepr(rsNat,%2)*pi(rsNat,%2)*kstock(rsNat,%2)
        - fdepr(rsNat,"%t0%")*pi(rsNat,"%t0%")*kstock(rsNat,"%t0%")
        + yTaxInd.l(rsNat,"%t0%"))/regy.l(rsNat,"%t0%");
$batinclude 'postModel/calcExp.gms' '%2' '%2' 'mm' 'va'

```

For depreciation:

```

*
* --- update only depreciation
*
    u.l(rsNat,%2) = u.l(rsNat,"%t0%") * (facty.l(rsNat,"%t0%")
        - fdepr(rsNat,%2)*pi(rsNat,%2)*kstock(rsNat,%2)
        + fdepr(rsNat,"%t0%")*pi(rsNat,"%t0%")*kstock(rsNat,"%t0%")
        + yTaxInd.l(rsNat,"%t0%"))/regy.l(rsNat,"%t0%");
$batinclude 'postModel/calcExp.gms' '%2' '%2' 'mm' 'dep'

```

And for indirect taxes:

```

$batinclude 'postModel/calcExp.gms' '%2' '%2' 'mm' 'dep'
*
* --- update only indirect taxes
*
    u.l(rsNat,%2) = u.l(rsNat,"%t0%") * (facty.l(rsNat,"%t0%") + yTaxInd.l(rsNat,%2))/regy.l(rsNat,"%t0%");
$batinclude 'postModel/calcExp.gms' '%2' '%2' 'mm' 'xt'

```

These three components above are also found in the EV decomposition by Huff and Hertel 2000.

The total impact of regional income changes is defined as:

```

    u.l(rsNat,%2) = u.l(rsNat,"%t0%") * regy.l(rsNat,%2)/regy.l(rsNat,"%t0%");
$batinclude 'postModel/calcExp.gms' '%2' '%2' 'mm' 'x'

```

For the individual factors, we loop over the factors and update their price and quantities from benchmark and shock

```

loop(f0,
*
*   --- update quantities and prices for current factor f0
*
    pft.l(rsNat,f0,%2) = pft.scale(rsNat,f0,%2);
    xft.l(rsNat,f0,%2) = xft.scale(rsNat,f0,%2);
    pf.l(rsNat,f0,a,%2) $ xfFlag(rsNat,f0,a) = pf.scale(rsNat,f0,a,%2);
    xf.l(rsNat,f0,a,%2) $ xfFlag(rsNat,f0,a) = xf.scale(rsNat,f0,a,%2);

```

From there, we recalculate the factor income:

```

factY.l(rsNat,%2) = [ sum(fm $ xfFlag(rsNat,fm), pft(rsNat,fm,%2)*xft(rsNat,fm,%2))
    + sum((f0,a) $ xfFlag(rsNat,f0,a), pf(rsNat,f0,a,%2)*xf(rsNat,f0,a,%2))
    - fdepr(rsNat,"%t0%")*pi(rsNat,"%t0%")*kstock(rsNat,"%t0%")
    ];

```

And calculate as discussed for other income contributions, the change in utility and the related EV:

```

--- update u and calculate money metric (homothetic preferences ...)
u.l(rsNat,%2) = u.l(rsNat,"%t0%") * (facty.l(rsNat,%2) + yTaxInd.l(rsNat,"%t0%"))/regy.l(rsNat,"%t0%");
$batinclude 'postModel/calcExp.gms' '%2' '%2' 'mm' f0

```



Afterwards, we reset the prices and quantities

```
pft.1(rsNat,f0,%2) = pft.1(rsNat,f0,"%t0%");
xft.1(rsNat,f0,%2) = xft.1(rsNat,f0,"%t0%");
pf.1(rsNat,f0,a,%2) $ xfFlag(rsNat,f0,a) = pf.1(rsNat,f0,a,"%t0%");
xf.1(rsNat,f0,a,%2) $ xfFlag(rsNat,f0,a) = xf.1(rsNat,f0,a,"%t0%");
);
```

### 6.1.3 Decomposition with a focus on prices in changes

In here, we calculate the approximate impact from individual price changes for commodities and factors. To do so for the product prices, the following steps are performed:

1. We update each Armington price, domestic or import price sequentially in the three steps from benchmark to simulation level.
2. Next, we update the price indices (respectively utility in case of the LES type demand system) for the private household, government and investments which enter the top-level utility function of the regional household.
3. Based on these updated price indices, we recalculate the change in utility of the regional household.
4. And calculate the EV equivalent to the simulated change in utility from updating only that single price.

### 6.1.4 Demand nests

The CES demand nest prices need to be updated first to capture their impact on the different demand systems. That requires a recursion, starting with the lowest to the highest level to capture the case of nests nested in other nests:

```
nestsToDo(dNest) = YES;
while( card(nestsToDo),
option kill=curNest;
curNest(nestsToDo) $ (not sum(dNest_n_fd(nestsToDo,nestsToDo1,fdn),1)) = YES;
pdNest(rNat,dNest,fdn,%2)
= [ prod(dNest_i_fd(dNest,i,fdn) $ alphaa.1(rNat,i,fdn,%2),
(m_pa(rNat,i,fdn,%2)/(alphaa(rNat,i,fdn,%2)*(1+(lambdai(rNat,i,%2)-1)$ sameas(fdn,"inv"))))*alphaa(rNat,i,fdn,%2))
* prod(dNest_n_fd(dNest,dNest1,fdn) $ alphaDN(rNat,dNest1,fdn,%2),
(pdNest(rNat,dNest1,fdn,%2)/alphaDN(rNat,dNest1,fdn,%2))*alphaDN(rNat,dNest1,fdn,%2))
* axdNestCD(rNat,dNest,fdn,%2)
] $ (sigmaFDNest(rNat,dNest,fdn) eq 1)
+ [ sum(dNest_i_fd(dNest,i,fdn) $ alphaa.1(rNat,i,fdn,%2),
alphaa(rNat,i,fdn,%2) * (m_pa(rNat,i,fdn,%2)/(1+(lambdai(rNat,i,%2)-1)$ sameas(fdn,"inv"))))* (1-sigmaFDNest(rNat,dNest,fdn)
+ sum(dNest_n_fd(dNest,dNest1,fdn) $ alphaDN(rNat,dNest1,fdn,%2),
alphaDN(rNat,dNest1,fdn,%2) * pdNest(rNat,dNest1,fdn,%2)) * (1-sigmaFDNest(rNat,dNest,fdn))
) * (1/(1-sigmaFDNest(rNat,dNest,fdn)))
] $ (sigmaFDNest(rNat,dNest,fdn) ne 1);
nestsToDo(curNest) = no;
);
```

The calculation in the loop requires different dual price aggregator for the CD case where the substitution elasticity is unity – first block – or not, as captured in the second block. Once the demand nest prices are updated, the price indices respectively utility for the three elements of the regional household – private, government and investment consumption can be calculated.

### 6.1.5 Private household

For the private household, the approach is different for the LES and the CDE demand.

1. LES

The LES demand has an indirect utility function which can be calculated from income and prices as shown below

```

uh.1(rsNat,"hhsld",%2)
= (yc(rsNat,"hhsld",%2)
  - sum(i $ (xaFlag(rsNat,i,"hhsld") $ alphaLES.1(rsNat,i,"hhsld",%2)),
    gammaLES(rsNat,i,"hhsld",%2)*pop(rsNat,"hhsld",%2) * pa.1(rsNat,i,"hhsld",%2))
  - sum(dNest_n_fd("top",dNest,"hhsld"),
    gammaLES(rsNat,dNest,"hhsld",%2)*pop(rsNat,"hhsld",%2) * pdNest(rsNat,dNest,"hhsld",%2))
)*( prod(i $ (xaFlag(rsNat,i,"hhsld") $ alphaLES.1(rsNat,i,"hhsld",%2)),
  (alphaLES.1(rsNat,i,"hhsld",%2)/pa.1(rsNat,i,"hhsld",%2))**alphaLES.1(rsNat,i,"hhsld",%2))
  * prod(dNest_n_fd("top",dNest,"hhsld") $ alphaLES.1(rsNat,dNest,"hhsld",%2),
  (alphaLES(rsNat,dNest,"hhsld",%2)/pdNest(rsNat,dNest,"hhsld",%2))**alphaLES(rsNat,dNest,"hhsld",%2)))

* 1/uhLES(rsNat,"hhsld",%2)
/ p_welfTemp(rsNat,"", "uh.1", %2)
* uh.1(rsNat,"hhsld", "%t0%");

```

The calculation faces the challenge that the even the relative impact of the price depends on the income level of the private household due to non-homothetic demands. If one uses as shown above in the first block the simulate income at the benchmark (%2 index) and updates all prices, the utility of the private household after all price change and all adjustments in income flows would be simulated, i.e. the full utility change and not only the one attributed to price changes. We therefore correct for the total change in utility from benchmark to the simulation outcome with the last term, even if that is somewhat ad-hoc.

The last statement calculates the changes in the price index for the private household necessary to reach that utility.

## 2. CDE

The expenditure function for the CDE case is only implicitly defined. We therefore refrain from fully consistent approach. Instead, we update the Z parameters based on the price change:

```

zCons.1(rsNat,i,"hhsld",%2) $ pa.1(rsNat,i,"hhsld", "%t0%")
= zCons.1(rsNat,i,"hhsld", "%t0%") * (pa.1(rsNat,i,"hhsld",%2)**bh(rsNat,i,%2))
  / (pa.1(rsNat,i,"hhsld", "%t0%")**bh(rsNat,i,%2));

zConsDN.1(rsNat,dNest,"hhsld",%2) $ pdNest.1(rsNat,dNest,"hhsld", "%t0%")
= zConsDN.1(rsNat,dNest,"hhsld", "%t0%") * (pdNest.1(rsNat,dNest,"hhsld",%2)**bhdN(rsNat,dNest,%2))
  / (pdNest.1(rsNat,dNest,"hhsld", "%t0%")**bhdN(rsNat,dNest,%2));

```

These parameters must add up to unity, see the following equation comprised in the GAMS code of the CDE demand system:

```

upeq(rs(r),h,ts(t)) $ sum(i,xaFlag(r,i,h)) ..
1 =e= sum(i $ (xaFlag(r,i,h) $ (not sum(dNest_i_fd(dNest,i,h),1))), zcons(r,i,h,t)/bh(r,i,t))
*
* --- add sub-nests under final demand function, not part of standard GTAP model
*
+ sum(dNest_n_fd("top",dNest,h) $ alphaDN(r,dNest,h,t), zconsDN(r,dNest,h,t)/bhdN(r,dNest,t));

```

We therefore calculate and apply a scaling factor:

```

zConsSum(r,"hhsld",%2)
= sum(i $ (xaFlag(r,i,"hhsld") $ (not sum(dNest_i_fd(dNest,i,"hhsld"),1))), zcons(r,i,"hhsld",%2)/bh(r,i,%2))
+ sum(dNest_n_fd("top",dNest,"hhsld") $ alphaDN(r,dNest,"hhsld",%2), zconsDN(r,dNest,"hhsld",%2)/bhdN(r,dNest,%2));

zCons.1(rsNat,i,"hhsld",%2) = zCons.1(rsNat,i,"hhsld",%2)/zConsSum(rsNat,"hhsld",%2);
zConsDN.1(rsNat,dNest,"hhsld",%2) = zConsDN.1(rsNat,dNest,"hhsld",%2)/zConsSum(rsNat,"hhsld",%2);

```

From there, the budget shares are defined:

```

zConsSum(rsNat,"hhsld",%2)
= sum(i $ (xaFlag(rsNat,i,"hhsld") $ (not sum(dNest_i_fd(dNest,i,"hhsld"),1))), zcons(rsNat,i,"hhsld",%2))
+ sum(dNest_n_fd("top",dNest,"hhsld") $ alphaDN(rsNat,dNest,"hhsld",%2), zconsDN(rsNat,dNest,"hhsld",%2));

xcshr(rsNat,i,"hhsld",%2) = zcons(rsNat,i,"hhsld",%2)/zConsSum(rsNat,"hhsld",%2);

```

Which allows the calculation of the changes in the price index:

```
p_welfTemp(rsNat,j,"pcons",%2)
= sum(i $ xaFlag(rsNat,i,"hhs1d"), xcshr(rsNat,i,"hhs1d",%2)*pa.1(rsNat,i,"hhs1d",%2)) ;
```

### 6.1.6 Government and investments

For the government, the consistent price CD or CES index is used:

```
p_welfTemp(rsNat,iIn(j),"pg",%2) $ axg(rsNat,%2)
= (sum( gov,prod( i $ ((not sum(dNest_i_fd(dNest,i,gov),1)) $ alphas.1(rsNat,i,gov,%2)),
(p_welfTemp(rsNat,j,i,gov)/alphaa(rsNat,i,gov,%2))**alphaa(rsNat,i,gov,%2))
* prod(dNest_n_fd("top",dNest,gov) $ alphaDN(rsNat,dNest,gov,%2),
(p_welfTemp(rsNat,j,dNest,gov)/alphaDN(rsNat,dNest,gov,%2))**alphaDN(rsNat,dNest,gov,%2))
)/axg(rsNat,%2) ) $ (sigmag(rsNat) eq 1)
+ (sum(gov,
sum(i $ ((not sum(dNest_i_fd(dNest,i,gov),1)) $ xaFlag(rsNat,i,gov)),
alphaa(rsNat,i,gov,%2)*p_welfTemp(rsNat,i,gov,%2)**(1-sigmat(rsNat)))
+ sum(dNest_n_fd("top",dNest,gov) $ alphaDN(rsNat,dNest,gov,%2),
alphaDN(rsNat,dNest,gov,%2)*p_welfTemp(rsNat,j,dNest,gov)**(1-sigmat(rsNat)))
)**(1/(1-sigmat(rsNat)))/axg(rsNat,%2) ) $ (sigmag(rsNat) ne 1);
```

The same holds for investments:

```
p_welfTemp(rsNat,iIn(j),"pSave",%2) $ axg(rsNat,%2)
= (sum( inv,prod( i $ ((not sum(dNest_i_fd(dNest,i,inv),1)) $ alphas.1(rsNat,i,inv,%2)),
(p_welfTemp(rsNat,j,i,inv)/alphaa(rsNat,i,inv,%2))**alphaa(rsNat,i,inv,%2))
* prod(dNest_n_fd("top",dNest,inv) $ alphaDN(rsNat,dNest,inv,%2),
(p_welfTemp(rsNat,j,dNest,inv)/alphaDN(rsNat,dNest,inv,%2))**alphaDN(rsNat,dNest,inv,%2))
)/axg(rsNat,%2) ) $ (sigmag(rsNat) eq 1)
+ (sum(inv,
sum(i $ ((not sum(dNest_i_fd(dNest,i,inv),1)) $ xaFlag(rsNat,i,inv)),
alphaa(rsNat,i,inv,%2)*p_welfTemp(rsNat,i,inv,%2)**(1-sigmat(rsNat)))
+ sum(dNest_n_fd("top",dNest,inv) $ alphaDN(rsNat,dNest,inv,%2),
alphaDN(rsNat,dNest,inv,%2)*p_welfTemp(rsNat,j,dNest,inv)**(1-sigmat(rsNat)))
)**(1/(1-sigmat(rsNat)))/axg(rsNat,%2) ) $ (sigmag(rsNat) ne 1);
```

### 6.1.7 Utility change for the regional household and EV

Next, we update the utility of the regional household under its CD-utility function. Note that we are not updating the shares to reflect that the elasticity of private expenditure to a utility change might be affected by the price change – which reflects the simplified treatment of private demand:

```
*
* --- calculate new utility based on updated prices
*
p_welfTemp(rsNat,iIn(j),"u.1",%2)
= au.1(rsNat,"%t0%")
* [uh.1(rsNat,"hhs1d", "%t0%")*p_pconsHlp(rsNat,%2)/p_welfTemp(rsNat,j,"pcons",%2)]**betap.1(rsNat,%2)
* [ug.1(rsNat,"%t0%")*pg.1(rsNat,"%t0%")/p_welfTemp(rsNat,j,"pg",%2)]**betag.1(rsNat,%2)
* [us.1(rsNat,"%t0%")*pSave.1(rsNat,"%t0%")/p_welfTemp(rsNat,j,"pSave",%2)]**betas.1(rsNat,%2);
```

From there, the equivalent variation is defined:

```
*
* --- EV
*
p_results(rsNat,"V","mm",iIn(j),"tot",%1,"%version%")
$ (p_results(rsNat,"V","ra","u", "%t0%") gt eps)
= ( p_welfTemp(rsNat,j,"u.1",%2)-u.1(rsNat,"%t0%") ) * p_results(rsNat,"V","au","mm","tot",%1,"%version%") *
( p_pConsHlp(rsNat,%2)/p_results(rsNat,"V","ra","betap","tot",%1,"%version%")
**p_results(rsNat,"V","ra","betap","tot",%1,"%version%")
* [ { ( pg.1(rsNat,"%t0%")/p_results(rsNat,"V","ra","betag","tot",%1,"%version%"))
**p_results(rsNat,"V","ra","betag","tot",%1,"%version%") } $ axg(rsNat,"%t0%") + 1 $ (not axg(rsNat,"%t0%")) ]
* [ { ( pi.1(rsNat,"%t0%")/p_results(rsNat,"V","ra","betas","tot",%1,"%version%"))
**p_results(rsNat,"V","ra","betas","tot",%1,"%version%") }
$ (p_results(rsNat,"V","ra","betas","tot",%1,"%version%") ge 0)
+ 1 $ (p_results(rsNat,"V","ra","betas","tot",%1,"%version%") lt 0) ];
```

As we sequentially updating over the product, the last product will measure the total EV of all price changes against the benchmark utility. We hence only consider differences between the sequential steps to the attribution to individual products:

```
p_results(rsNat,"V",j,"mm","tot",%1,"%version%")
= p_results(rsNat,"V",j,"mm","tot",%1,"%version%")
- p_results(rsNat,"V",j-1,"mm","tot",%1,"%version%");
```

In order to capture differences between the in parts ad-hoc treatment and prices, the difference to the total is calculated:

```
p_results(rsNat,"V","diff","mm","tot",%1,"%version%")
= p_results(rsNat,"V","ra","mm","tot",%1,"%version%")
- sum(j, p_results(rsNat,"V",j,"mm","tot",%1,"%version%"))
- sum(f0, p_results(rsNat,"V",f0,"mm","tot",%1,"%version%"))
- p_results(rsNat,"V","dep","mm","tot",%1,"%version%")
- p_results(rsNat,"V","xt","mm","tot",%1,"%version%");
```

The results can be found in the following tables in the interface:

Note that due to the interaction between income and price changes, the total EV is not equal to the sum of the individual contributing elements. The results can be found in the following tables in the interface:

| Money metric decomposition, by product and factor, per capita [0] |        |                                                               |
|-------------------------------------------------------------------|--------|---------------------------------------------------------------|
|                                                                   | Design | Sectors and institutions                                      |
| Model overview                                                    | >      | Money metric [Dis. \$]                                        |
| Welfare                                                           | >      | Income                                                        |
| Time series                                                       | >      | Price indices                                                 |
| Households                                                        | >      | Money metric decomposition                                    |
| Markets                                                           | >      | Money metric decomposition, node graph                        |
| Sectors                                                           | >      | Money metric decomposition, per capita                        |
| Trade                                                             | >      | Money metric decomposition, by product and factor             |
| No table                                                          | >      | Money metric decomposition, by product and factor, per capita |

The first table reports the contribution of the overall changes in the prices indices, in factor income and indirect tax revenues:

Money metric decomposition [0]

Scenarios: Tariffs  
 Sectors and institutions: mm  
 Years shock

|                                 | World   | Australia, New Zealand | East Asia | Southeast Asia | South Asia | North America | Latin America |
|---------------------------------|---------|------------------------|-----------|----------------|------------|---------------|---------------|
| Total [Bio \$]                  | 55.17   | 2.48                   | 54.39     | 4.56           | -0.81      | -11.63        |               |
| Overall price index [Bio \$]    | 644.96  | 3.80                   | -66.02    | 7.99           | 60.00      | 204.21        |               |
| Private price index [Bio \$]    | 477.05  | 2.16                   | -31.50    | 3.26           | 41.54      | 159.01        |               |
| Government price index [Bio \$] | 69.10   | -0.16                  | -29.90    | -0.92          | 2.54       | 31.65         |               |
| Savings price index [Bio \$]    | 95.40   | 1.79                   | -4.76     | 5.64           | 15.36      | 13.06         |               |
| Income total [Bio \$]           | -578.73 | -1.42                  | 120.19    | -3.66          | -58.97     | -213.38       |               |
| Factor income [Bio \$]          | 0.00    | 4.45                   | 118.65    | 15.88          | -4.95      | -44.63        |               |
| Depreciation [Bio \$]           | 81.90   | 1.17                   | -2.92     | 2.35           | 7.05       | 20.91         |               |
| Indirect tax income [Bio \$]    | -660.63 | -7.04                  | 4.46      | -21.89         | -61.06     | -189.66       |               |

Note that the table was pivoted to show a comparison across countries. The global factor income change impact is by definition zero in the GTAP standard model as (1) factor stocks are fixed, (2) the global factor price index is used as the numeraire.

The decomposition discussed above is found in the “by product and factor” table, where “Sum” captures the total as shown above. It repeats the impacts of changes in nominal income (factor income, indirect taxes etc.):

Money metric decomposition, by product and factor [0]

Regions: World  
 Sectors and institutions: Money metric [Bio \$]  
 Years shock

|                             | World   |
|-----------------------------|---------|
| All commodities             | 642.15  |
| Total factor demand         | 0.00    |
| Sum                         | 55.17   |
| Indirect taxes              | -660.63 |
| Depreciation                | 81.90   |
| Interaction effects         | -8.25   |
| All commodities             | 642.15  |
| Grains and Crops            | 21.37   |
| Livestock and Meat Products | 27.46   |
| Mining and Extraction       | 2.50    |
| Processed Food              | 76.85   |
| Textiles and Clothing       | 51.84   |
| Light Manufacturing         | 88.16   |
| Heavy Manufacturing         | 94.03   |
| Utilities and Construction  | 43.77   |
| Transport and Communication | 69.95   |
| Other Services              | 166.23  |
| Total factor demand         | 0.00    |
| Land                        | -9.67   |
| Capital                     | 11.49   |
| UnSkLab                     | -0.00   |
| SkLab                       | -0.00   |
| NatRes                      | -1.81   |

### 6.1.8 Decomposition following Huff and Hertel, 2000

The decomposition departs from the definition of regional income and allocates shares in total EV changes to changes in income contributions from changes in factor endowments at factor prices at the base year, from changes in quantities on which taxes are levied, from the trade balance etc.. The main differences to the approach above are that the change in EV is attributed in shares to components of income.

For endowment changes, we use the absolute changes in endowment xft times factor prices pft at the benchmark:

```
p_deltaEV(rNat,"xft",%arg2%) = sum(f,pft(rNat,f,"%t0%") * (xft(rNat,f,%arg2%)-xft(rNat,f,"%t0%")));
```

Similar, for depreciation in case of changes in the capital stock:

```
p_deltaEV(rNat,"depr",%arg2%) = - fdepr(rNat,"%t0%")*pi(rNat,"%t0%")*(kstock(rNat,%arg2%)-kstock(rNat,"%t0%"));
```

For the taxes, the GAMS code copies the equation defining the different tax revenues into assignments, where the quantity is replaced by the quantity change, e.g.:

Tax revenues from production tax, equation (13) in VDM 2018

```
[ sum(a $ prdtx.1(rNat,a,"%t0%"), prdtx(rNat,a,"%t0%")*(m_px(rNat,a,"%t0%")*(m_xp(rNat,a,"%t0%")-m_xp(rNat,a,"%t0%"))
```

For the terms of trade effects, changes in f.o.b. and c.i.f. prices at benchmark trade flows are taken into:

```
p_deltaEV(rNat,"t.o.t",%arg2%) = sum(i,rNat1, (%pefob(rNat,i,rNat1,%arg2%)-%pefob(rNat,i,rNat1,"%t0%")) *xw(rNat,i,rNat1,"%t0%"))
- sum(i,rNat1, (%pefob(rNat1,i,rNat,%arg2%)-%pefob(rNat1,i,rNat,"%t0%")) *xw(rNat1,i,rNat,"%t0%"))
```

Plus the impact of changes in the trade margins for exports:

```
+ sum(m,i,rNat1) $ amgm(m,rNat,i,rNat1),
    amgm(m,rNat,i,rNat1)
    * ( m_xws(rNat,i,rNat1,"%t0%") $ rrComb(rNat,rNat1)
        $onexpand
        + m_xws1(rNat,i,rNat1,"%t0%") $ (not rrComb(rNat,rNat1))
        $offexpand
    )
    * tmarg(rNat,i,rNat1,"%t0%")/m_lambdamg(m,rNat,i,rNat1,"%t0%")
--- margin prices in region r ???
    * (ptmg(m,%arg2%)-ptmg(m,"%t0%"))

)
```

And imports:

```
--- value of transport margins for imports of rNat times changes in margins price
- sum(m,i,rNat1) $ amgm(m,rNat1,i,rNat),
    amgm(m,rNat1,i,rNat)
    * ( m_xws(rNat1,i,rNat,"%t0%") $ rrComb(rNat1,rNat)
        $onexpand
        + m_xws1(rNat1,i,rNat,"%t0%") $ (not rrComb(rNat1,rNat))
        $offexpand
    )
    * tmarg(rNat1,i,rNat,"%t0%")/m_lambdamg(m,rNat1,i,rNat,"%t0%")
    * (ptmg(m,%arg2%)-ptmg(m,"%t0%"))

);
```

The final changes stems from updating prices of investments:

```
p_deltaEV(rNat,"fSav",%arg2%) = netInv(rnat,"%t0%") * (pi(rNat,%arg2%)-pi(rNat,"%t0%"))
- yi(rNat,"%t0%") * (psave(rNat,%arg2%)-psave(rNat,"%t0%"));
```

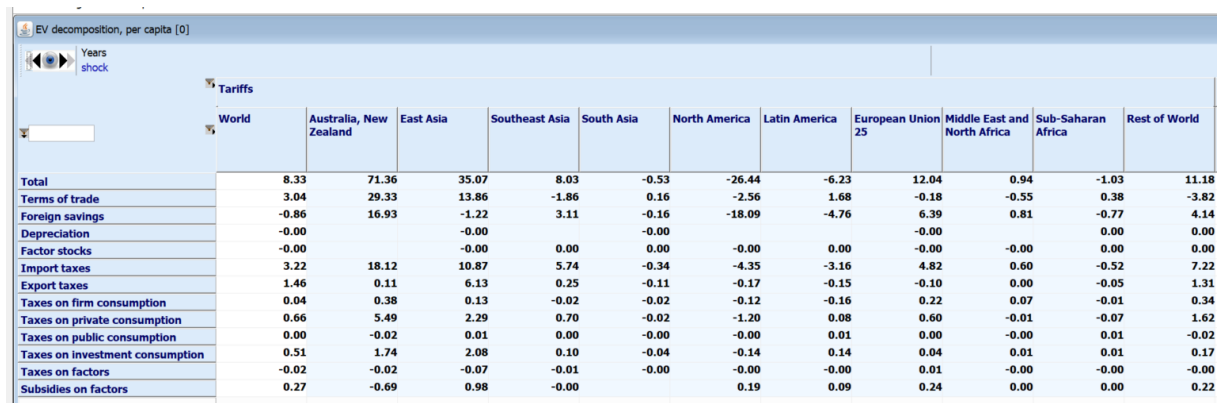
We assign the total EV in shares to these individual effects:

```
p_results(rnat,"v","ev",evEff,"tot",%arg1%, "%version%") $ p_deltaEV(rNat,"tot",%arg2%)
= p_results(rNat,"v","ra","mm","tot",%arg1%, "%version%")/p_deltaEV(rNat,"tot",%arg2%)*p_deltaEV(rNat, evEff,%arg2%);
```

And aggregate also to regional aggregates:

```
p_results(rAgg,"v","ev",evEff,"tot",%arg1%, "%version%")
= sum(ragg_r(ragg,rnat), p_results(rNat,"v","ev",evEff,"tot",%arg1%, "%version%"));
```

The results can be inspected in the inspected on a per capita basis:



|                                 | World | Australia, New Zealand | East Asia | Southeast Asia | South Asia | North America | Latin America | European Union 25 | Middle East and North Africa | Sub-Saharan Africa | Rest of World |
|---------------------------------|-------|------------------------|-----------|----------------|------------|---------------|---------------|-------------------|------------------------------|--------------------|---------------|
| Total                           | 8.33  | 71.36                  | 35.07     | 8.03           | -0.53      | -26.44        | -6.23         | 12.04             | 0.94                         | -1.03              | 11.18         |
| Terms of trade                  | 3.04  | 29.33                  | 13.86     | -1.86          | 0.16       | -2.56         | 1.68          | -0.18             | -0.55                        | 0.38               | -3.82         |
| Foreign savings                 | -0.86 | 16.93                  | -1.22     | 3.11           | -0.16      | -18.09        | -4.76         | 6.39              | 0.81                         | -0.77              | 4.14          |
| Depreciation                    | -0.00 |                        | -0.00     |                | -0.00      |               |               | -0.00             |                              | 0.00               | 0.00          |
| Factor stocks                   | -0.00 |                        | -0.00     | 0.00           | 0.00       | -0.00         | 0.00          | -0.00             | -0.00                        | 0.00               | 0.00          |
| Import taxes                    | 3.22  | 18.12                  | 10.87     | 5.74           | -0.34      | -4.35         | -3.16         | 4.82              | 0.60                         | -0.52              | 7.22          |
| Export taxes                    | 1.46  | 0.11                   | 6.13      | 0.25           | -0.11      | -0.17         | -0.15         | -0.10             | 0.00                         | -0.05              | 1.31          |
| Taxes on firm consumption       | 0.04  | 0.38                   | 0.13      | -0.02          | -0.02      | -0.12         | -0.16         | 0.22              | 0.07                         | -0.01              | 0.34          |
| Taxes on private consumption    | 0.66  | 5.49                   | 2.29      | 0.70           | -0.02      | -1.20         | 0.08          | 0.60              | -0.01                        | -0.07              | 1.62          |
| Taxes on public consumption     | 0.00  | -0.02                  | 0.01      | 0.00           | -0.00      | -0.00         | 0.01          | 0.00              | -0.00                        | 0.01               | -0.02         |
| Taxes on investment consumption | 0.51  | 1.74                   | 2.08      | 0.10           | -0.04      | -0.14         | 0.14          | 0.04              | 0.01                         | 0.01               | 0.17          |
| Taxes on factors                | -0.02 | -0.02                  | -0.07     | -0.01          | -0.00      | -0.00         | -0.00         | 0.01              | -0.00                        | -0.00              | -0.00         |
| Subsidies on factors            | 0.27  | -0.69                  | 0.98      | -0.00          |            | 0.19          | 0.09          | 0.24              | 0.00                         | 0.00               | 0.22          |

Or as totals:



|                                 | World |
|---------------------------------|-------|
| Total                           | 55.17 |
| Terms of trade                  | 20.14 |
| Foreign savings                 | -5.72 |
| Depreciation                    | -0.00 |
| Factor stocks                   | -0.00 |
| Import taxes                    | 21.34 |
| Export taxes                    | 9.67  |
| Taxes on firm consumption       | 0.29  |
| Taxes on private consumption    | 4.36  |
| Taxes on public consumption     | 0.02  |
| Taxes on investment consumption | 3.37  |
| Taxes on factors                | -0.11 |
| Subsidies on factors            | 1.80  |



## 6.2 Trade in VA indicators

### 6.2.1 Background

Trade has a double role by firstly allowing consumers to not only demand domestic produce, but also imports and by secondly by generating revenues and thus income by exports. The first role can be easily assessed in a CGE framework by reporting import shares on demand. The second one is harder to assess as export revenues comprise both direct value-added from the exporting sector and indirect value-added comprised in domestic intermediates. In parallel, the use of import intermediates in production implies that value-added is also generated in other than the exporting country. In the following, we present the implementation of a well-established framework drawing on Leontief Multiplier analysis into the post-processing of CGEBox to assess the income generate role of exports.

We start with an illustrative example of an open economy with 2 two sectors and no foreign savings. Imports serve in our example only as intermediates inputs are not detailed by commodity:

|                            | Agr | NonAgr |
|----------------------------|-----|--------|
| Agr share                  | 0,1 | 0,05   |
| NonAgr share               | 0,4 | 0,5    |
| Import share               | 0,2 | 0,1    |
| VA share                   | 0,3 | 0,35   |
| Output                     | 30  | 80     |
| VA                         | 9   | 28     |
| Final demand incl. Exports | 23  | 28     |
| Exports                    | 2   | 12     |
| Final domestic demand      | 21  | 16     |

The example is balanced as the export revenues of 14 (=2+12) are equal to the value of imports ( $0,2 \cdot 30 + 0,1 \cdot 80$ )=12. The question is about the contribution of the exports (or domestic sales) to value added. Based on usual Leontief approach, we first calculate  $(I-A)$ , where  $I$  is the identity matrix and  $A$  the matrix of domestic input coefficients:

In our example that yields the following 2x2 table:

|      |       |
|------|-------|
| 0,9  | -0,05 |
| -0,4 | 0,5   |

Inverting that matrix delivers the Leontief multipliers  $(I-A)^{-1}$ :

|            |            |
|------------|------------|
| 1,1627907  | 0,11627907 |
| 0,93023256 | 2,09302326 |



Multiplying the value-added shares of the sectors with the Leontief multipliers delivers the VA multipliers for each commodity:

0,6744186 0,76744186

If we now multiply these with final demand and export, we derive the VA contribution of the demand contributions:

|                                    |            |
|------------------------------------|------------|
| Final domestic demand contribution | 26,4418605 |
| Export contribution                | 10,5581395 |
| <b>Total Value added</b>           | <b>37</b>  |

Which are equal to the value added as reported in our simple SAM and, as the SAM is balanced also equal to final domestic demand. The aim is now to implement the above steps into the CGEBOX framework.

### 6.2.2 Post-Model processing

As a first step, we define (I-A) as follows, where the *p\_results* parameter was populated before from the model results. Note that we are using the “V” field which captures value (= quantity x price):

```
--- invert (I-A)
execute_unload '%scrdir%\p_IA.gdx',i,p_IA;
execute '=invert.exe %scrdir%\p_IA.gdx i p_IA %scrdir%\p_IAInv.gdx p_IAInv';
execute_load '%scrdir%\p_IAInv.gdx',p_IAInv;
```

Next, we use the “INVERT” utility from GAMS to define (I-A)<sup>-1</sup>:

```
--- invert (I-A)
execute_unload '%scrdir%\p_io.gdx',i,p_IA;
execute '=invert.exe %scrdir%\p_IA.gdx i p_IA %scrdir%\p_IAInv.gdx p_IAInv';
execute_load '%scrdir%\p_IAInv.gdx',p_IAInv;
```

From there, we calculate multipliers for each primary factor – commodity combination:

```
--- define multipliers for each commodity j and each primary factor f
p_results(rs,"Q",j,f,"tradeInVa",%1,"%version%")
= sum( (a,i) $ xFlag(rs,a,i),
      p_results(rs,"U",f,a,"dom",%1,"%version%")/p_results(rs,"U","out",a,"dom",%1,"%version%") * p_IAInv(i,j));
```

Which can be aggregated to yield value added multipliers

```
--- aggregate to VA multipliers
p_results(rs,"Q",j,"va","tradeInVa",%1,"%version%") = sum(f, p_results(rs,"Q",j,f,"tradeInVa",%1,"%version%"));
```

Using these multipliers and export volumes allows to calculate the VA contribution of exports:

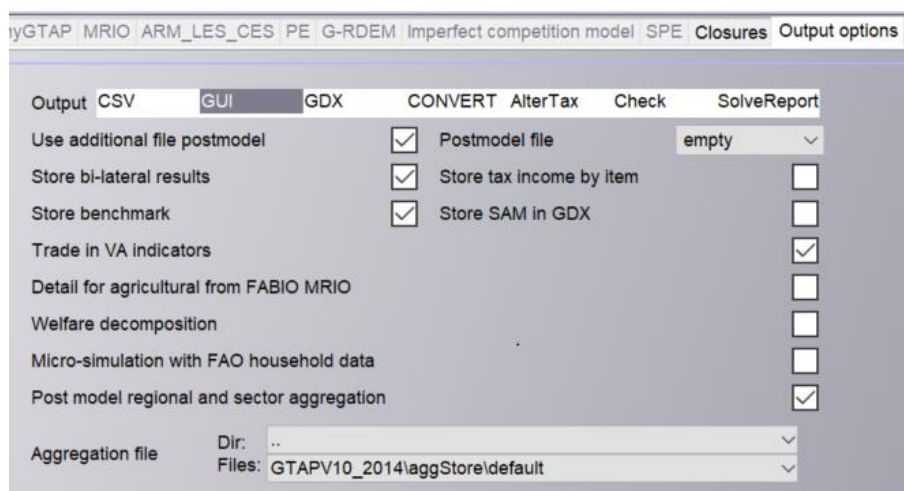
```
*
* --- calculate income by primary factor embodied in trade
*
p_results(rs,"U",j,f,"tradeInVa",%1,"%version%")
= p_results(rs,"Q",j,f,"tradeInVa",%1,"%version%") * p_results(rs,"Q",j,"exp","dom",%1,"%version%");
*
* --- calculate value added embodied in trade
*
p_results(rs,"U",j,"va","tradeInVa",%1,"%version%") = sum(f, p_results(rs,"U",j,f,"tradeInVa",%1,"%version%"));
```

Further code lines not shown here aggregate the information over regions and commodities.

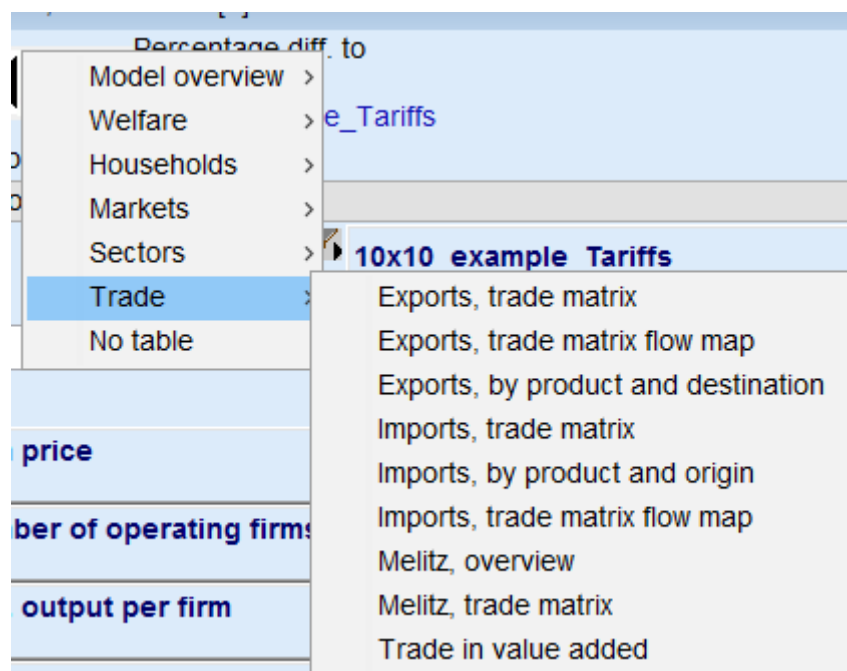
The TIVA approach also defines output multipliers. This allows assigning CO2 equivalent discharges (taxed and emitted) allocated to production activities (linked to intermediate or factor use, process emissions) in delivering countries to final demand in use countries. The resulting embodied emissions per constant unit dollar are stored in temporary file “*embed.gdx*” in the scratch directory and can be used, for instance, for simulation of a Carbon Border Adjustment mechanism.

### 6.2.3 GUI

In order to generate the information, the “*GUI*” output option must be switched on as well as the “*Trade in VA indicators*” checkbox:



That combination produces a table under the theme “trade”:



Which reports both the multipliers and the contribution of trade to value added. Additionally, the share from export related value added on total income is reported:

Trade in value added [0]

Percentage diff. to  
Scenarios  
10x10\_example\_Tariffs

Exporter: World Sectors and institutions: Total Years: shock

|                             | Per unit of output | In total trade | In trade, relative to GDP |
|-----------------------------|--------------------|----------------|---------------------------|
| Total                       | 0,7508             | 9889,7246      | 0,1986                    |
| Total                       | 0,7508             | 9889,7246      | 0,1986                    |
| Grains and Crops            | 0,9726             | 257,4330       | 0,0052                    |
| Livestock and Meat Products | 0,8502             | 122,8059       | 0,0025                    |
| Mining and Extraction       | 0,8705             | 1186,6638      | 0,0238                    |
| Processed Food              | 0,7638             | 425,7714       | 0,0085                    |
| Textiles and Clothing       | 0,6706             | 400,4073       | 0,0080                    |
| Light Manufacturing         | 0,6488             | 1697,5337      | 0,0341                    |
| Heavy Manufacturing         | 0,6005             | 3742,0999      | 0,0751                    |
| Utilities and Construction  | 0,7302             | 113,5394       | 0,0023                    |
| Transport and Communication | 0,7690             | 747,9837       | 0,0150                    |
| Other Services              | 0,8304             | 1195,4865      | 0,0240                    |

As seen below, the table also provides detail by factor:

Trade in value added [0]

Percentage diff. to  
Scenarios  
10x10\_example\_Tariffs

Exporter: World Sectors and institutions: Land Years: shock

|                  | Per unit of output |  |          |
|------------------|--------------------|--|----------|
| Total            |                    |  | 118,6026 |
| Total            | 0,0092             |  | 118,6026 |
| Grains and Crops | 0,2257             |  | 53,2937  |

## 6.3 Altertax

### 6.3.1 Background

There are various use cases where it makes sense to construct a new benchmark from the GTAP data base before shocking the model. One option to construct a new benchmark first changes exogenously entries in the global SAM and next re-balances it. That approach is especially suitable if benchmark changes are motivated by new or updated data. The filtering approach implemented in CGEBox can be seen as a specific variant of that approach.

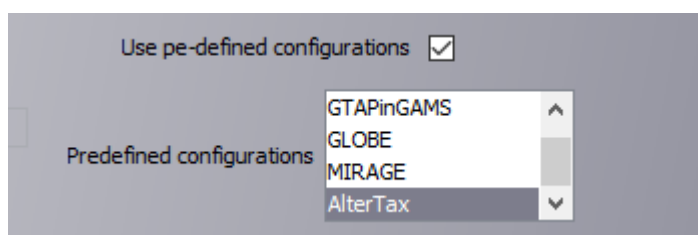
The option discussed in here uses the results of the simulation model itself. There are two possible motivations for it. First, this approach is especially appropriate for changes where the behavioral responses embodied in the equation structure provide a better way to re-adjust the benchmark compared to the more mechanical re-balancing the approaches. Second, one might want to conduct comparative-static analysis to the outcome of baseline process with the G-RDEM model.

Conceptually, re-using the simulation outcome as a benchmark data base set is straightforward as the accounting identities underlying a balanced SAM must also be fulfilled by a simulation with the CGE itself. One main challenge is to map the simulated results back in same format used on the input side of the simulation model. As CGEBox is a SAM based CGE approach, that step is not overly complex. Second, additional data sets, for instance, relating to land use or emissions, might not have included in the simulation process but shall still be made available for subsequent benchmarking. The brief following documentation will discuss the basic methodology and technical implementation

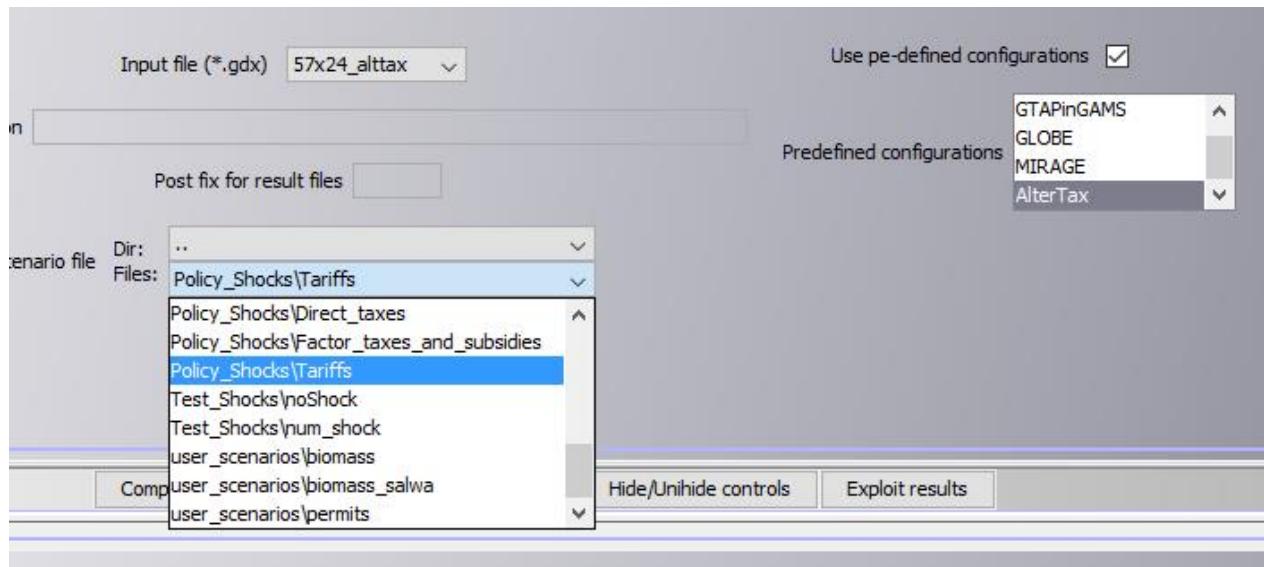
### 6.3.2 Technical realization

The technical realization re-uses existing elements of CGEBox to implement alterTax:

1. A predefined configuration which can be chosen from the interface set-ups a simulation model useful for alterTax, i.e. the parameterization of the model is changed such that most substitution elasticities are unity (or close to unity).



The user can select any shock file:



6. As part of post-processing, the relevant entries for a start of the model define a new benchmark, including the global SAM.

Note: The alterTax output can be also generated from any normal model run, in case the pre-configuration given next is not suitable.

### 6.3.3 Pre-configuration

The pre-configuration changes the following elements compared to the standard GTAP Model:

- The demand system for private households is changed to a CD.
- The production nesting using substitution elasticities of unity between VA and the intermediate composite, between primary factors and between intermediate.
- The Armington elasticities are set close to unity (a CD implementation is currently not supported).
- All factors are mobile

The use of the CD functional form allows to keep the value shares constant during the alterTax simulations which thus maintains basic relations between transactions in the SAM. If highly detailed global SAMs are subject to AlterTax, the per-solve mechanism can be used as well.

### 6.3.4 Post-Model processing

The SAM output option is switched on and the results for the shock are stored again as a SAM in a new data file. The same holds for various sub-matrices, for instance:

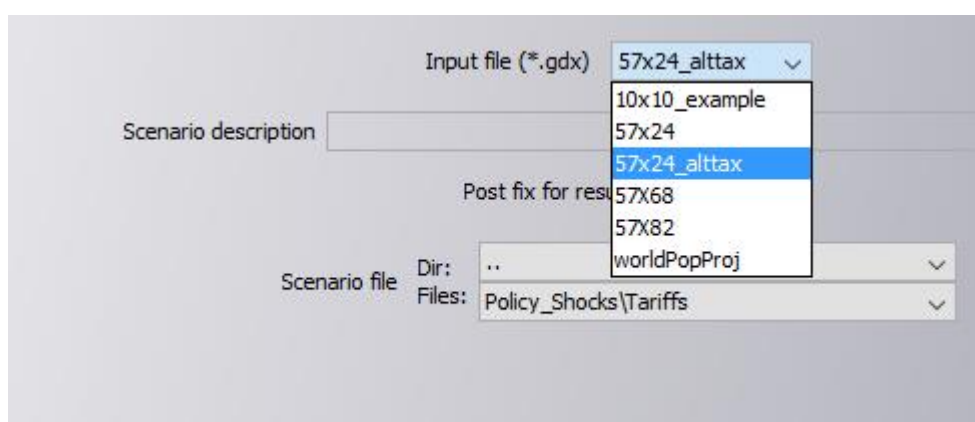
```
xda0(r,i,aa)    $ xda0(r,i,aa)    = xd.l(r,i,aa,"shock") * pdp.l(r,i,aa,"shock") / gblscale;
```

The update symbols are stored to a new data file:

```
execute_unload "%datdir%\%dataset%\%postfix%\alttax.gdx", is, aa, a, i, r, f, l, cap, lnd, nr, fd, h, gov, inv,
sam0, xda0, xdn0, xna0, xmm0, xmarg0,
eh0, bh0, sigmap0, sigmav0, sigmai0, sigman0, sigmaw0, omegaf0, RoRflex0,
kstock0, depry0, pop0,
emid0, emii0,
exptxY0, imptxY0, fcttxY0, fcttsY0
mapi=ii, mapr=rr, m, h, mapa=aa1, ff,
$$ifthen defined rAez
    rAez, rAez_r_aez, landUse, AEZ, land_cover
$$endif
*$$ifi "%filterMethod%"=="Rebalancing" itrlog, keepCor.1
*$$ifi "%loadLandUse%"=="on" landUse, AEZ, land_cover
    $$ifthen defined s_meta
        s_meta
    $$endif
;
```

### 6.3.5 Use of the updated benchmark

The updated benchmark data set is stored in the data directory with the post-fix “alttax” appended and can be used as any other data set:



## **6.4 Using FAOSTAT data for splits**

- Wolfgang Britz, Martin Bruckner, Salwa Haddad –

### **6.4.1 Introduction**

FAOSTATs offers a wealth of data on global agriculture including herd sizes and harvested area for UN Member countries and many agri-food products. The so-called Food-Balance Sheets report also on processed food products in rich detail.

## 7. GUI and installation

### 7.1 Software requirements and access to the code

GAMS version 24.5 (released 2015) or higher is necessary to run the GAMS code, independently from using the code with or without the GUI. For GAMS versions before 24.8 (released 2016), it is not recommended to use the grid solve for the pre-solves or during data preparation. CONOPT4 is only available with newer GAMS version, if it is found, CONOPT should be used instead. The GUI requires a Java 8 run-time installation on the computer, for more detail see section 7.4 below, pages 428 ff.

The code and GUI are distributed in two ways. First, a zip archive can be downloaded from the GTAP web site (<https://www.dropbox.com/s/olcqbqprk6i9qth/CGEBOX.zip?dl=0>). The GAMS code reflects the status of the project in December 2018. In order to benefit from code improvements and bug-fixes, it is recommend to use the SVN option, see section 7.4 below, pages 428 ff.

### 7.2 Some comments for runGTAP and GEMPACK users

CGEBox is not thought or designed as a tool to easily switch from runGTAP or a model written in GEMPACK to a GAMS based CGE modelling platform. Even if CGEBox can replicate the structure of the standard GTAP model – however with equations written in levels and not in log-linearized form – the GAMS code uses different mnemonics and a different coding style, following Van der Mensbrugghe 2018<sup>124</sup>. That is not only due to differences between GAMS and GEMPACK, but also reflects a different style how to code a complex model and to allow for modularity. Please refer to the equation by equation documentation of the core of CGEBox in the next section “Basic model equations”. Equally, the GUI does not attempt to mimic the touch and feel of runGTAP or other GEMPACK utilities.

Users which so far have worked with runGTAP and GEMPACK can clearly expect to benefit from their knowledge about CGE modeling and the concept of GTAP, and their experience of having worked with a modelling language. With CGEBox, they will face however a new modeling platform which will most probably lead also to disappointments if things work differently than presumed, if usability is for some aspects assessed as lower and time must be spent with the documentation and code to get things going. It might be best to expect that things work differently and to have a look at the documentation first, instead of simply starting to work with the system along the lines used from runGTAP.

Note specifically that upper and lower case letters in symbol names have no meaning in the sense of indicating variables in levels or in percentage changes. Equally, note that GAMS makes a distinction between parameters which are always constants in a model and can never be changed by the solver and variables. While variables in GAMS can be technically turned into constants during a model solve by fixing them, parameters can never be turned in endogenous variables. This implies that symbols defined as parameters can never be subject to a closure swap.

Sparsity might warrant some further comments. In most GTAP data bases, all transactions are non-zero, but many are often very tiny. CGEBox offers a filter approach to remove such tiny transactions from the global SAM. Doing so reduces data base and model size, thus speeding up model generation

---

<sup>124</sup> Van der Mensbrugghe, D. (2018). The Standard GTAP Model in GAMS, Version 7. *Journal of Global Economic Analysis*, 3(1), 1-83



and solution as well as memory needs. However, allowing for sparsity in the data base requires that in most cases zero transactions are excluded in the GAMS code from data transformation and from entering the model. Otherwise, math trap errors such as divisions by zero are easily provoked. GAMS throws run time errors if that occurs and will as a default not solve any model after such an error occurs. That must be especially reflected when the user develops its own shock files.

### 7.3 Background

Working with a larger existing model such as GTAP can provoke serious entry barriers for newcomers, partly due to the need to learn specific model mnemonics underlying the code. That knowledge is necessary to analyze results if only model output directly produced by the modeling language is available, e.g. listings or proprietary binary result files. Equally, the sheer number of symbols found in larger models (parameters, variables, equations, labels for regions, sectors, factors etc.) and the dimensionality of results can overwhelm beginners. This might shy them away from a fascinating research field, even if excellent model documentation is available, such as in the case of GTAP. Memorizing model mnemonics and partly model structure is rather independent from the modeling language used –beginners will always have to digest a larger amount of information to make the first steps with a larger economic simulation model.

It might therefore be useful to reduce entry barriers such that students might become enthusiastic about the model's capabilities and then take the more tedious steps to familiarize themselves with the model's code. Therefore, CGEBox comes with a Graphical User Interface (GUI) which complements the GAMS code. The reader should however note that the GUI linked to CGEBox follows a somewhat different approach compared to runGTAP, the GUI available for the GEMPACK version (see also Britz 2014<sup>125</sup> and Britz et al. 2015<sup>126</sup>) or, for instance, the GAMSIDE often used with economic models realized in GAMS. As a consequence, some functionalities found in runGTAP are not comprised in the GUI linked to the CGEBox and vice versa. The development of the GUI for CGEBox and its GAMS code for post-model processing feeding into the exploitation part of the GUI went partly in parallel with the development of the model codes to ensure seamless inter-operability. However, it is important to note that the GAMS code of CGEBox can be used completely independent from the GUI.

The GUI provides two major functionalities: firstly, it allows running the model while also liberating a user to carry out changes that are otherwise necessary in the GAMS code—such as to select a specific data base or shock file. Instead, controls on the GUI allow steering the model run, thereby avoiding code edits. That include configuration of the model by selecting modules and parameters files. Secondly, the GUI offers tools to view and analyze model results. Specifically, a post-model reporting part maps back the variables into a structure similar to a SAM, and additionally stores quantities, prices, values, tax rates and tax incomes. That structure can then be viewed in a flexible exploitation tool.

The post model code also provides some eventually useful aggregations, e.g. to total intermediate input, total factor input, total output, total demand and world totals and provides a welfare

---

<sup>125</sup> Britz, W. (2014): A New Graphical User Interface Generator for Economic Models and its Comparison to Existing Approaches, *German Journal for Agricultural Economics*, 63(4)

<sup>126</sup> Britz, W., Domínguez, I. P. and Gopalakrishnan, B. N. (2015): Analyzing results from agricultural large-scale Economic Simulation Model: Recent Progress and the Way Ahead, *German Journal for Agricultural Economics*, 107-119

decomposition. The user can also provide based on an aggregation definition file generated with GTAPAgg a second, more aggregated sectoral and/or regional aggregation which is used for post-model aggregation of these results.

The exploitation tools available with the GUI provide interesting features: they allows to view simultaneously results from different shocks without running additional programs, support maps and a wider range of graphics as well as basic statistics and thus might hence also be of interest to seasoned modelers. Furthermore, the GUI comprises some additional utilities: a tool similar to “EXAMINER” in GTAP which shows the linearized version of the model’s equations and allows inspecting parameters and variables found in these equations. Similar to “AnalyzeGE”, that tool also allows decomposing changes of the LHS of each equation to changes of the variables on the RHS. Another utility allows documenting the code in HTML pages. Finally, instead to using the GUI in interactive model, a batch mode allows to perform model runs. The batch reports these run in a HTML page and allows parallel execution of different shocks or of the same shock in different model variants.

This section provides a hand-on guide on how to use the model in conjunction with the GUI. The latter is based on a GUI generator for GAMS projects ([http://www.ilr.uni-bonn.de/agpo/staff/britz/ggig\\_e.htm](http://www.ilr.uni-bonn.de/agpo/staff/britz/ggig_e.htm), Britz 2010<sup>127</sup>), which is also used for some larger projects such as CAPRI (<http://www.capri-model.org/dokuwiki/doku.php?id=start>), but also a range of other modeling system such as the Policy Evaluation Model (PEM) by the OECD; a dynamic single farm model, an Agent Based Model of structural change in agriculture or a recursive-dynamic Hydro-economic river basin model. For those interested in a discussion about interfaces used with economic simulation models, see for example <https://www.gtap.agecon.purdue.edu/resources/download/5864.pdf> (Pérez Domínguez et al. 2012) or Britz et al. 2015. The GUI generator is flexible enough to also host extensions as shown by the implementations of GTAP-AEZ, GTAP-AGR and GTAP-E.

A first version of the GUI was developed for a class at Bonn University, Institute for Food and Resource Economics, in the winter term 2013/2014, building on Tom Rutherford’s GTAPinGAMS version. The students had already some knowledge of GAMS and CGEs, and wanted to use a global CGE in a project. A year later, again a class of students wanted to use a global CGE, which led to further improvements such as flexible aggregation of regions in the mapping viewer or the introduction of sensitivity analysis. In 2014, when the GTAP Center decided to release a full-fledged version of the standard GTAP model in GAMS, it was decided to adjust the interface to operate with that version. The further work on GTAP led to improvements to the GUI generator, also to the benefit of other modeling systems applying the GUI. The GUI itself is distributed for free. However, the underlying Java code is currently not open source.

## 7.4 SVN checkout

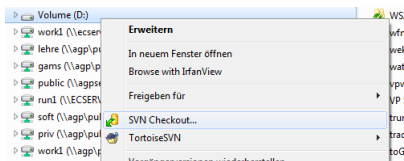
The preferred option to access code and GUI is to benefit from SVN, a software version system on which the model codes are hosted on the repository <https://svn1.agp.uni-bonn.de/svn/cgebox/>. The SVN option is especially recommended for users who both want to use the GUI and foresee also own changes to the code. In the SVN repository, the development of CGEBox continues. We expect that over the next years additional modules will be integrated into the code and clearly, bugs corrected and improvements implemented once the code is more widespread used. Using SVN allows updating more easily to bug fixed or new extensions.

---

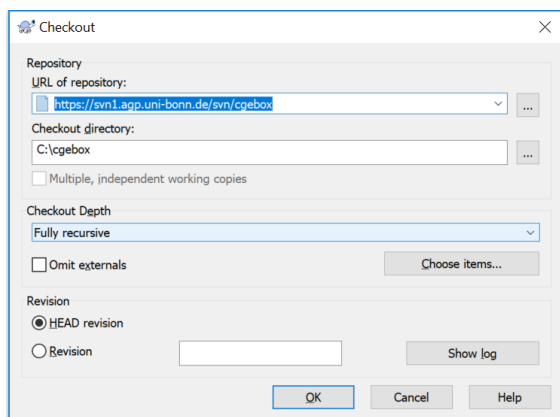
<sup>127</sup> Britz, W. (2010): GAMS Graphical Interface Generator Programming guide. Institute for Food and Resource Economics, University Bonn. In: [http://www.ilr.uni-bonn.de/agpo/staff/britz/GGIG\\_programming\\_guide.pdf](http://www.ilr.uni-bonn.de/agpo/staff/britz/GGIG_programming_guide.pdf)

Therefore, as a first step, we recommend installing a SVN client such as TortoiseSVN (<http://tortoisesvn.net/index.de.html>, freeware) on your machine. Afterwards, make a checkout of the repository <https://svn1.agp.uni-bonn.de/svn/cgebox/> (i.e. download the code from the server to your local machine), using the user id “cgebox” and the pass word “cgebox”. Using SVN will allow you to keep your local copy automatically updated to code improvements, to check which changes you introduced in the code and to switch easily back to original versions of all or specific files if deemed useful, if necessary on a line by line basis.

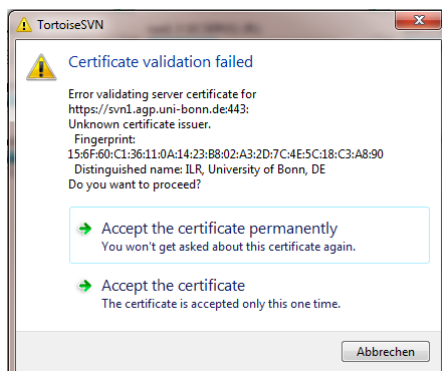
The basic steps for using SVN to install CGEBox based on TortoiseSVN are detailed below. First, navigate in the windows explorer to the disk / directory where you want to install GTAP, and open the context menu with a right mouse click:



Chose, “SVN Checkout ...” and copy the URL “<https://svn1.agp.uni-bonn.de/svn/cgebox/>” in the first field as shown below and press “OK”.

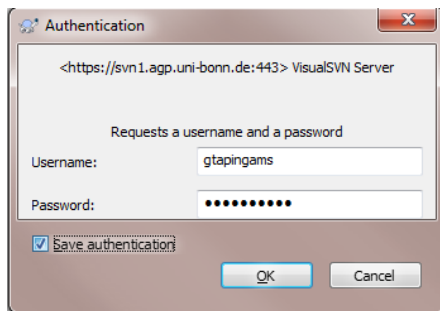


In case you get a warning such as:

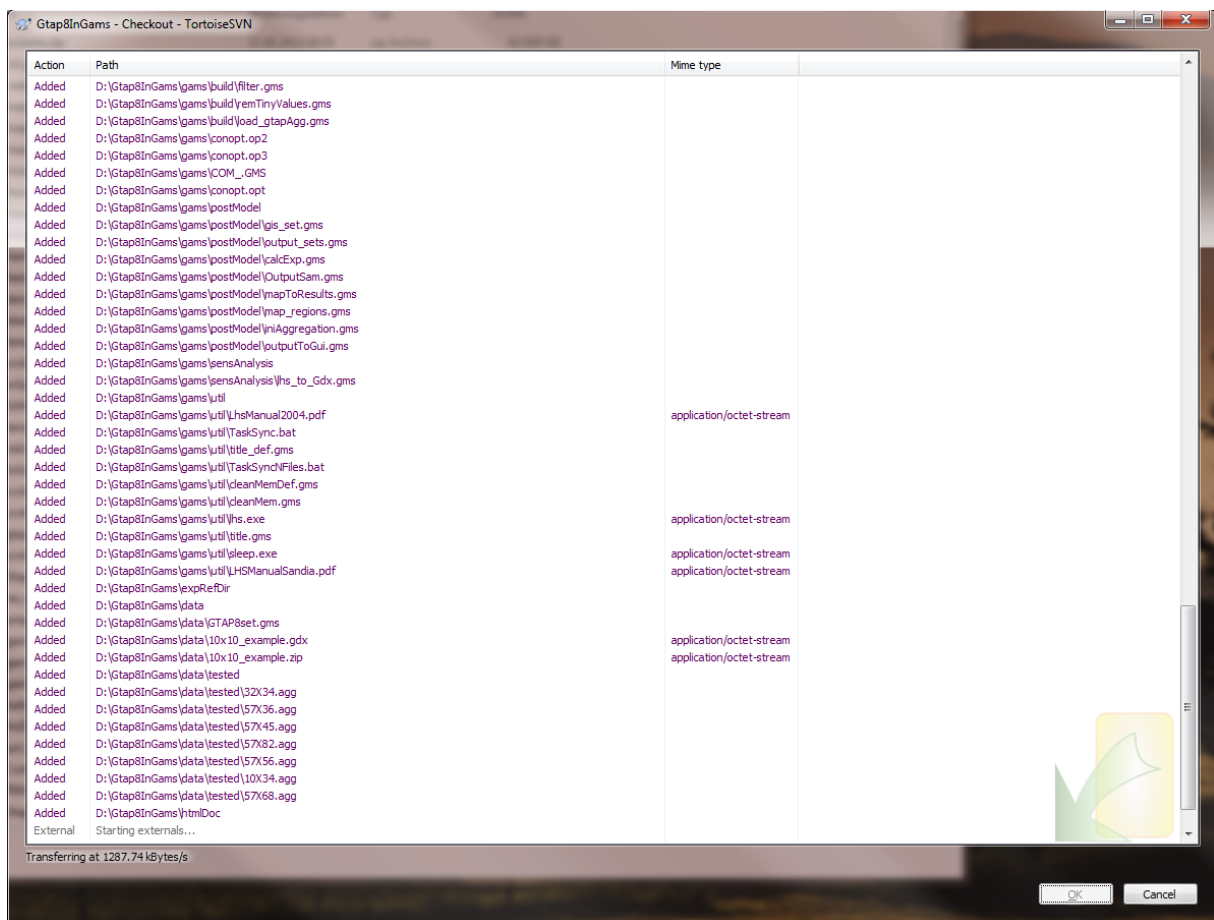


choose “Accept the certificate permanently”. In the next dialogue, enter the username and password “cgebox” ( not the one shown below), check “Save authentication” in the checkbox in the lower part of the dialog and press “OK”.

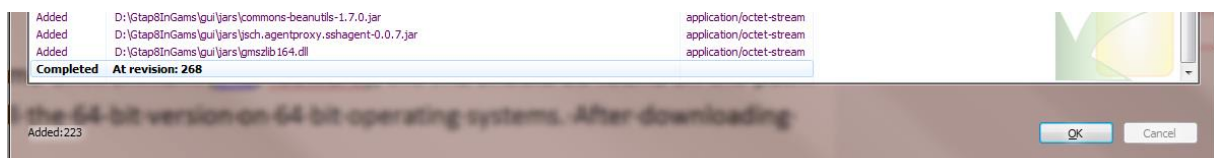
## SVN checkout



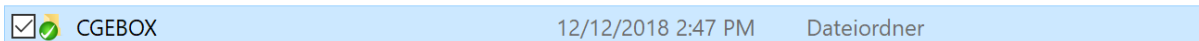
The SVN client will next download the newest release of all files to your computer which might take some time:



And should end with a “Completed At revision ..” message.



After you press “OK”, you should find the new folder on your computer:

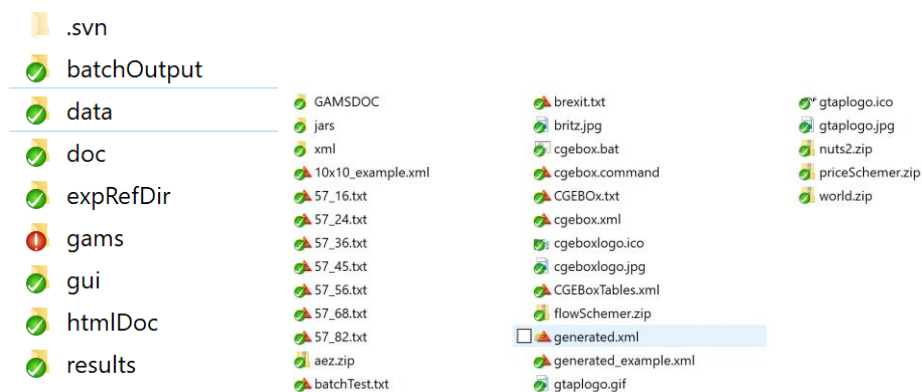


The green mark indicates that there are no local modifications to files found in the folder, i.e. that the files on your disk are identical to those just downloaded from the repository. If you introduce changes in any file, a different symbol will be shown.

For further information on how to benefit from TortoiseSVN for your daily work with CGEBox for other projects, please refer to its documentation available on its web-site. We would like to mention here only that you have always two versions of each file on your computer: (1) the so-called local working copy – these are the files in your normal directory with which you can work as usual and (2) a second version in a hidden data base which reflects the latest download from the repository. That allows the system to e.g. find out which files changed and to highlight changes line by line. It is important to note that you will never lose edits on your local computing when you perform “updates”, i.e. download newer versions from the central repository.

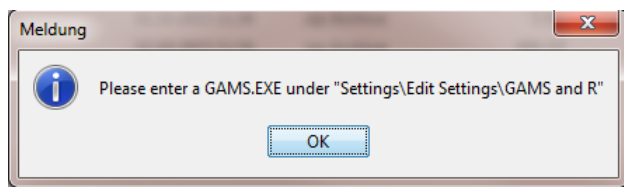
## 7.5 Starting the GUI

The GUI requires a Java 8 run time environment ([JRE](#), feeware); the JRE should be found on the path. Please make sure that you install the 64-bit version of the JRE on 64 bit operating systems. After downloading the GAMS sources and other files via SVN or having unzipped the downloaded folder, navigate to the GUI folder:



and double-click in the GUI folder on the file “cgebox.bat”; it should open the interface<sup>128</sup>. You might want to put a shortcut to that batch file on your desktop to easily startCGEBox.

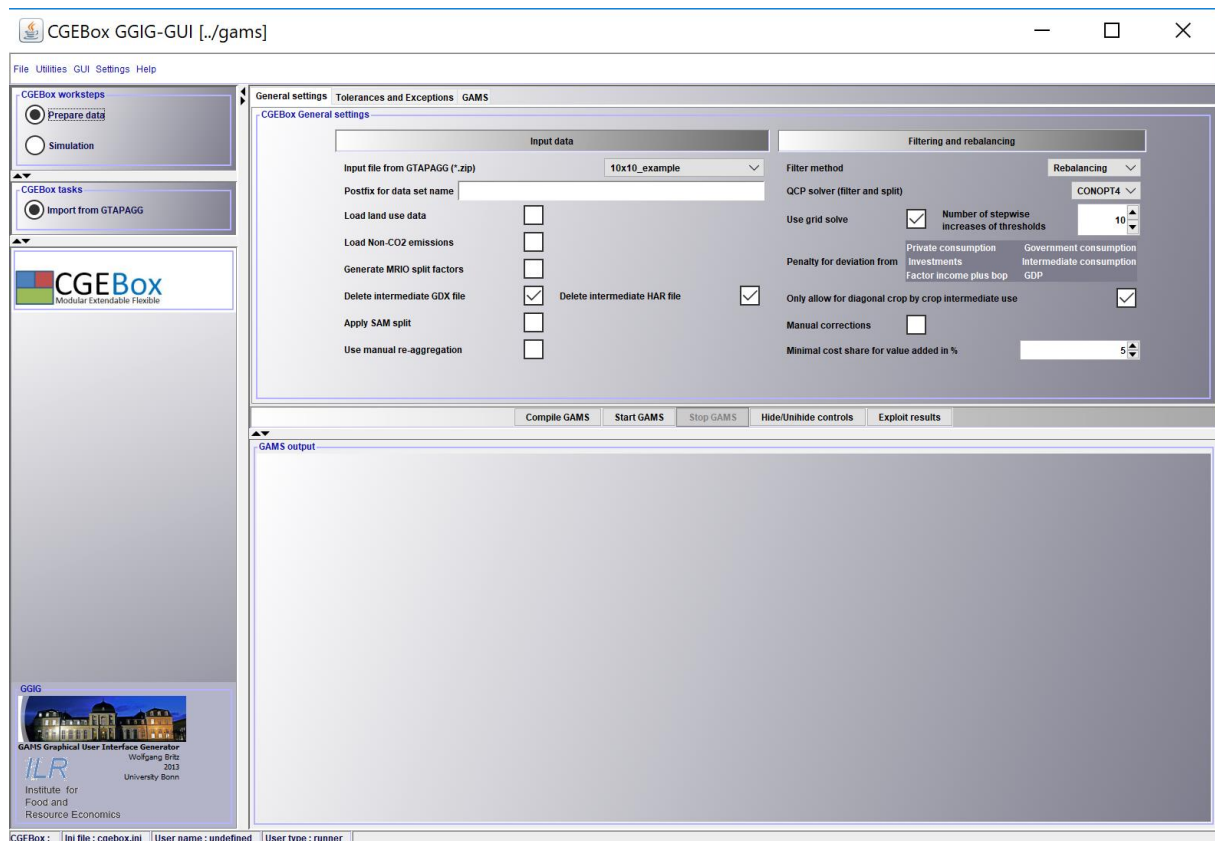
You might first see a message which tells you that a GAMS executable is not yet registered:



You can first ignore that warning for now. After pressing “OK”, you should see a program opened as below:

---

<sup>128</sup> If that does not work, open a command prompt, start from there the “cgebox.bat” batch file and analyze the error. Most probably, JAVA is not found. In that case, either put it on the path or change the batch file such that it calls Java from the directory where it is installed.



**Figure: Interface at first start**

**Note:** The layout shown in screen shots might not be completely identical to the distribution version

The left hand side allows selecting work steps and the tasks comprised in them; the right hand side carries controls to steer these tasks. The bar in the middle comprises buttons to start the GAMS code of a task and exploit the resulting output.

Typical reasons why the **GUI does not open** and possible remedies are described briefly below:

- **Java is not installed** or not registered on the PATH. In order to test that, open a DOS prompt and type Java. If you receive a message that the command is not found, (re-)install Java and make sure that Java also updates the PATH during the installation.
- A **wrong Java Runtime Environment** is installed. Please first check if Java version 8 is installed and not some older version (that check is not only useful for the GUI of CGEBox, but also generally for security reasons if one uses Java plugins in web pages). On 64 bit Windows operation systems, a 64 bit version of Java should be used. You can test that again in a DOS prompt by typing “java -version”. The response should look similar as shown below (note the string “64-Bit”). Unfortunately, many 64 bit Windows machines use a 32 bit browser, and if you install Java from such as browser, the 32 bit version is downloaded. The link [JRE](#) points both to a 32 and 64 bit version, make sure you select the 64 bit version on a computer with a 64 bit OS. Installing a 64 bit version on a machine where the 32 bit version is already installed does no harm and makes the 64 bit version the default.

```
java version "1.8.0_25"
Java(TM) SE Runtime Environment (build 1.8.0_25-b18)
Java HotSpot(TM) 64-Bit Server VM (build 25.25-b02, mixed mode)
```

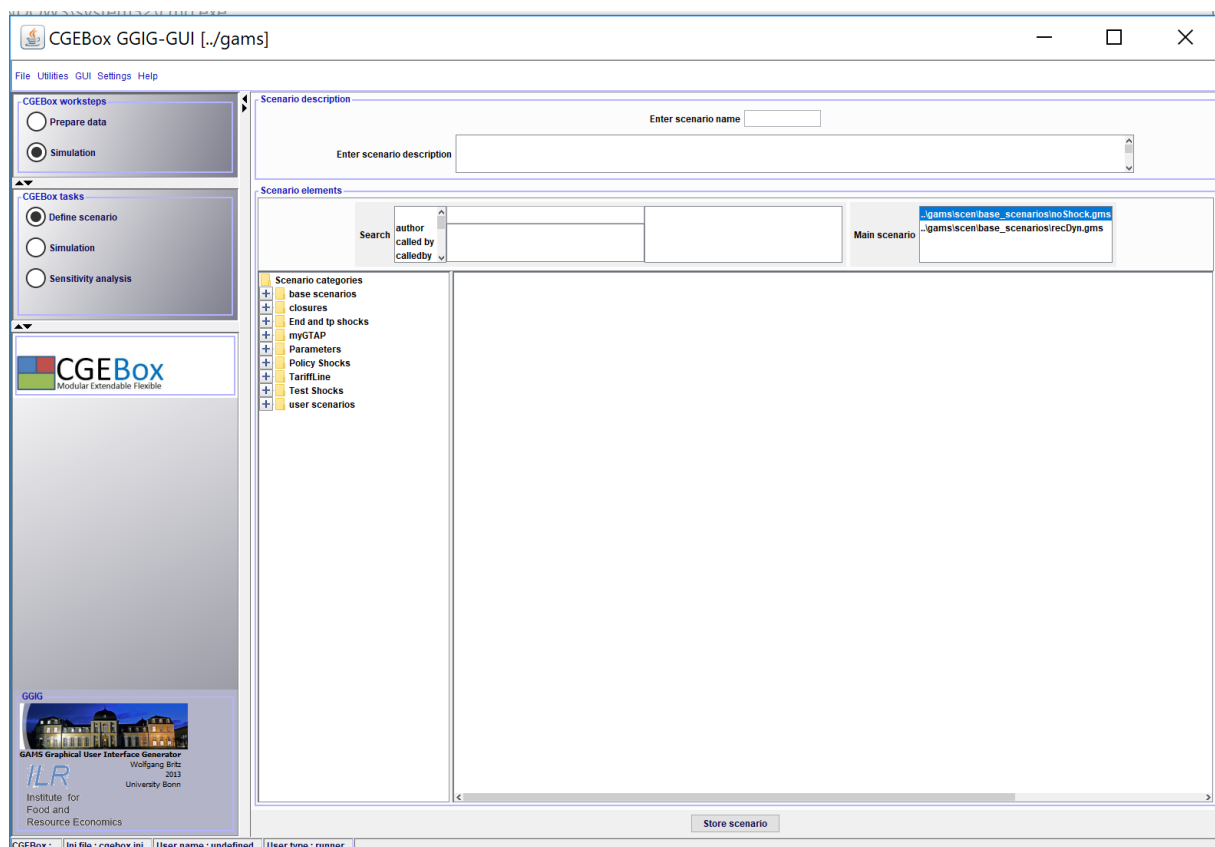
- If that still does not help, open a DOS prompt, navigate to the GUI folder and type “cgebox.bat” – possible errors will be shown in the window. In case you cannot find help in your team with regard to these error messages, contact the author.

General information on how to work with the GUI can be found in the GGIG user guide at [http://www.ilr.uni-bonn.de/agpo/staff/britz/GGIG\\_user\\_Guide.pdf](http://www.ilr.uni-bonn.de/agpo/staff/britz/GGIG_user_Guide.pdf). That pdf document can also found in the folder “GUIjars” in your installation. Please note that the hints given below in the document on using the interface are only thought as a first introduction, refer to the GGIG user guide for any further details or use the inbuilt help system by pressing F1.

We would finally like to note that we did our best to test the combination of the CGEBox code and the GUI, but that improvements are certainly possible.

## 7.6 First steps

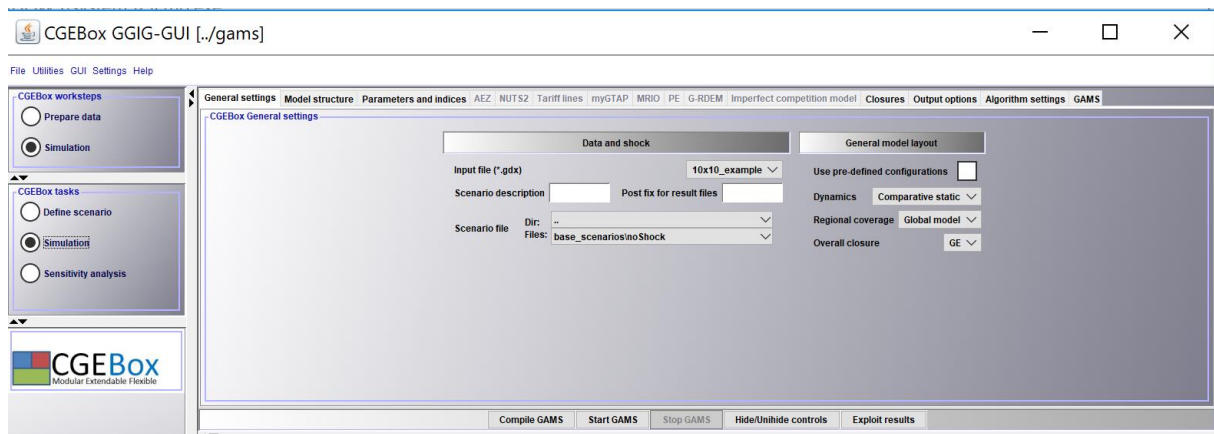
The installation comprises a set of results for testing the interface without the need of actually running GAMS code. You can now already look at these results by selecting on the left hand side the work step “Simulation”:



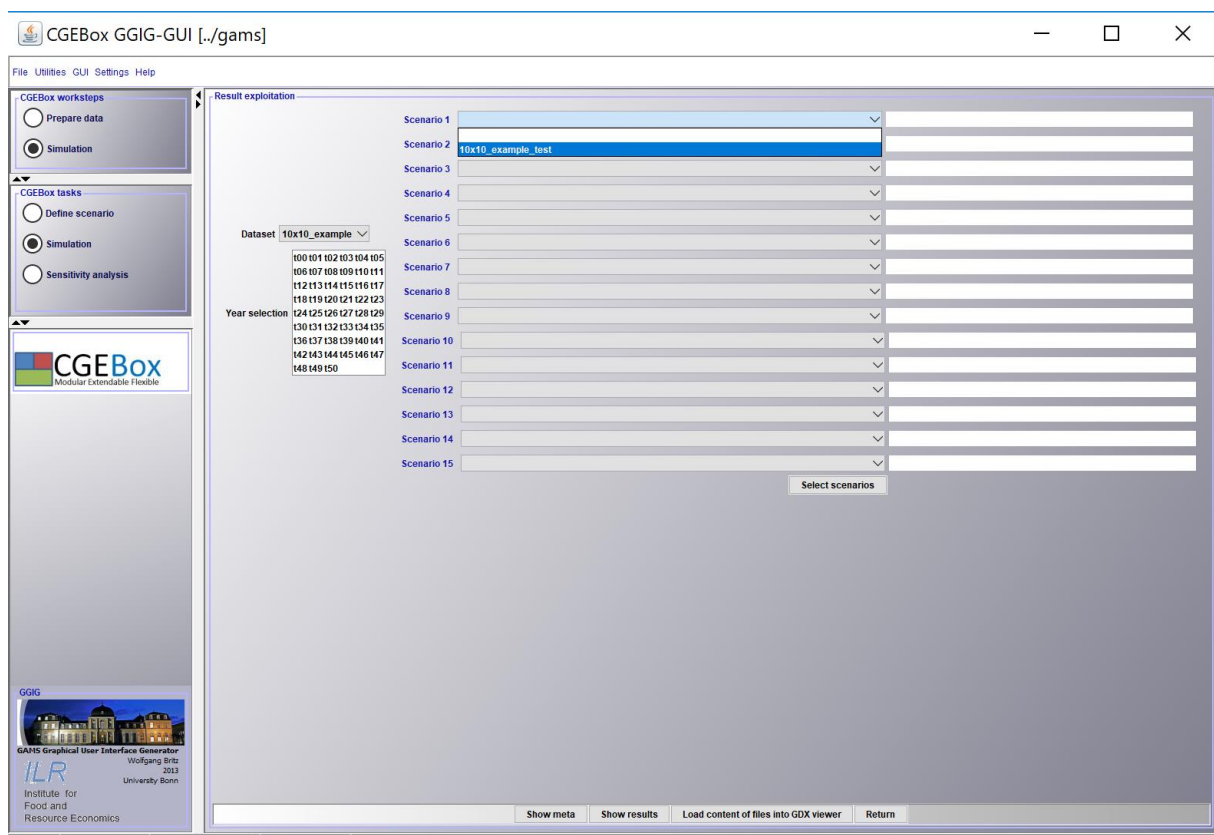
You will see the interface for the “Scenario editor” which is discussed later. We skip that functionality here and move directly to the task “Simulation”:



## First steps



and next pressing the “Exploit results” button. The layout will change as seen below. Please click now once in the “Dataset” selection (otherwise you will get a warning message later):

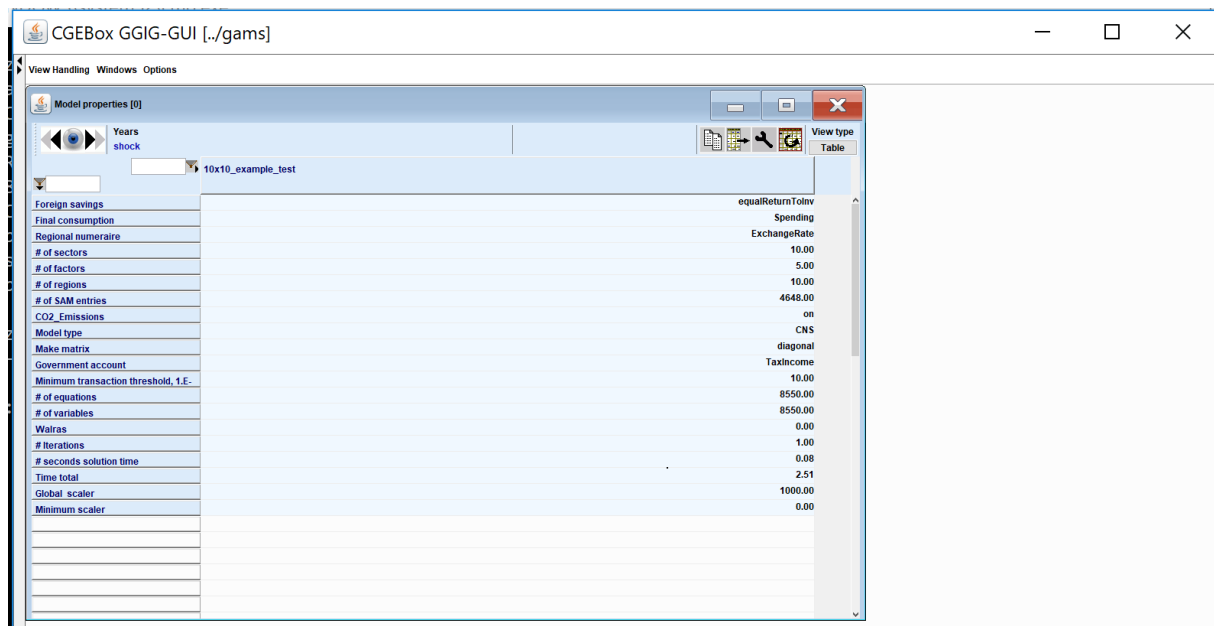


**Figure: Scenario selection for exploitation**

Select the scenario “10x\_10\_example\_test” and press “show results”, you should see a table similar as the one shown below:

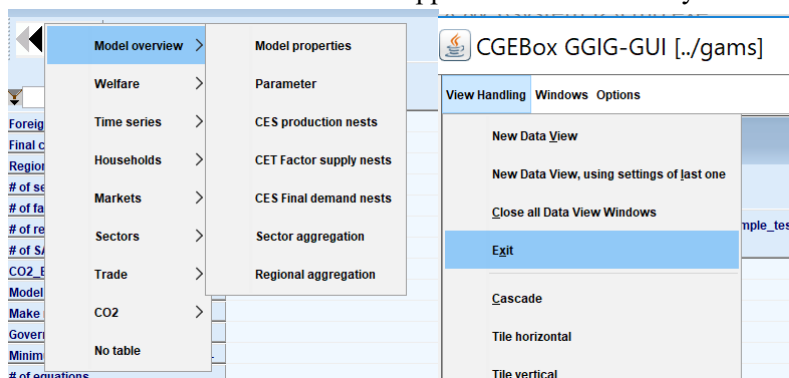
Note: In the year selection panel, nothing should be selected for comparative-static runs.





(Once you have produced your own results, you can find the data bases you used and the scenarios generated in the drop-down boxes, and you might combine different scenarios for comparison.)

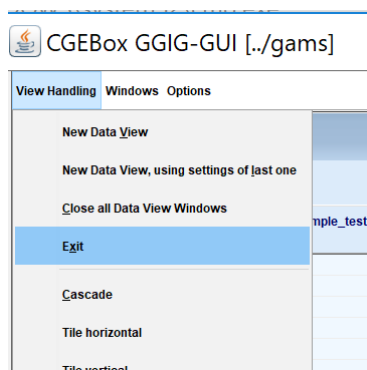
The “” view button in the upper left corner lets you select different views on your results:

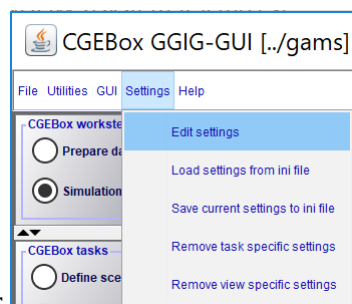


How to produce your own results is discussed in the following. Information on the different views is given in a section below.

## 7.7 Making the interface working with GAMS

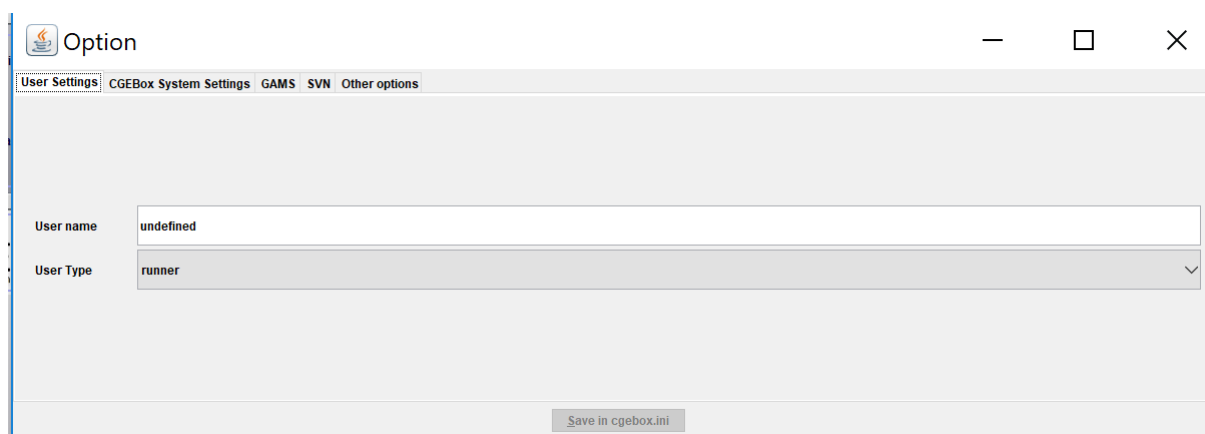
In order to use the GUI to generate data sets or for simulation runs, you have first to register a GAMS executable. In order to do so, first leave the result viewer as shown below



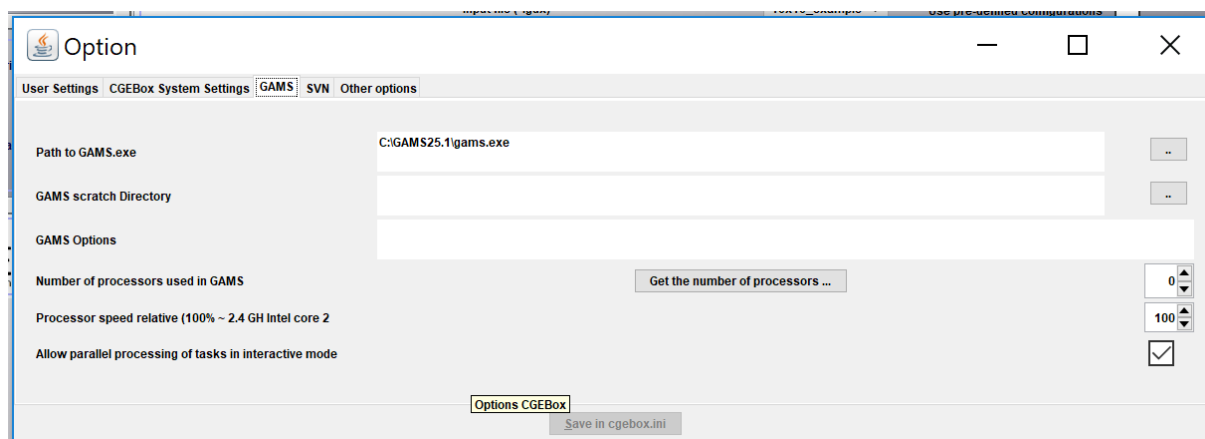


And open the setting dialogue from the menu bar

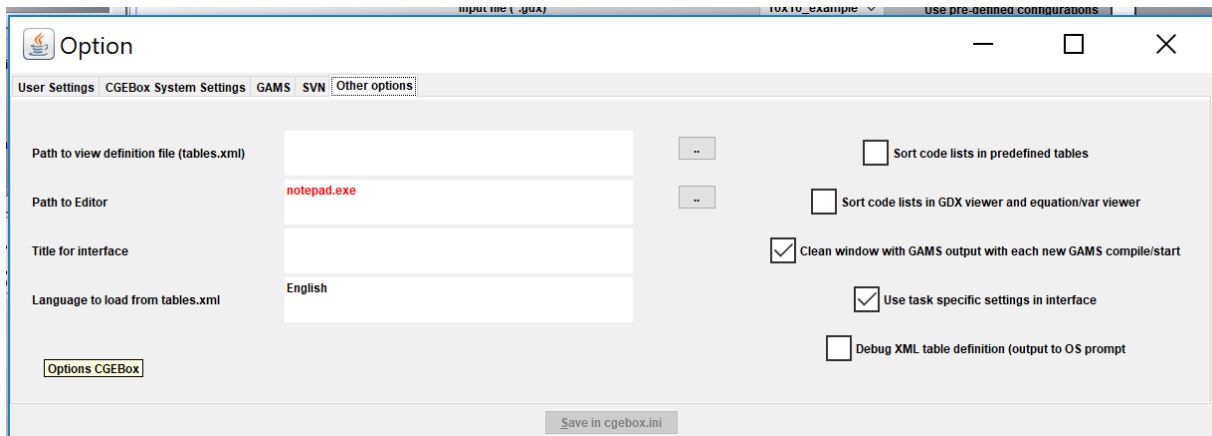
First, you should enter your name:



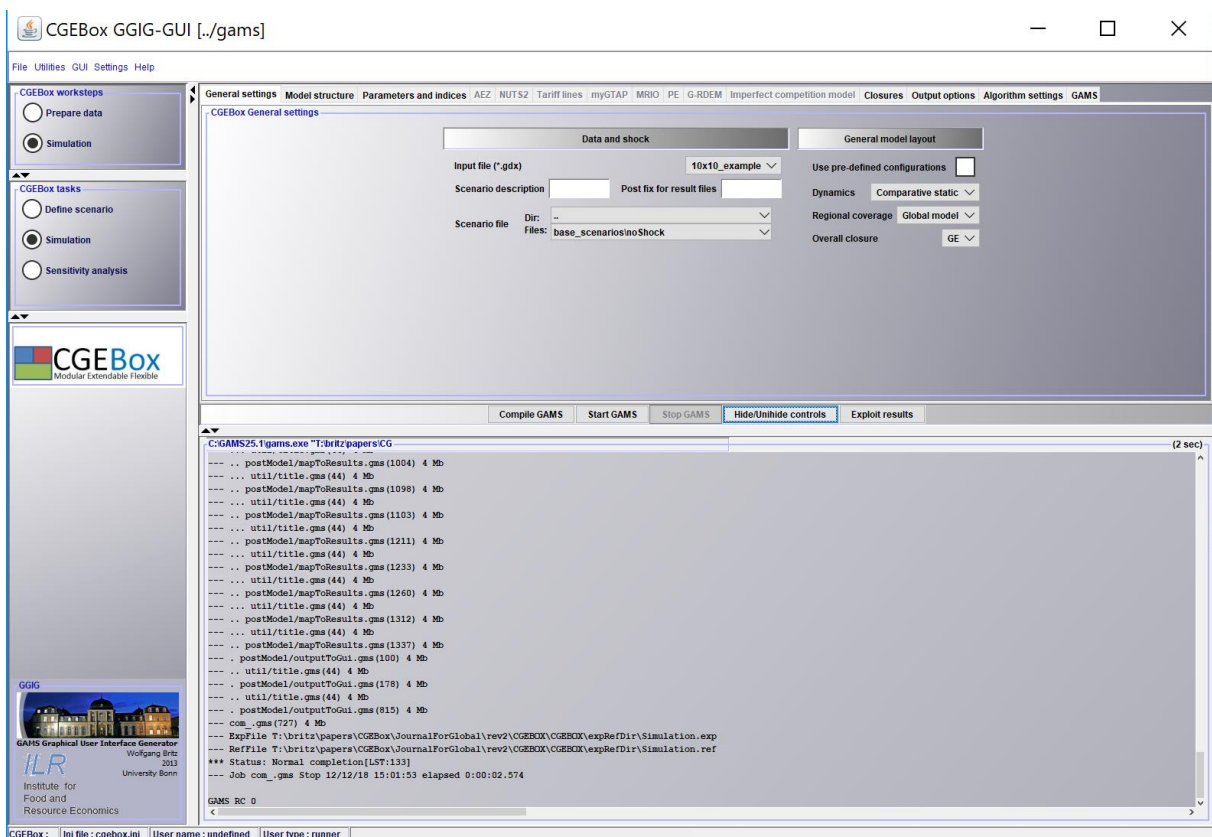
Next, click on the tab “GAMS” and either type in text field next to “Path to GAMS.EXE” which “GAMS.EXE” you would like to use. Alternatively, you can use the button to the right of the field to navigate to the directory where “GAMS.EXE” is found via a file selection dialogue. Please do not only enter a directory, but the full file name as shown below (and choose GAMS.EXE, not the user interface of GAMS, GAMSIDE.EXE). Depending how you install GAMS on your machine, the correct GAMS executable might already be entered by the Java code.



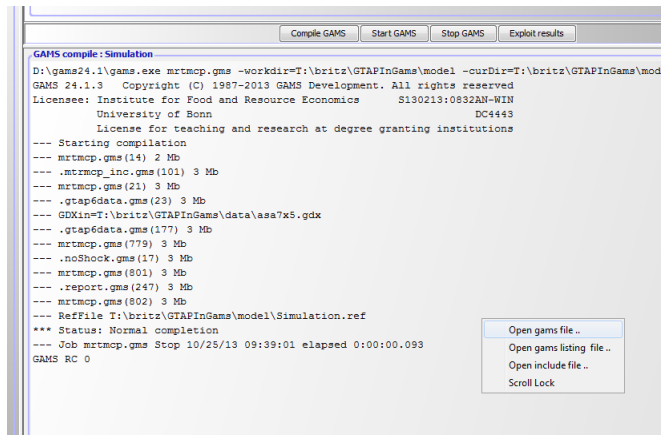
If you are using regularly a text editor, you can register it under “Other options”



In order to check if it worked, press the “Compile GAMS” button. You should see now the compile output from GAMS in the window below the button.



If you have registered an editor, you can also check its proper functioning: a right mouse click in the window with the output should open up a pop-up menu:



That pop-up menu offers you the options to open in your favorite editor the GAMS file, the listing produced by GAMS and the include file generated by the interface.

### Notes:

- The use of the GTAP-AGR or GTAP-E module works best with a data base with a full disaggregation of primary agriculture respectively detail for the energy sector.
- Some pre-defined configuration such as “MIRAGE” comprise also list of sector which assume that you use a data base with full sector resolution.
- You cannot use the global bank mechanism (Global equal returns to capital) if capital is declared immobile. There are further combinations of options where warnings are raised.
- A license either for PATH or CONOPT is required to solve the model.

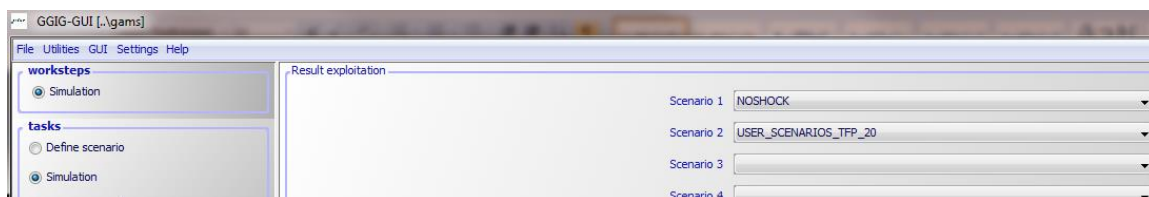
## 7.8 Inspecting the results

Before we turn to the exploitations tool of the GUI, the reader should be reminded that CGEBox allows for a wider range of possibilities to look at results:

1. As with any GAMS model, solprint=on will produce a listing of the model solution showing variables and equations, see the GAMS documentation for details. That option can be switched on from the interface as discussed above.
2. If additionally, LIMROW and LIMCOL are non-zero, the individual equations will be shown with their Jacobian entries at current variable levels. The numbers entered under LIMROW and LIMCOL will limit the output of individual instances shown and can be set by the GUI. For more information on how to interpret the resulting output, please refer to the GAMS manual.
3. If the output option GDX is switched on, all GAMS symbols (sets, parameters, variables) will be outputted to a GDX container and can be analyzed in a GDX viewer, for instance, the one delivered with the GAMS IDE or the GDX viewer built into the GUI.
4. The model's results can be mapped back into a SAM, and the SAM will be stored in the GDX.
5. The information from 1. -4. can also be inspected with the “Equation and Variable viewer” discussed in another section, which also provides a decomposition of each equation similar to “analyzeGE”.
6. Finally, the exploitation tools can be used, which also cover the SAM, as discussed in the following.

Which option works best depends on the use case and user preferences. Model listings with “limcol” and “limrow” switched on are extremely useful during model development and debugging, whereas the exploitation tools are probably the best approach to systematically analyze results from a shock. Individual equation decomposition can complement both approaches. We can only recommend trying out all approaches at least once to find out which one fits best one's preferences under specific use scenarios.

Inspecting results with GUI is nothing new: You have done that already if you followed the short introduction ... so press on the “Exploit results” button and next on “Show results” again. Use the “NoShock” and your own scenario:



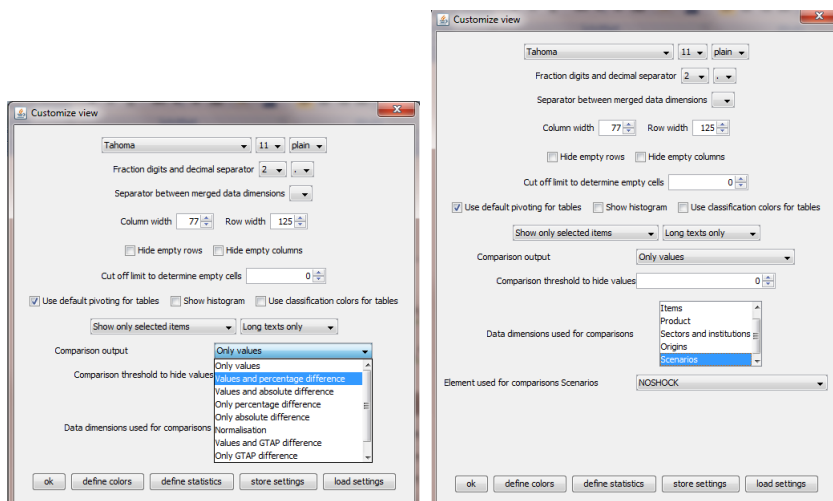
And now, you can see the counterfactual and the simulation side by side in the tables:

## Inspecting the results

The screenshot shows the GGIG-GUI interface with a table of economic indicators. The table has columns for 'Regions' (World), 'Total or household groups' (Total), and 'USER\_SCENARIOS\_TFP\_20'. The 'Customize Table' option is highlighted in the context menu.

| Indicator                            | World     | Total     | USER_SCENARIOS_TFP_20 |
|--------------------------------------|-----------|-----------|-----------------------|
| Total utility [Index]                | 1.000     | 1.218     |                       |
| Utility of private household [Index] | 1.000     | 1.215     |                       |
| Utility of government [Index]        | 1.000     | 1.225     |                       |
| Utility of saving [Index]            | 1.000     | 1.220     |                       |
| Total income [Mtl \$]                | 55518.188 | 56632.293 |                       |
| Savings price [Index]                | 7.000     | 5.801     |                       |
| Factor income [Mtl \$]               | 39484.758 | 40464.602 |                       |
| Current account balance [Mtl \$]     |           | 0.000     |                       |
| Tax income [Mtl \$]                  | 16033.430 | 16167.688 |                       |
| Population size [Mtl heads]          | 6620.293  | 6620.293  |                       |

You might now want to add relative differences: click with the mouse in the table or use the “Customize” (  ) button in order to customize your view (see screenshot below). Choose “Values and percentage difference” as the “Comparison output”, select “Scenarios” as the “Data dimensions used for comparisons” and make sure that the “NOSHOCK” scenario is used as the comparison element:



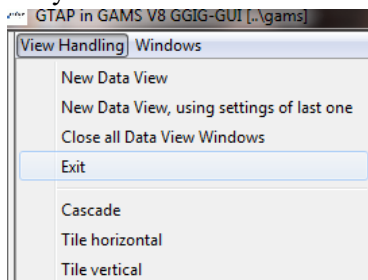
After closing the dialogue, you find now relative difference in small numbers beyond the simulated results in each cell of the table.

| Indicator                            | Value     | Percentage diff. to Scenarios NOSHOCK |
|--------------------------------------|-----------|---------------------------------------|
| Total utility [Index]                | 1.000     | 1.218                                 |
| Utility of private household [Index] | 1.000     | 21.753%                               |
| Utility of government [Index]        | 1.000     | 1.215                                 |
| Utility of saving [Index]            | 1.000     | 21.467%                               |
| Total income [Mil \$]                | 55518.188 | 1.225                                 |
| Consumer price [Index]               |           | 22.499%                               |
| Government price [Index]             |           | 1.220                                 |
| Savings price [Index]                | 7.000     | 21.981%                               |
| Factor income [Mil \$]               | 39484.758 | 56632.293                             |
| Current account balance [Mil \$]     |           | 2.007%                                |
| Tax income [Mil \$]                  | 16033.430 |                                       |
| Population size [Mil heads]          | 6620.293  | 5.801                                 |
|                                      |           | -17.137%                              |
|                                      |           | 40464.602                             |
|                                      |           | 2.462%                                |
|                                      |           | 0.000                                 |
|                                      |           | inf.                                  |
|                                      |           | 16167.688                             |
|                                      |           | 0.837%                                |
|                                      |           | 6620.293                              |
|                                      |           | 0.000%                                |

(Note: as expected, total utility increases, but prices drop ...).

There are many other functionalities found in the exploitation tools, the GGIG user manual discusses these options: [http://www.ilr.uni-bonn.de/agpo/staff/britz/GGIG\\_user\\_Guide.pdf](http://www.ilr.uni-bonn.de/agpo/staff/britz/GGIG_user_Guide.pdf).

If you want to leave the viewer again, use “Exit” from “View Handling” menu:

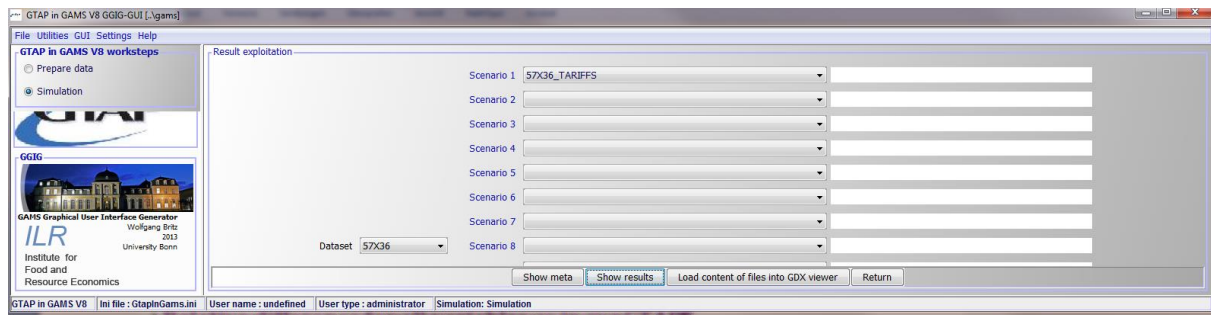


### 7.8.1 Notes for GEMPACK users or users familiar with other CGE models

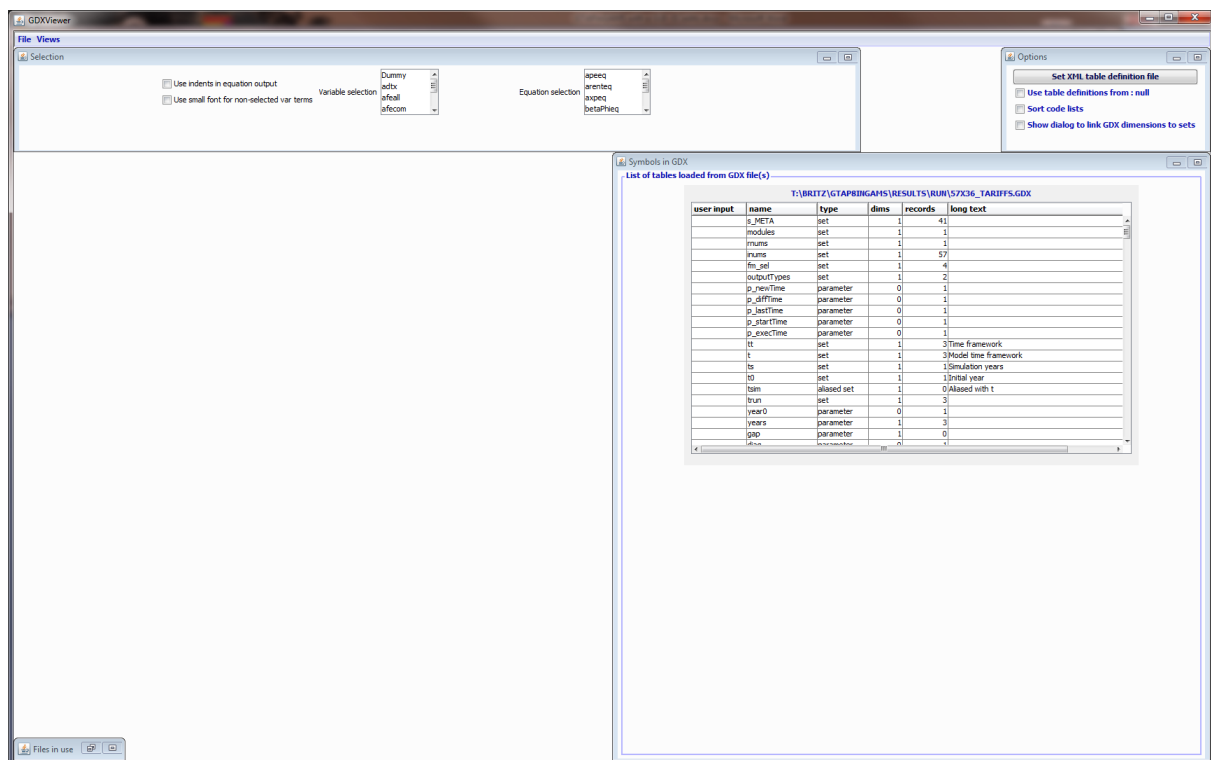
- The view “Model overview, SAM” shows a SAM generated from the model results (including row and column sums and potential imbalances) if the option “Store SAM in GDX” is switched on.
- Many tables comprise additional information on aggregates on top of the other rows. Only in some cases is what is called “total” a simple sum of the rows below. Thus, the information should generally not be confused with the “sum” information shown in “viewwhar”.
- There are only a few cases with a 1:1 relation between a variable in the model and numbers shown in table; typically the results shown in the different views stem from some post-model processing (e.g. multiplying a quantity index with a price).
- Instead of using the button, “Show results” which presents the different views discussed below, the “Load content of files in GDX viewer” allows to inspect all symbols used in the code, including all variables and the generated SAM (see next section). The GDX files are stored under “results/run” and can be naturally be opened with another GDX viewer such as the one comprised in the GAMSIDE.
- If the user want to inspect variables similar to “AnalyzeGE”, including decomposing equations, she can use the “Equation and variable viewer” (see below).
- Note again that all model equations are written in levels and not in relative changes.

### 7.8.2 Relative difference for all variables as in runGTAP

GEMPACK users typically compare simulated against benchmark levels for model variables. That is also possible with the interface. In order to do so, select the file with the result and press the “Load content of files into GDX viewer” button:



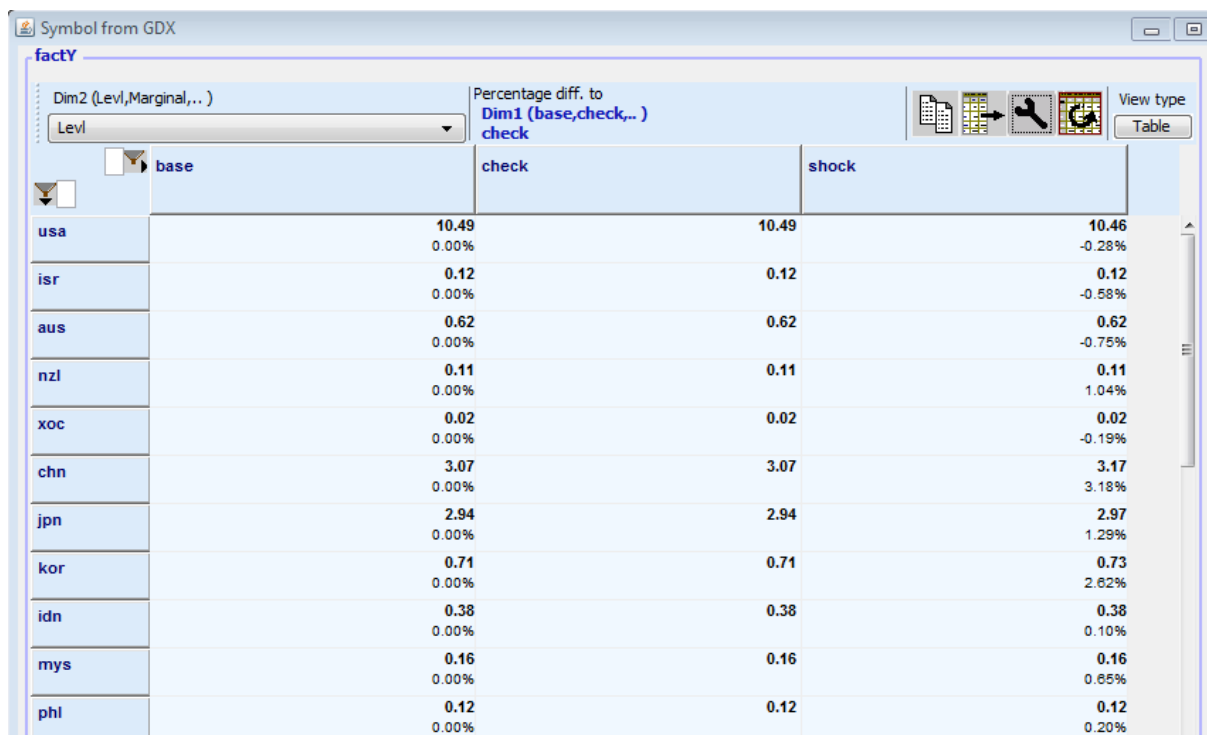
That will open a new window as shown below:



The window “List of table loaded from GDX file(s)” shows the status of all GAMS symbols (variable, equations, sets, parameters) after the solution of the shock. Double-click on any symbol you are interested in, e.g. the factor income “facty” as shown in the screen shot below. That will open a window where the results are shown, for the base case (= as initialized from the data base), the benchmark test (= check) and the shock.







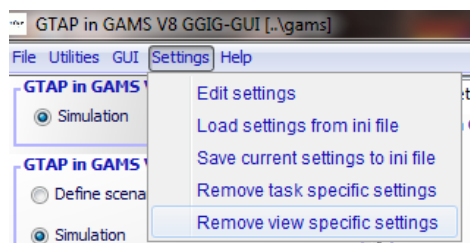
| Dim2 (Levl,Marginal,...) | base  | check | shock | Percentage diff. to Dim1 (base,check,...) |
|--------------------------|-------|-------|-------|-------------------------------------------|
| usa                      | 10.49 | 10.49 | 10.46 | -0.28%                                    |
| isr                      | 0.12  | 0.12  | 0.12  | -0.58%                                    |
| aus                      | 0.62  | 0.62  | 0.62  | -0.75%                                    |
| nzl                      | 0.11  | 0.11  | 0.11  | 1.04%                                     |
| xoc                      | 0.02  | 0.02  | 0.02  | -0.19%                                    |
| chn                      | 3.07  | 3.07  | 3.17  | 3.18%                                     |
| jpn                      | 2.94  | 2.94  | 2.97  | 1.29%                                     |
| kor                      | 0.71  | 0.71  | 0.73  | 2.62%                                     |
| idn                      | 0.38  | 0.38  | 0.38  | 0.10%                                     |
| mys                      | 0.16  | 0.16  | 0.16  | 0.65%                                     |
| phl                      | 0.12  | 0.12  | 0.12  | 0.20%                                     |

## 7.8.3 Trouble shooting with the viewer

If you are not sure what controls are for, try pressing the F1 button when the mouse is hovering over the control. In most cases, the PDF user guide will open at a page offering information.

Especially when working in the beginning with the viewer, one often ends up with a table showing no longer any data, an awkward selection, an unsuitable pivot etc.. The following strategies can be used to overcome such problems:

1.  use the “close” button to remove the view, and use “View Handling”, “New Data View” to open a fresh one.
2. Leave the viewer (“View Handling”, “Exit”), and click on “Remove view specific settings” under “Settings” in the menu bar. That will set the viewer back to the original defaults.



3. If that does also not help, try to close and re-open the application.
4. If the viewer still shows curious things, it is most probably some programming error ... contact the author.

### 7.8.4 Overview on existing views for exploitation

**Model overview/Model properties:** reports some basic attributes of the model instance as seen below, such as the modules switched one, closures used or the size of the model:

| Property                | Value            |
|-------------------------|------------------|
| Foreign savings         | equalReturnToInv |
| CO2_Emissions           | on               |
| # of sectors            | 10.00            |
| # of factors            | 5.00             |
| # of regions            | 10.00            |
| # of SAM entries        | 4648.00          |
| Model type              | CNS              |
| Make matrix             | diagonal         |
| Global scaler           | 1000.00          |
| # of equations          | 12380.00         |
| # of variables          | 12380.00         |
| # Iterations            | 15.00            |
| # seconds solution time | 0.53             |
| Time total              | 3.43             |

There are also two tables which report the aggregation of the data base under **Model overview/regional aggregation** and **Model overview/sectoral aggregation**:

| Region                       | Members                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Australia, New Zealand       | (Australia, New Zealand, Rest of Oceania, )                                                                                                                                                                                                                                                                                                                                |
| East Asia                    | (China, Hong Kong, Japan, Korea, Mongolia, Taiwan, Rest of East Asia, )                                                                                                                                                                                                                                                                                                    |
| Southeast Asia               | (Cambodia, Indonesia, Lao People's Democratic Republ, Malaysia, Philippines, Singapore, Thailand, Viet Nam, Rest of Southeast Asia, )                                                                                                                                                                                                                                      |
| South Asia                   | (Bangladesh, India, Nepal, Pakistan, Sri Lanka, Rest of South Asia, )                                                                                                                                                                                                                                                                                                      |
| North America                | (Canada, United States of America, Mexico, Rest of North America, )                                                                                                                                                                                                                                                                                                        |
| Latin America                | (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Paraguay, Peru, Uruguay, Venezuela, Rest of South America, Costa Rica, Guatemala, Honduras, Nicaragua, Panama, El Salvador, Rest of Central America, Caribbean, )                                                                                                                                                   |
| European Union 25            | (Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, United Kingdom, )                                                                                                                          |
| Middle East and North Africa | (Rest of Western Asia, Egypt, Morocco, Tunisia, Rest of North Africa, )                                                                                                                                                                                                                                                                                                    |
| Sub-Saharan Africa           | (Benin, Burkina Faso, Cameroon, Cote d'Ivoire, Ghana, Guinea, Nigeria, Senegal, Togo, Rest of Western Africa, Central Africa, South Central Africa, Ethiopia, Kenya, Madagascar, Malawi, Mauritius, Mozambique, Rwanda, Tanzania, Uganda, Zambia, Zimbabwe, Rest of Eastern Africa, Botswana, Namibia, South Africa, Rest of South African Customs, )                      |
| Rest of World                | (Oman, Israel, Switzerland, Norway, Rest of EFTA, Albania, Bulgaria, Belarus, Croatia, Romania, Russian Federation, Ukraine, Rest of Eastern Europe, Rest of Europe, Kazakhstan, Kyrgyzstan, Rest of Former Soviet Union, Armenia, Azerbaijan, Georgia, Bahrain, Iran Islamic Republic of, Kuwait, Qatar, Saudi Arabia, Turkey, United Arab Emirates, Rest of the World, ) |

## Inspecting the results

Sector aggregation [0]

Scenarios  
TEST

|                                                                                                                                                                                         |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| Grains and Crops<br>(Paddy rice, Wheat, Cereal grains nec, Vegetables, fruit, nuts, Oil seeds, Sugar cane, sugar beet, Plant-based fibers, Crops nec, Processed rice, )                 |  |
| Livestock and Meat Products<br>(Cattle,sheep,goats,horses, Animal products nec, Raw milk, Wool, silk-worm cocoons, Meat: cattle,sheep,goats,horse, Meat products nec, )                 |  |
| Mining and Extraction<br>(Minerals nec, Forestry, Fishing, Coal, Oil, Gas, )                                                                                                            |  |
| Processed Food<br>(Vegetable oils and fats, Dairy products, Sugar, Food products nec, Beverages and tobacco products, )                                                                 |  |
| Textiles and Clothing<br>(Textiles, Wearing apparel, )                                                                                                                                  |  |
| Light Manufacturing<br>(Leather products, Wood products, Paper products, publishing, Metal products, Motor vehicles and parts, Transport equipment nec, Manufactures nec, )             |  |
| Heavy Manufacturing<br>(Petroleum, coal products, Chemical,rubber,plastic prods, Mineral products nec, Ferrous metals, Metals nec, Electronic equipment, Machinery and equipment nec, ) |  |
| Utilities and Construction<br>(Electricity, Gas manufacture, distribution, Water, Construction, )                                                                                       |  |
| Transport and Communication<br>(Trade, Transport nec, Sea transport, Air transport, Communication, )                                                                                    |  |
| Other Services<br>(Insurance, Financial services nec, Business services nec, Recreation and other services, PubAdmin/Defence/Health/Educat, Dwellings, )                                |  |

**Model Overview/SAM** reports the results in a SAM, including row and columns sum and potential imbalances. Due to rounding errors and feasibility tolerances of the solvers, small imbalances in relative terms are possible and not of concern:

SAM [1]

Regions: World  
Origins: Total  
Years: shock

10x10\_Tariffs

| Grain                        | Oceania  | EastAsia   | SEAsia    | SouthAsia | NAmerica   | LatinAmer | EU_25      | MEIA      | SSA       | RestofWorld | Grains and Crops | Liv Me |
|------------------------------|----------|------------|-----------|-----------|------------|-----------|------------|-----------|-----------|-------------|------------------|--------|
| Australia, New Zealand       |          |            |           |           |            |           |            |           |           |             | 5470.86          |        |
| East Asia                    |          |            |           |           |            |           |            |           |           |             | 13403.21         |        |
| Southeast Asia               |          |            |           |           |            |           |            |           |           |             | 19017.48         |        |
| South Asia                   |          |            |           |           |            |           |            |           |           |             | 13406.56         |        |
| North America                |          |            |           |           |            |           |            |           |           |             | 78980.38         |        |
| Latin America                |          |            |           |           |            |           |            |           |           |             | 53563.66         |        |
| European Union 25            |          |            |           |           |            |           |            |           |           |             | 72707.31         |        |
| Middle East and North Africa |          |            |           |           |            |           |            |           |           |             | 7597.30          |        |
| Sub-Saharan Africa           |          |            |           |           |            |           |            |           |           |             | 17189.90         |        |
| Rest of World                |          |            |           |           |            |           |            |           |           |             | 27491.80         |        |
| Grains and Crops             | 1628.41  | 45166.89   | 14725.96  | 10667.06  | 34534.85   | 13419.00  | 96181.22   | 16060.23  | 8774.68   | 33307.83    |                  |        |
| Livestock and Meat Products  | 1257.37  | 27600.54   | 3567.70   | 710.45    | 17218.91   | 4356.69   | 63032.86   | 2422.86   | 2899.07   | 19383.95    |                  |        |
| Mining and Extraction        | 9455.21  | 416827.66  | 68202.07  | 90717.98  | 297363.31  | 35377.76  | 423494.28  | 8742.51   | 15718.52  | 102812.74   |                  |        |
| Processed Food               | 9131.17  | 65037.52   | 25261.48  | 12142.52  | 80825.78   | 22016.18  | 235287.66  | 15308.34  | 18758.60  | 61598.29    |                  |        |
| Textiles and Clothing        | 7148.06  | 79334.26   | 22636.69  | 11149.33  | 125823.12  | 19808.73  | 218646.17  | 13868.54  | 12829.08  | 73258.90    |                  |        |
| Light Manufacturing          | 50377.86 | 226457.47  | 79897.42  | 39464.47  | 648833.81  | 97419.51  | 1151365.38 | 38315.93  | 59522.32  | 333534.00   |                  |        |
| Heavy Manufacturing          | 96718.02 | 1261821.50 | 387924.38 | 154015.48 | 1176961.00 | 240079.50 | 2372662.00 | 101033.31 | 119092.05 | 698601.50   |                  |        |
| Utilities and Construction   | 1444.59  | 19565.64   | 6879.54   | 1591.53   | 9805.61    | 4257.45   | 75993.51   | 3427.62   | 6088.09   | 36543.52    |                  |        |
| Transport and Communication  | 19522.31 | 173821.56  | 50788.73  | 18943.08  | 137347.89  | 34221.13  | 460401.50  | 11411.07  | 19856.88  | 100875.77   |                  |        |
| Other Services               | 22045.32 | 165100.83  | 64577.36  | 40070.86  | 252669.31  | 46522.71  | 692118.50  | 21865.73  | 33369.14  | 158218.98   |                  |        |
| Unskilled labor              |          |            |           |           |            |           |            |           |           |             |                  |        |
| Skilled labor                |          |            |           |           |            |           |            |           |           |             |                  |        |
| Capital                      |          |            |           |           |            |           |            |           |           |             |                  |        |
| Land                         |          |            |           |           |            |           |            |           |           |             |                  |        |
| Household                    |          |            |           |           |            |           |            |           |           |             |                  |        |

Note: the SAM can only be shown if SAM output was activated on the interface when the simulation was run.

**Model Overview/CES products nests** reports the nesting structure used in the production module:

**CES production nests [0]**

Regions: Oceania    Items: dNests    sectors: Paddy rice    Years: shock

|                                | ND   | VA | Labor | CAP+EME | energy | non-electric | non-coal |
|--------------------------------|------|----|-------|---------|--------|--------------|----------|
| Paddy rice                     | 1.00 |    |       |         |        |              |          |
| Wheat                          | 1.00 |    |       |         |        |              |          |
| Cereal grains nec              | 1.00 |    |       |         |        |              |          |
| Vegetables, fruit, nuts        | 1.00 |    |       |         |        |              |          |
| Oil seeds                      | 1.00 |    |       |         |        |              |          |
| Sugar cane, sugar beet         | 1.00 |    |       |         |        |              |          |
| Plant-based fibers             | 1.00 |    |       |         |        |              |          |
| Crops nec                      | 1.00 |    |       |         |        |              |          |
| Cattle,sheep,goats,horses      | 1.00 |    |       |         |        |              |          |
| Animal products nec            | 1.00 |    |       |         |        |              |          |
| Raw milk                       | 1.00 |    |       |         |        |              |          |
| Wool, silk-worm cocoons        | 1.00 |    |       |         |        |              |          |
| Forestry                       | 1.00 |    |       |         |        |              |          |
| Fishing                        | 1.00 |    |       |         |        |              |          |
| Coal                           |      |    |       |         |        | 0.50         |          |
| Oil                            |      |    |       |         |        |              | 1.00     |
| Gas                            |      |    |       |         |        |              | 1.00     |
| Minerals nec                   | 1.00 |    |       |         |        |              |          |
| Meat: cattle,sheep,goats,horse | 1.00 |    |       |         |        |              |          |
| Meat products nec              | 1.00 |    |       |         |        |              |          |
| Vegetable oils and fats        | 1.00 |    |       |         |        |              |          |
| Dairy products                 | 1.00 |    |       |         |        |              |          |
| Processed rice                 | 1.00 |    |       |         |        |              |          |
| Sugar                          | 1.00 |    |       |         |        |              |          |

Similar views are available for the nesting used for factor supply and final demand:

**CES Factor supply nests [0]**

Regions: Oceania    Items: fNests    xft

| Paddy rice                     | on |
|--------------------------------|----|
| Wheat                          | on |
| Cereal grains nec              | on |
| Vegetables, fruit, nuts        | on |
| Oil seeds                      | on |
| Sugar cane, sugar beet         | on |
| Plant-based fibers             | on |
| Crops nec                      | on |
| Cattle,sheep,goats,horses      | on |
| Animal products nec            | on |
| Raw milk                       | on |
| Wool, silk-worm cocoons        | on |
| Forestry                       | on |
| Fishing                        | on |
| Coal                           | on |
| Oil                            | on |
| Gas                            | on |
| Minerals nec                   | on |
| Meat: cattle,sheep,goats,horse | on |
| Meat products nec              | on |
| Vegetable oils and fats        | on |
| Dairy products                 | on |
| Processed rice                 | on |
| Sugar                          | on |
| Food products nec              | on |
| Beverages and tobacco products | on |
| Textiles                       | on |
| Wearing apparel                | on |

**CES Final demand nests [0]**

Regions: Oceania    Items: dNests    Final demand: Private consumption    Years: shock

|                                | energy | non-electric | non-coal |
|--------------------------------|--------|--------------|----------|
| Paddy rice                     | on     |              |          |
| Wheat                          | on     |              |          |
| Cereal grains nec              | on     |              |          |
| Vegetables, fruit, nuts        | on     |              |          |
| Oil seeds                      | on     |              |          |
| Sugar cane, sugar beet         | on     |              |          |
| Plant-based fibers             | on     |              |          |
| Crops nec                      | on     |              |          |
| Cattle,sheep,goats,horses      | on     |              |          |
| Animal products nec            | on     |              |          |
| Raw milk                       | on     |              |          |
| Wool, silk-worm cocoons        | on     |              |          |
| Forestry                       | on     |              |          |
| Fishing                        | on     |              |          |
| Coal                           |        | on           |          |
| Oil                            |        |              | on       |
| Gas                            |        |              | on       |
| Minerals nec                   | on     |              |          |
| Meat: cattle,sheep,goats,horse | on     |              |          |
| Meat products nec              | on     |              |          |
| Vegetable oils and fats        | on     |              |          |
| Dairy products                 | on     |              |          |
| Processed rice                 | on     |              |          |
| Sugar                          | on     |              |          |
| Food products nec              | on     |              |          |
| Beverages and tobacco products | on     |              |          |
| Textiles                       | on     |              |          |
| Wearing apparel                | on     |              |          |

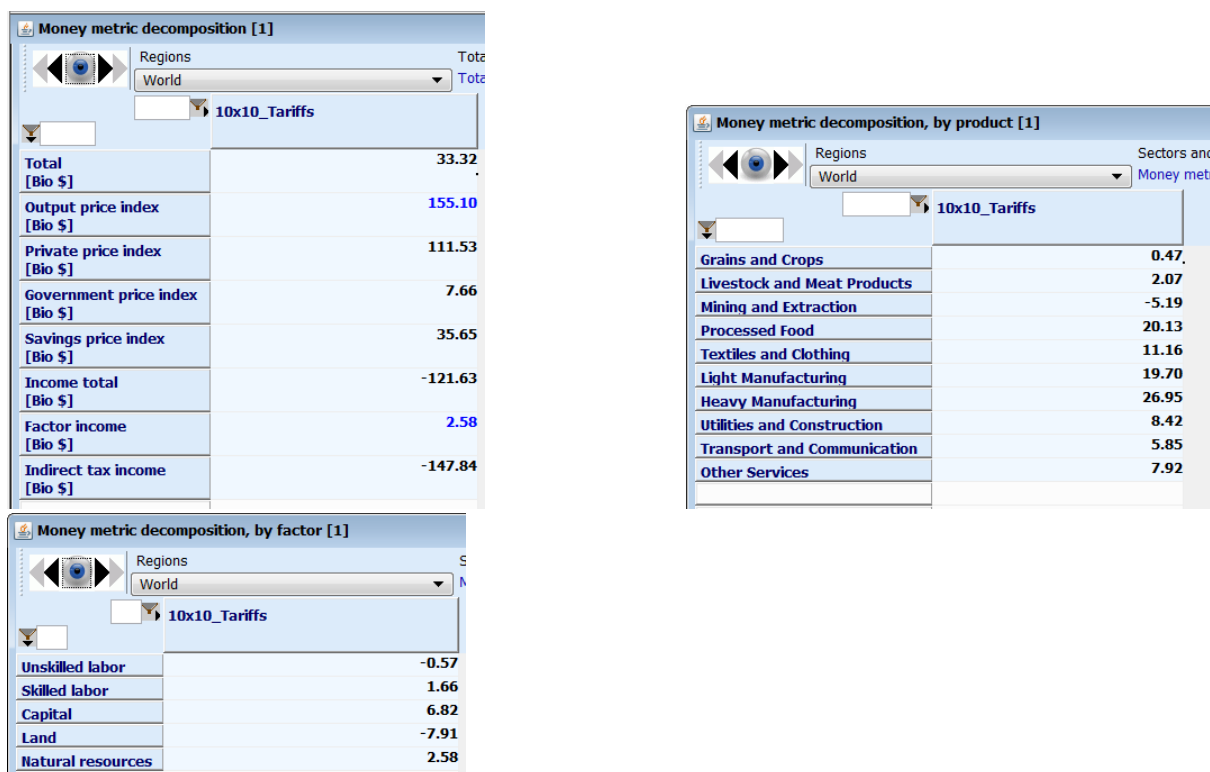
The table **Welfare/Income** provides an overview in key results such as GDP and price indices:

**Income [1]**

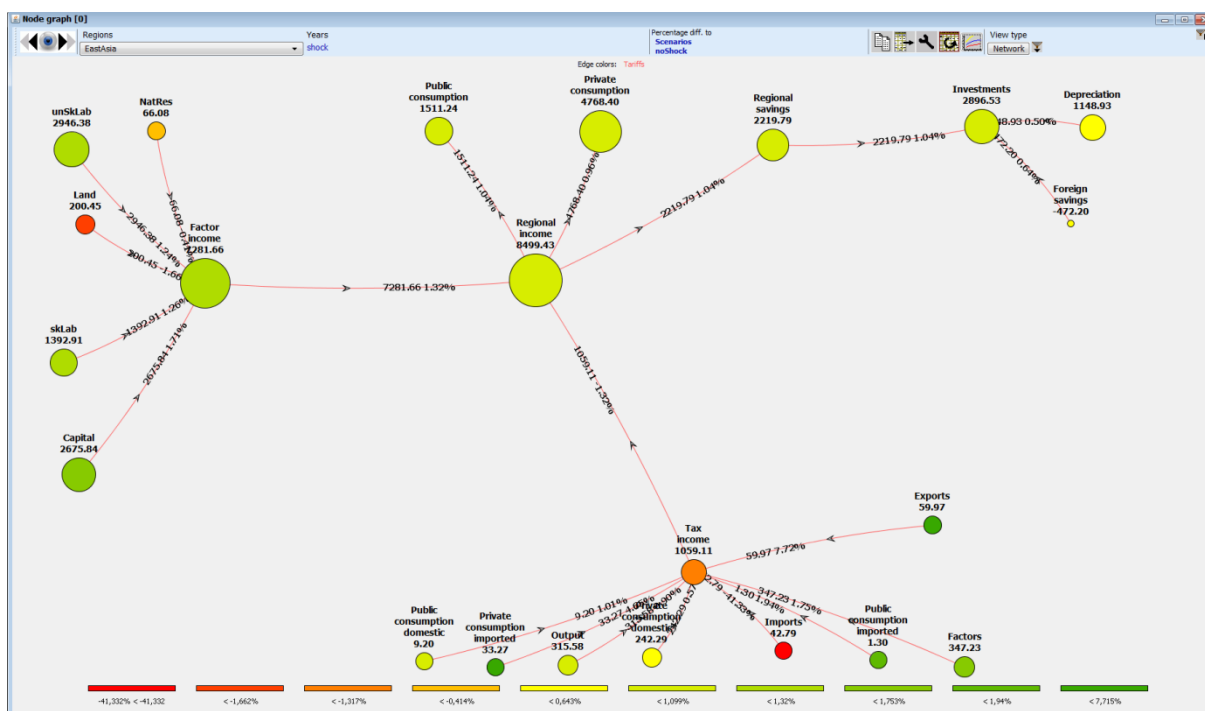
Regions: World    Total or household: Total    10x10\_Tariffs

|                                      |          |
|--------------------------------------|----------|
| Money metric [Bio \$]                | 33.32    |
| Total utility [Index]                | 1.00     |
| Utility of private household [Index] | 1.00     |
| Utility of government [Index]        | 1.00     |
| Utility of saving [Index]            | 1.00     |
| GDP [Bio \$]                         | 49780.15 |
| Consumer price [Index]               | 0.99     |
| Government price [Index]             | 1.00     |
| Factor income [Bio \$]               | 39510.97 |
| Current account balance [Bio \$]     | -0.00    |
| Tax income [Bio \$]                  | 9547.97  |
| Population size [Mil heads]          | 6620.29  |

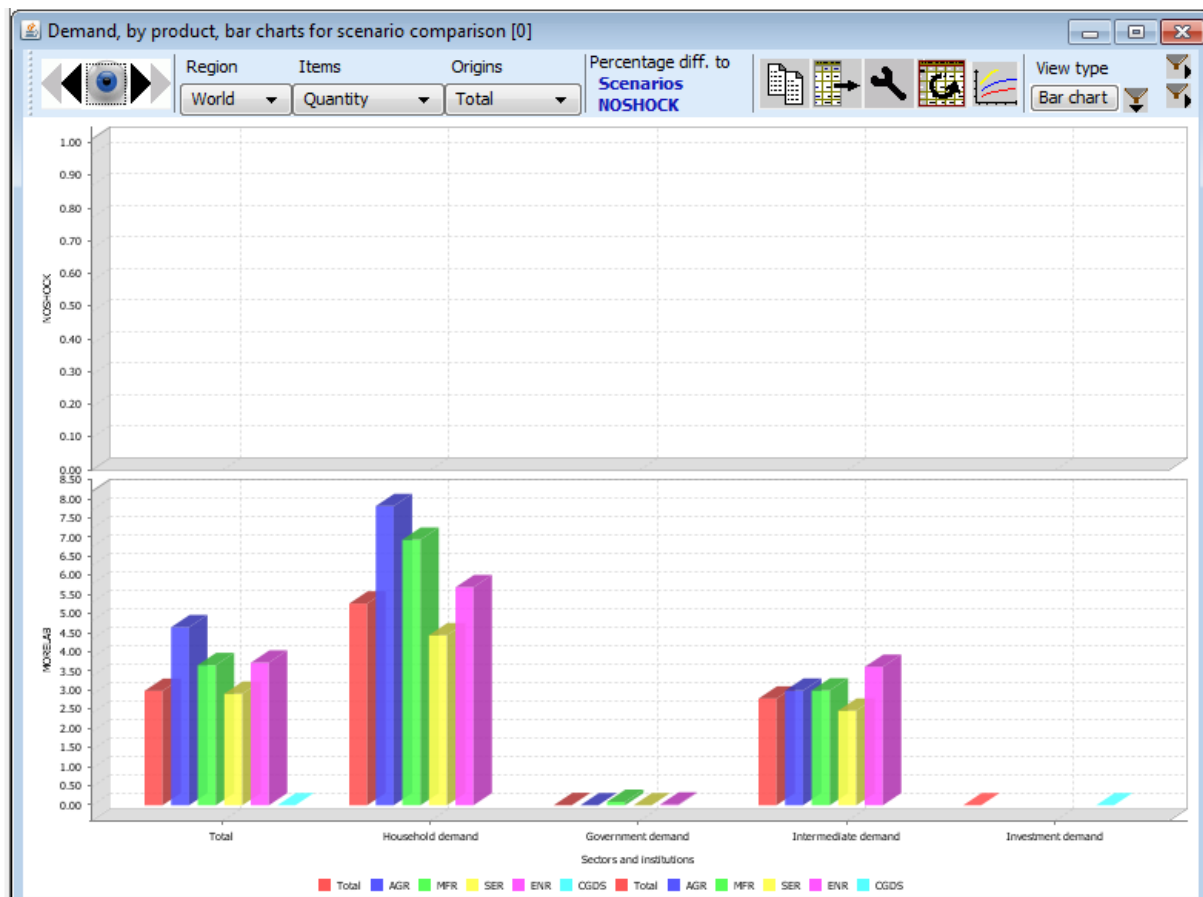
The three tables **Welfare/Money metric decomposition**, **Welfare/Money metric decomposition, by product** and **Welfare/ Money metric decomposition, by factor** support an welfare analysis. “By products” reports welfare gains or losses related to output price changes, by factor linked to changes in factor prices or quantities:



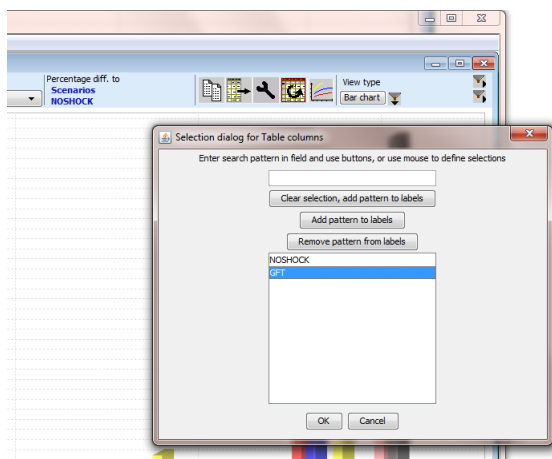
A graphical view on income generation and use provides **Welfare/Node graph**, the following example uses advanced options, such as coloring of the nodes by simulated change and sizing them according to simulated value:







Note: you use the second “filter” bottom next to view type to “view type” (double click on it) to open a dialogue which lets you remove the benchmark scenario:



That will produce a nicer graph.

**Markets/Demand by institution:** same content as before, but different pivot (could also be quickly produced by using the pivot possibility which can be applied to any view)



Demand, by institution [0]

Region: USA, Items: Quantity, Origins: Total

NOSHOCK

|                     | Total    | AGR    | MFR     | SER      | ENR    | Other   | Total    | AGR   |
|---------------------|----------|--------|---------|----------|--------|---------|----------|-------|
| Total               | 21263.64 | 242.58 | 5951.03 | 12460.10 | 619.30 | 1990.64 | 21931.60 | 261.5 |
| Household demand    | 6925.93  | 48.30  | 1389.33 | 5352.55  | 135.75 |         | 6623.22  | 42.5  |
| Government demand   | 1528.65  | 1.78   | 228.70  | 1298.10  | 0.06   |         | 1529.02  | 1.7   |
| Intermediate demand | 9910.88  | 161.00 | 3706.44 | 5571.60  | 471.84 |         | 10417.64 | 181.1 |
| Investment demand   | 1990.64  |        |         |          |        | 1990.64 | 1990.64  |       |
| Export demand       | 888.88   | 31.49  | 626.56  | 219.18   | 11.64  |         | 1350.66  | 35.9  |

**Markets/Final demand nests:** shows the demand aggregates –quantities, prices and values – according to the active demand nesting:

Final demand nests [0]

Regions: Oceania, Items: Quantity, Origins: Demand

|              | Private consumption | Government | Investment |
|--------------|---------------------|------------|------------|
| top          | 576.57              | 184.14     | 271.59     |
| energy       | 18.43               | 0.01       | 11.03      |
| non-electric | 13.66               | 0.00       | 0.00       |
| non-coal     | 13.65               | 0.00       | 0.00       |

**Markets/Balance:** Shows total use and how it is sourced:

Markets, balance [0]

Region: World, Items: Quantity, Origins: Total

|                                   | Use       | Domestic  | Imports  |
|-----------------------------------|-----------|-----------|----------|
| Total                             | 136815.61 | 121112.05 | 15703.55 |
| Grains and Crops                  | 2194.30   | 1864.09   | 330.21   |
| Livestock and Meat Products       | 1931.71   | 1786.12   | 145.59   |
| Mining and Extraction             | 4864.46   | 3282.53   | 1581.93  |
| Processed Food                    | 4500.96   | 3911.90   | 589.05   |
| Textiles and Clothing             | 2405.43   | 1787.59   | 617.84   |
| Light Manufacturing               | 12850.76  | 9985.44   | 2865.32  |
| Heavy Manufacturing               | 26660.95  | 19783.81  | 6877.14  |
| Utilities and Construction        | 11185.15  | 11018.84  | 166.31   |
| Transport and Communication       | 19005.00  | 17974.03  | 1030.97  |
| Other Services                    | 38384.38  | 36885.19  | 1499.19  |
| That must be the investment good? | 12832.52  | 12832.52  |          |

**Demand, per product, per capita in \$** is an example for a tabled definition which uses the in-built expression evaluator, it divides total demand by the population size:

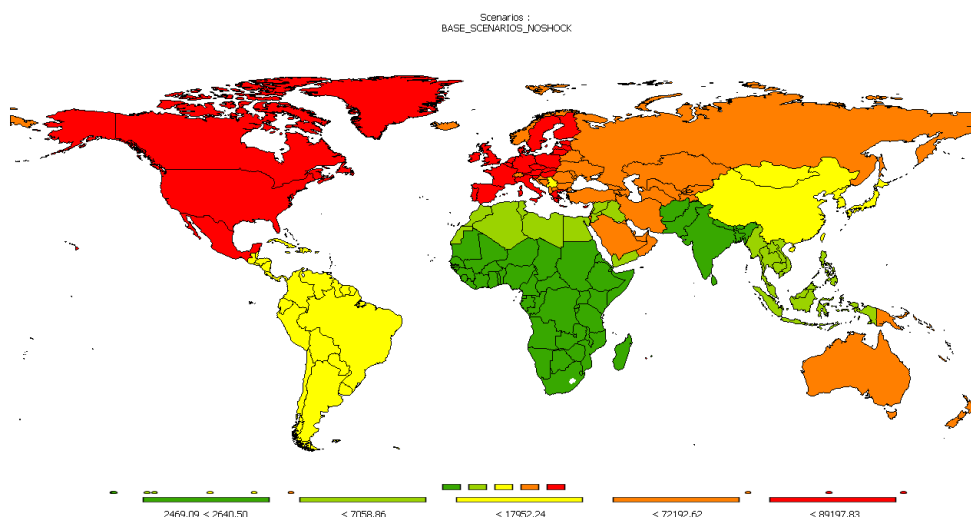
Demand, by product, per capita in \$ [0]

Region: World Items: Q Origins: Total View type: Table

BASE\_SCENARIOS\_NOSHOCK

|                                   | Total    | Household demand | Government demand | Intermediate demand | Investment demand | Export demand |
|-----------------------------------|----------|------------------|-------------------|---------------------|-------------------|---------------|
| Total                             | 19660.35 | 4602.31          | 1424.48           | 9700.63             | 1938.36           | 1924.02       |
| Grains and Crops                  | 217.49   | 89.83            |                   | 111.79              |                   | 15.87         |
| Livestock and Meat Products       | 212.19   | 101.69           |                   | 100.34              |                   | 10.15         |
| Mining and Extraction             | 638.36   | 7.88             |                   | 444.66              |                   | 185.82        |
| Processed Food                    | 635.65   | 372.83           |                   | 191.11              |                   | 71.71         |
| Textiles and Clothing             | 244.44   | 103.20           |                   | 89.15               |                   | 52.09         |
| Light Manufacturing               | 1901.86  | 352.23           | 4.66              | 1157.64             |                   | 387.32        |
| Heavy Manufacturing               | 3883.43  | 345.09           | 2.69              | 2614.66             |                   | 920.99        |
| Utilities and Construction        | 1635.37  | 125.83           | 1.69              | 1507.48             |                   | 0.37          |
| Transport and Communication       | 2765.58  | 1183.57          | 15.02             | 1371.13             |                   | 125.32        |
| Other Services                    | 5587.62  | 1920.16          | 1400.42           | 2112.67             |                   | 154.38        |
| That must be the investment good? | 1938.36  |                  |                   |                     | 1938.36           |               |

That view can alternatively be visualized as a map in **Demand, per product, per capita in \$, map**:



Maps can be made from any table by putting the regions in the rows and choosing “Map” under “View type”.

**Markets/Intermediate demand:** break of the intermediate demand by sector. Note: the total in the columns is an aggregation over all sectors.

Intermediate demand [0]

Region: USA Items: Quantity Origin: Total

NOSHOCK

World

USA

EUR

CHN

ROW

MORELAB

|       | SER     | ENR    | CGDS    | Total   | AGR    | MFR     | SE       |
|-------|---------|--------|---------|---------|--------|---------|----------|
| Total | 9910.88 | 129.12 | 2976.38 | 4501.76 | 312.98 | 1990.64 | 10417.64 |
| AGR   | 161.00  | 18.24  | 125.81  | 15.86   | 0.03   | 1.05    | 181.18   |
| MFR   | 3706.44 | 49.22  | 1818.59 | 1029.74 | 12.45  | 796.45  | 3989.68  |
| SER   | 5571.60 | 58.84  | 922.14  | 3300.12 | 97.36  | 1193.13 | 5739.79  |
| ENR   | 471.84  | 2.82   | 109.84  | 156.04  | 203.14 | 0.00    | 506.98   |

**Markets/Factor demand:** report primary factor use, price etc. by sector

Factor demand [0]

Region: USA Items: Quantity Origins: Factor demand

NOSHOCK

MORELAB

|         | AGR     | MFR   | SER    | ENR     | CGDS   | Total   | AGR |
|---------|---------|-------|--------|---------|--------|---------|-----|
| LND     | 42.10   | 42.10 |        |         |        | 42.10   | 4   |
| LAB     | 3156.50 | 32.24 | 674.81 | 2414.43 | 35.02  | 3472.15 | 3   |
| SKL     | 2230.39 | 2.46  | 381.57 | 1830.26 | 16.09  | 2230.39 |     |
| Capital | 3644.02 | 38.89 | 694.84 | 2779.58 | 130.72 | 3644.02 | 4   |
| RES     | 22.61   | 2.09  | 2.35   | 18.17   |        | 22.61   |     |

**Sectors/Sector overview:** reports for each sector total output, output taxes, total intermediate and total factor demand, as well the individual intermediate and factor demands.

Sector overview [0]

Region: World Items: Quantity Origins: Total

NOSHOCK

MORE

|                                | AGR      | MFR     | SER      | ENR      | CGDS    | Total   |
|--------------------------------|----------|---------|----------|----------|---------|---------|
| Total output                   | 65315.98 | 2009.74 | 19605.71 | 34325.31 | 2633.08 | 6742.14 |
| Output taxes                   | 1050.87  | -5.75   | 320.25   | 645.96   | 90.41   |         |
| Total intermediate demand      | 35040.87 | 904.67  | 12931.68 | 13104.96 | 1517.95 | 6581.61 |
| Total intermediate input taxes | 491.28   | -2.40   | 110.98   | 207.20   | 14.97   | 160.53  |
| Total factor demand            |          |         |          |          |         |         |
| Total factor taxes             | 2813.62  | -2.29   | 661.88   | 2106.04  | 47.99   |         |
| AGR                            | 1329.38  | 282.08  | 889.01   | 115.06   | 13.98   | 29.24   |
| MFR                            | 14741.90 | 325.16  | 8106.09  | 3651.91  | 172.65  | 2486.09 |
| SER                            | 16812.19 | 249.35  | 3336.39  | 8756.92  | 403.27  | 4066.27 |
| ENR                            | 2157.39  | 48.08   | 600.19   | 581.06   | 928.05  | 0.01    |
| LND                            | 294.81   | 294.81  |          |          |         |         |
| LAB                            | 8605.80  | 484.29  | 2134.70  | 5844.96  | 141.86  |         |
| SKL                            | 5205.31  | 11.98   | 920.29   | 4201.65  | 71.39   |         |
| Capital                        | 11613.01 | 283.11  | 2510.56  | 8214.55  | 604.79  |         |
| RES                            | 200.41   | 41.32   | 15.38    |          | 143.71  |         |

**Sectors/Sector cost shares:** a selection of the value item of the view above, normalized by total output:

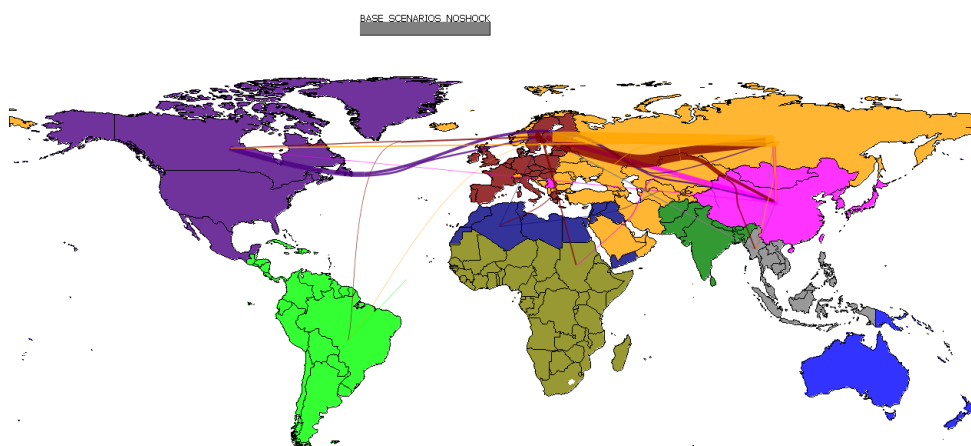
Cost shares [0]

Region: World | Items: Value | Origins: Domestic

NOSHOCK

|                                | Total | AGR   | MFR  | SER  | ENR  | CGDS |
|--------------------------------|-------|-------|------|------|------|------|
| Total output                   | 1.00  | 1.00  | 1.00 | 1.00 | 1.00 | 1.00 |
| Total intermediate demand      | 0.45  | 0.40  | 0.52 | 0.34 | 0.42 | 0.82 |
| Total intermediate input taxes | 0.01  | -0.00 | 0.00 | 0.01 | 0.00 | 0.02 |
| Total factor demand            |       |       |      |      |      |      |
| Total factor taxes             | 0.04  | -0.00 | 0.03 | 0.06 | 0.02 |      |
| AGR                            | 0.02  | 0.13  | 0.04 | 0.00 | 0.01 | 0.00 |
| MFR                            | 0.16  | 0.13  | 0.30 | 0.09 | 0.04 | 0.22 |
| SER                            | 0.24  | 0.12  | 0.16 | 0.24 | 0.14 | 0.60 |
| ENR                            | 0.03  | 0.02  | 0.03 | 0.02 | 0.22 | 0.00 |
| LND                            | 0.00  | 0.15  |      |      |      |      |
| LAB                            | 0.13  | 0.24  | 0.11 | 0.17 | 0.05 |      |
| SKL                            | 0.08  | 0.01  | 0.05 | 0.12 | 0.03 |      |
| Capital                        | 0.18  | 0.14  | 0.13 | 0.24 | 0.23 |      |
| RES                            | 0.00  | 0.02  | 0.00 |      | 0.05 |      |

There are further tables available to bi-lateral trade which can also be visualized as flow maps as shown in the screen shot below,



and more might be added in the future. But the examples above might be sufficient to judge if these views (in addition of using a GDX viewer to look out variables) are a useful addition to result analysis.

| Meta     | Items    | Origins |
|----------|----------|---------|
| Markets  | Quantity | Total   |
| Sectors  |          |         |
| Trade    |          |         |
| No table |          |         |

- Exports, trade matrix
- Exports, by product and destination
- Imports, trade matrix
- Imports, by product and origin

## 8. Sensitivity analysis

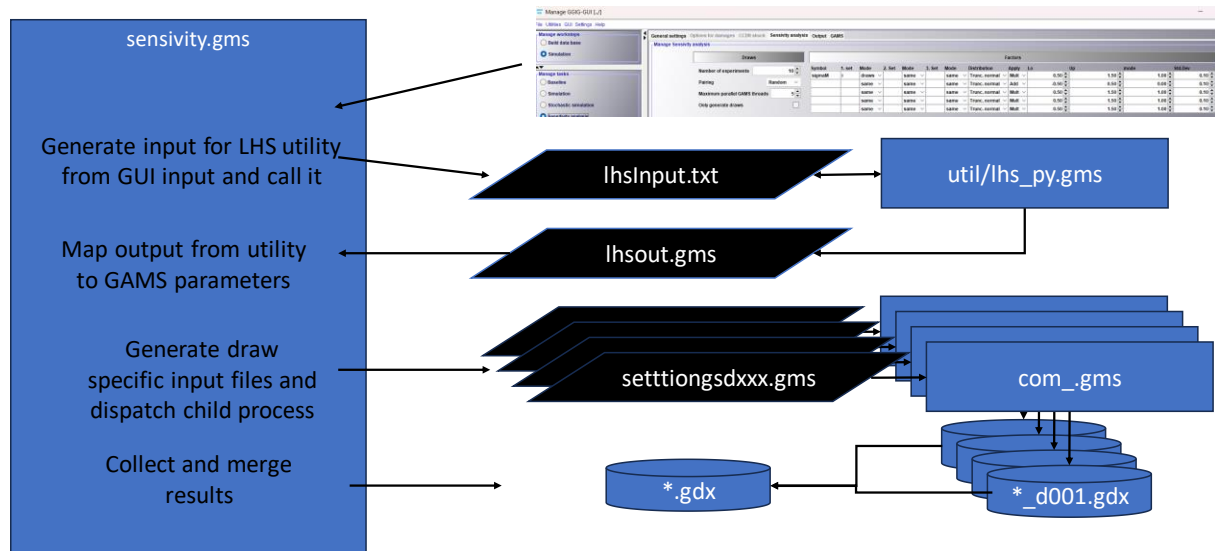
One of the daunting aspects of using larger simulation models is the high number of parameters needed, both behavioral ones and parameters linked to technology. Generally, most of these parameters are not known with certainty which raises the question how robust results from a shock are. The same often applies to the exact definition of a shock. The idea of sensitivity analysis is to repeatedly solve the same structural model under the same (structural) shock with different parameters to assess the distribution of results. That can help to pinpoint parameters which strongly drive results (which explains the notion of “sensitivity”) but also to learn if the model’s structure with the underlying behavioral and technology assumptions or its parameterization ultimately determine the direction and size of simulated impacts. If the latter are not changing much under different set of parameters, the results can be termed robust. Note here the difference to an econometric exercise where the distribution of the error terms gives immediate information on how robust results are; in most cases, it is even possible to calculate confidence intervals for simulated results. That is generally not possible with large-scale simulation models, not at least as we have not overall knowledge about the “true” multi-dimensional distributions of the parameter space of the model. Still, it is reassuring to find that core results do not differ much even if parameters considered sensitive are changed.

GEMPACK and thus using runGTAP with GEMPACK upports systematic sensitivity analysis (Channing 1996<sup>129</sup>). CGEBox also offers similar functionality drawing on experiences with other GAMS-based models. In order to keep the technical solution relatively easy, the following approach is chosen:

1. The GAMS program which is used to run a single shock is also used for each single sensitivity experiment, in order to use as much as possible existing code and to avoid duplicating code fragments. Note that some parameter or shock changes will requiring a new benchmarking of the model which is supported by the code automatically, even if that drives up somewhat the overall time needed.
2. A second GAMS program “sensitivity.gms” acts as the master program. In combination with Python code, it defines the draws and executes in parallel GAMS children. It let each child solve the same (structural) shock. Each child uses its own parameter settings; changes in these settings are defined by the experiments.
3. This second GAMS mother program collects the results from the children and combines them into one GDX file which can be inspected later. It always reports the mean and standard deviations over the draws, but can also store the results for each draw which clearly requires from memory and disk space.

---

<sup>129</sup> Arndt, Channing. "An introduction to systematic sensitivity analysis via Gaussian quadrature." GTAP Technical Papers (1996): 2., <https://www.gtap.agecon.purdue.edu/resources/download/1209.pdf>



**Figure 25: Overview on sensitivity analysis**

Currently, the sensitivity analysis allows changing any parameters with up to three driving sets. The user can define for each set on each parameter if each element is subject to a specific random shock or if the same shock is applied to all elements on the set.

Whereas Arndt 1996 uses Gaussian Quadratures, the approach in here is based on Latin Hypercube Sampling (LHS). The user chooses the number of draws, and the LHS uses stratified random sampling for space filling design. The LHS is chosen as it can be easily applied also in case of non-symmetric, truncated distributions. In the default case, the LHS as a stratified random sampling mechanism simply returns (more or less) uniform distributed draws over the chosen potentially truncated PDF. The mean of the draws is generally close to the mode. The user should keep in mind that non-symmetric bounds will lead to a systematic deviation from the mean parameter used without sensitivity analysis.

Draws by LHS will generally recover the basic character of a distribution even with a limited number of draws such as ten. Ten draws will imply that for each parameter, the PDF of its distribution is sliced into ten bins, and one draw is made from each bin. If you draw from a uniform between 0.5 and 1.5, you will hence get one draw from the interval [0.5,0.6], one from [0.6,0.7] .... The GAMS code will scale the draws from the LHS such that the average draw for each factor is equal to its desired mean. The user has the choice between three different distributions, all truncated at the chosen lower and upper limit:

1. Uniform
2. Truncated Normal distribution
3. Triangular distribution

Equally, the user can determine if the draws are added to the base case or multiplied with it. The combination of user set truncation points and distributions allows for a rather flexible design of the experiments. If the lower is equal to the upper bound, the parameter will not be subject to sensitivity analysis. Clearly, at least for one parameter, the upper and lower bounds must differ. The following section details the options.

## 8.1 Controls for sensitivity experiments

The left-hand panel labelled “Draws” determines

- The **number of draws**. For each draw, a full model solve is performed, increasing the number of draws implies a higher sampling accuracy but also longer execution times. If the individual draws are not stored, more draws will not imply higher memory or disk space use for the final result.
- **Start number**: This will determine the start number for the experiments. For instance, if the start number of 2 is chosen, the first draw will receive the number 11 if 10 experiments are chosen. This allows to generate many samples of the same size and combines them into one large observation sample without the need to relabel individual draws.
- How many **GAMS child processes are used in parallel**. It is generally not recommended to use more parallel processes than cores on the computer. As CONOP4 solves in parallel, it might be even best to use only one or two parallel processes for models where the final solve of the full model takes a long time. For larger models, pre-solve can be used which in most cases will help to speed up the overall time especially in sensitivity analysis. Keep also in mind that modern processor layouts allow to temporarily increase the frequency of single cores which can speed up tremendously non-parallelized processing on multi-core machines. The higher frequency will heat up that specific core, therefore, the BIOS will automatically re-schedule threads to another core once a critical temperature is reached which requires to reduce its frequency. Forcing your computer to run many threads in parallel might hence mean that none of your core can boost its frequency. Here, careful listening to your computer might help: if your computer permanently makes a lot of fan noise, reduce the number of parallel threads. The least wanted outcome from a sensitivity analysis is a destroyed motherboard ... (that would give the term “robustness” a new meaning ...).
- The user might also choose “**Only generate draws**” to prevent the actual model runs, in order to solely check the layout of the scenarios. The file “LHS.lst” found in the scratch directory will show you detailed statistics such as the individual draws, the correlation between the factors etc.. The resulting draws for the LHS are also reported in the GAMS listing file (gams\sensitivity.lst): A look here might help to get an idea if the distributions are sensibly chosen, keeping in mind that it will not report the parameter values themselves, but either the multiplicative change (as below) or additive ones.

```
----- 5745 PARAMETER p_report draws
```

|          | sigmaW<br>Oceania<br>GrainsCrops-c | sigmaW<br>Oceania<br>MeatLstk-c | sigmaW<br>Oceania<br>Extraction-c | sigmaW<br>Oceania<br>ProcFood-c |
|----------|------------------------------------|---------------------------------|-----------------------------------|---------------------------------|
| draws001 | 1.00                               | 1.06                            | 0.96                              | 0.91                            |
| draws002 | 1.22                               | 0.96                            | 1.02                              | 1.04                            |
| draws003 | 0.97                               | 1.04                            | 0.85                              | 0.94                            |
| draws004 | 1.11                               | 0.99                            | 1.04                              | 0.98                            |
| draws005 | 0.96                               | 1.10                            | 1.16                              | 0.86                            |
| draws006 | 0.88                               | 0.93                            | 1.06                              | 1.23                            |
| draws007 | 1.03                               | 1.20                            | 1.09                              | 1.05                            |
| draws008 | 0.93                               | 1.01                            | 0.99                              | 1.00                            |
| draws009 | 1.07                               | 0.81                            | 0.88                              | 0.91                            |
| draws010 | 0.82                               | 0.90                            | 0.93                              | 1.08                            |

If you only want information on the mean and standard deviation of the results, keep “**Store each draw**” off. Otherwise, all results from all draws are collected in the GDX container. That allows you to generate, for instance, histograms or scatter plots as discussed below but can also lead to very large GDX containers. In doubt, run the base case with a normal simulation and open the resulting GDX container from the run under “*results\run*” with the GAM Studio or similar. Check the number of non-empty elements in the parameter *p\_results* and its size on disk.

| Entry | Symbol    | Type | Dim | Nr Elem |
|-------|-----------|------|-----|---------|
| 1     | p_results | Par  | 7   | 376,655 |

Multiply the numbers reported there by the number of draws to get an idea of the size of the resulting combined result set. Don’t count on being able to load more than say 10-20 Mio non-zeros into a viewer without manually changing run-time options for the Java run time engine or even running out of memory on your machine.

You might also consider switching some part of the post-model processing off to reduce the size of the result set. For instance, if you are not interested in the distribution of bi-lateral trade flows, tick off “Store bi-lateral results”; you can also reduce the size by only reporting tax rates, but not tax income for each item or you might not need the distributions of the welfare decomposition. Reducing post-model reporting clearly also speeds up running the individual experiments.

The **right-hand panel “Factors”** steers the the design of experiments for each parameter subject to sensitivity analysis. The standard example, delivered as the default with the GUI, changes the second level Armington elasticities *sigmaW* around their given levels, using multiplicative draws with a truncated normal distribution.

| Factors  |        |      |        |       |        |      |               |       |      |      |      |         |
|----------|--------|------|--------|-------|--------|------|---------------|-------|------|------|------|---------|
| Symbol   | 1. set | Mode | 2. Set | Mode  | 3. Set | Mode | Distribution  | Apply | Lo   | Up   | mode | Std.Dev |
| p_sigmaW | rNat   | same | i      | draws |        | same | Trunc. normal | Mult  | 0.50 | 1.50 | 1.00 | 0.10    |

The “**symbol**” field as the first text field defines the name of the GAMS parameter subject to sensitivity analysis, here *p\_sigmaW*. The three following text fields (**1. Set**, **2. Set**, **3. Set**) are the names of the sets of this symbol, in order they appear on the parameter. In our example, *p\_sigmaW* is defined over the nations *rNat* and the commodities *i* such that the third field has to be left empty.

If we want to restrict the draw to only some regions or commodities, we need to define sub-sets in our shock file and use it in the table above. If a parameter has less than three dimension, leave the last field



empty etc.. If a parameter has more than three dimensions, you need to use a helper parameter instead as discussed below.

For each set, we have the **mode** choice between “same” or “draws”, where “same” means that the same random shock being used for all elements in each experiment for this set, and “draws” means that each set elements receives its own shock in each experiment. If you use a 57x10 data base and you chose “draws” both for regions and commodities, you will hence have 570 independent factors in our LHS.

Let’s start with “same”, “same” which implies that we have only one factor level for each draw:

| Factors  |        |      |        |      |        |      |               |       |      |      |         |
|----------|--------|------|--------|------|--------|------|---------------|-------|------|------|---------|
| Symbol   | 1. set | Mode | 2. Set | Mode | 3. Set | Mode | Distribution  | Apply | Lo   | Up   | mode    |
| p_sigmaW | rNat   | same | i      | same |        | same | Trunc. normal | Mult  | 0.50 | 1.50 | 1.00    |
|          |        |      |        |      |        |      |               |       |      |      | Std.Dev |
|          |        |      |        |      |        |      |               |       |      |      | 0.10    |

The include file for a specific experiment will comprise statements as shown below, i.e. all 10x10 = 100 elements in p\_sigmaW will for one experiment (i.e. model solve) be changed by the same relative random change, i.e. the LHS will work with one factor only:

```
sigmaW('Oceania','GrainsCrops-c') = 1.0800361320*sigmaW('Oceania','GrainsCrops-c');
sigmaW('Oceania','MeatLstk-c') = 1.0800361320*sigmaW('Oceania','MeatLstk-c');
sigmaW('Oceania','Extraction-c') = 1.0800361320*sigmaW('Oceania','Extraction-c');
sigmaW('Oceania','ProcFood-c') = 1.0800361320*sigmaW('Oceania','ProcFood-c');
sigmaW('Oceania','TextWapp-c') = 1.0800361320*sigmaW('Oceania','TextWapp-c');
sigmaW('Oceania','LightMnfc-c') = 1.0800361320*sigmaW('Oceania','LightMnfc-c');
sigmaW('Oceania','HeavyMnfc-c') = 1.0800361320*sigmaW('Oceania','HeavyMnfc-c');
sigmaW('Oceania','Util_Cons-c') = 1.0800361320*sigmaW('Oceania','Util_Cons-c');
sigmaW('Oceania','TransComm-c') = 1.0800361320*sigmaW('Oceania','TransComm-c');
sigmaW('Oceania','OthServices-c') = 1.0800361320*sigmaW('Oceania','OthServices-c');
sigmaW('EastAsia','GrainsCrops-c') = 1.0800361320*sigmaW('EastAsia','GrainsCrops-c');
sigmaW('EastAsia','MeatLstk-c') = 1.0800361320*sigmaW('EastAsia','MeatLstk-c');
sigmaW('EastAsia','Extraction-c') = 1.0800361320*sigmaW('EastAsia','Extraction-c');
sigmaW('EastAsia','ProcFood-c') = 1.0800361320*sigmaW('EastAsia','ProcFood-c');
sigmaW('EastAsia','TextWapp-c') = 1.0800361320*sigmaW('EastAsia','TextWapp-c');
sigmaW('EastAsia','LightMnfc-c') = 1.0800361320*sigmaW('EastAsia','LightMnfc-c');
sigmaW('EastAsia','HeavyMnfc-c') = 1.0800361320*sigmaW('EastAsia','HeavyMnfc-c');
sigmaW('EastAsia','Util_Cons-c') = 1.0800361320*sigmaW('EastAsia','Util_Cons-c');
```

If we put instead “**draws**” on the region set rNat:

| Symbol | 1. set | Mode  | 2. Set | Mode | 3. Set |
|--------|--------|-------|--------|------|--------|
| sigmaW | rNat   | draws | i      | same |        |

We get for each experiment a different random change for each region as the LHS will work with 10 factors in our 10x10 example data set:

```

sigmaW('Oceania','GrainsCrops-c') = 1.0022894030*sigmaW('Oceania','GrainsCrops-c');
sigmaW('Oceania','MeatLstk-c') = 1.0022894030*sigmaW('Oceania','MeatLstk-c');
sigmaW('Oceania','Extraction-c') = 1.0022894030*sigmaW('Oceania','Extraction-c');
sigmaW('Oceania','ProcFood-c') = 1.0022894030*sigmaW('Oceania','ProcFood-c');
sigmaW('Oceania','TextWapp-c') = 1.0022894030*sigmaW('Oceania','TextWapp-c');
sigmaW('Oceania','LightMnfc-c') = 1.0022894030*sigmaW('Oceania','LightMnfc-c');
sigmaW('Oceania','HeavyMnfc-c') = 1.0022894030*sigmaW('Oceania','HeavyMnfc-c');
sigmaW('Oceania','Util_Cons-c') = 1.0022894030*sigmaW('Oceania','Util_Cons-c');
sigmaW('Oceania','TransComm-c') = 1.0022894030*sigmaW('Oceania','TransComm-c');
sigmaW('Oceania','OthServices-c') = 1.0022894030*sigmaW('Oceania','OthServices-c');
sigmaW('EastAsia','GrainsCrops-c') = 1.0563722580*sigmaW('EastAsia','GrainsCrops-c');
sigmaW('EastAsia','MeatLstk-c') = 1.0563722580*sigmaW('EastAsia','MeatLstk-c');
sigmaW('EastAsia','Extraction-c') = 1.0563722580*sigmaW('EastAsia','Extraction-c');
sigmaW('EastAsia','ProcFood-c') = 1.0563722580*sigmaW('EastAsia','ProcFood-c');
sigmaW('EastAsia','TextWapp-c') = 1.0563722580*sigmaW('EastAsia','TextWapp-c');
sigmaW('EastAsia','LightMnfc-c') = 1.0563722580*sigmaW('EastAsia','LightMnfc-c');
sigmaW('EastAsia','HeavyMnfc-c') = 1.0563722580*sigmaW('EastAsia','HeavyMnfc-c');
sigmaW('EastAsia','Util_Cons-c') = 1.0563722580*sigmaW('EastAsia','Util_Cons-c');

```

And finally, choosing “**draws**” for both sets will generate 10 regions x10 commodities =100 factors

| Symbol | 1. set | Mode    | 2. Set | Mode    | 3 |
|--------|--------|---------|--------|---------|---|
| sigmaW | rNat   | draws ▾ | i      | draws ▾ | : |

```

sigmaW('Oceania','GrainsCrops-c') = 1.1184663500*sigmaW('Oceania','GrainsCrops-c');
sigmaW('Oceania','MeatLstk-c') = 0.9331679670*sigmaW('Oceania','MeatLstk-c');
sigmaW('Oceania','Extraction-c') = 0.9607824656*sigmaW('Oceania','Extraction-c');
sigmaW('Oceania','ProcFood-c') = 0.9866457350*sigmaW('Oceania','ProcFood-c');
sigmaW('Oceania','TextWapp-c') = 0.9877797127*sigmaW('Oceania','TextWapp-c');
sigmaW('Oceania','LightMnfc-c') = 1.0682751920*sigmaW('Oceania','LightMnfc-c');
sigmaW('Oceania','HeavyMnfc-c') = 1.0975363260*sigmaW('Oceania','HeavyMnfc-c');
sigmaW('Oceania','Util_Cons-c') = 0.9429919894*sigmaW('Oceania','Util_Cons-c');
sigmaW('Oceania','TransComm-c') = 0.9942992523*sigmaW('Oceania','TransComm-c');
sigmaW('Oceania','OthServices-c') = 0.9188807855*sigmaW('Oceania','OthServices-c');
sigmaW('EastAsia','GrainsCrops-c') = 0.9263090962*sigmaW('EastAsia','GrainsCrops-c');
sigmaW('EastAsia','MeatLstk-c') = 1.0423022530*sigmaW('EastAsia','MeatLstk-c');

```

If you want to work, for instance, with a variable or parameter which has more then three dimensions, you can simply type in a new symbol name:

|         |      |        |   |        |       |        |    |
|---------|------|--------|---|--------|-------|--------|----|
| p_shock | rNat | same ▾ | i | same ▾ | rnat1 | same ▾ | Ti |
|---------|------|--------|---|--------|-------|--------|----|

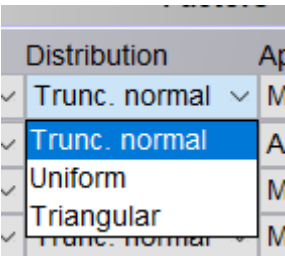
The program code will put random draws according to your desired layout in a run specific include file; values which you can pick up in your shock file.

```

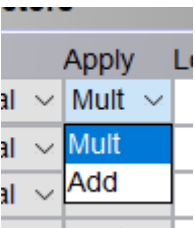
-----
parameter p_shock;
p_shock('Oceania','GrainsCrops-c','Oceania') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','EastAsia') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','SEAsia') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','SouthAsia') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','NAmerica') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','LatinAmer') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','EU_25') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','MENA') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','SSA') = 0.7841810194;
p_shock('Oceania','GrainsCrops-c','RestofWorld') = 0.7841810194;
p_shock('Oceania','MeatLstk-c','Oceania') = 0.7841810194;
p_shock('Oceania','MeatLstk-c','EastAsia') = 0.7841810194;

```

Next, we detail how to define the individual draws. For each symbol subject to sensitivity analysis, we have the choice between three **distributions**:



Next, we can determine if the random draws are multiplied with the base case of the parameter or added to it (that choice is only relevant if the parameter is already declared such that are values to add something to or multiply with, otherwise, the draws are stored as shown for *p\_shock* above):



For a multiplicate draw, the range of the draws defined up “*lo*” and “*up*” should set around unity which the “*mode*” close to unity. For additive draws, the “*lo*” and “*up*” should be around zero and the mode close to zero. The “Std.Dev” field is only relevant for the truncated normal. The mode and std.dev. in case of a truncated normal refer to the underlying not-truncated normal distribution.

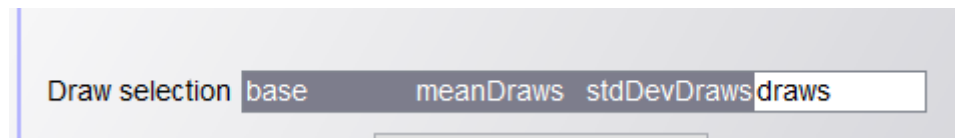
| Apply  | Lo    | Up   | mode | Std.Dev |
|--------|-------|------|------|---------|
| ✓ Mult | 0.50  | 1.50 | 1.00 | 0.10    |
| ✓ Add  | -0.50 | 0.50 | 0    | 0.10    |

Note that bounds which are not symmetric around a mode of unity for multiplicative or a mode of zero in case of additive draws will imply a systematic difference between the paramters of the base case and the mean of the draws for each parameter (besides the fact that the mode of the draws will differ from the parameter inputted).

8.2 Parallel solves and analyzing the results

The actual solve for each individual experiment starts from a savepoint generated after loading the data base, such that execution time is only spend on benchmarking, model generation, solving and post-model reporting. Fifty draws of the 10x10 model under a 50% multi-lateral trade simulation experiment took about 30 seconds to solve on my rather fast laptop, another 5 seconds are spent to collect and combine the results. 20 draws with a 57x10 model under the same shock took about two minutes, with the 57x24 model, time goes up to about 6 minutes.

When you inspect the results, the interface will allow showing either all draws and/or the standard case, the average of the draws and the standard deviation:



If only the mean over the draws is selected, the result tables look almost identical to a single simulation run. The tables will report two outcomes: (1) the mean of the draws labeled “meanDraws” and the (2) result at the original parameters “base” – what one would simulate without any sensitivity analysis and (3) the standard deviation over the draws:

| Demand, by product [1] |             | Items                               |
|------------------------|-------------|-------------------------------------|
| Regions                | World       | Quantity                            |
|                        |             | 10x10_example_Policy_Shocks_Tariffs |
|                        |             | Total Household demar               |
| All commodities        | base        | 124270.91                           |
|                        | meanDraws   | 124270.88                           |
|                        | stdDevDraws | 0.65                                |

If the draws are selected, you get views such as below:

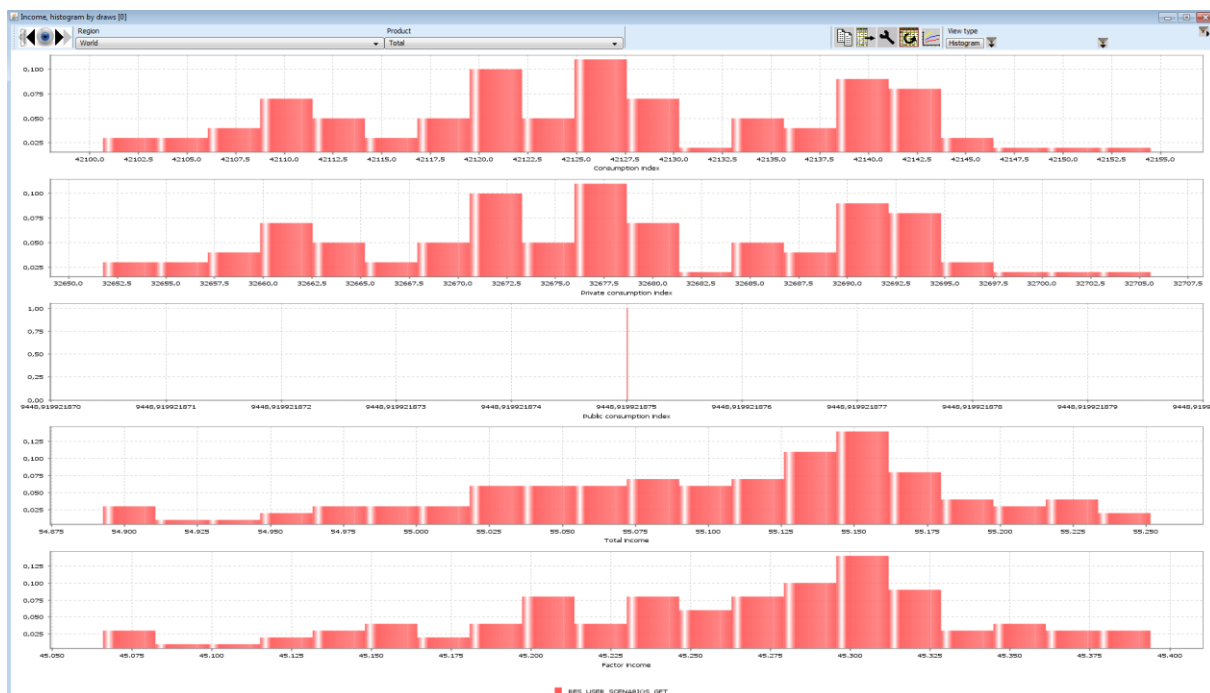
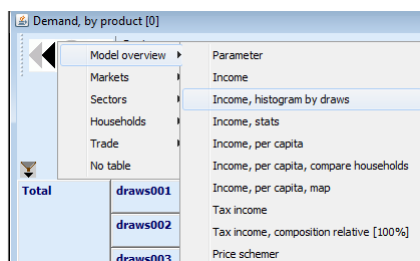
| Income [0]                |                        | Product       |
|---------------------------|------------------------|---------------|
| RES_BASE_SCENARIOS_FUTURE |                        |               |
|                           | Australia, New Zealand | East Asia     |
|                           | Southeast Asia         | South Asia    |
|                           | North America          | Latin America |
| Private consumption index | draws001               | 797.37        |
|                           | draws002               | 778.09        |
|                           | draws003               | 810.42        |
|                           | draws004               | 603.39        |
|                           | draws005               | 556.75        |
|                           | draws006               | 71.61         |
|                           | draws007               | 849.09        |
|                           | draws008               | 903.13        |
|                           | draws009               | 795.44        |
|                           | draws010               | 778.99        |
|                           | draws011               | 886.35        |
|                           | draws012               | 851.03        |
|                           | draws013               | 800.89        |
|                           | draws014               | 801.67        |
|                           | draws015               | 851.37        |
|                           | draws016               | 789.14        |
|                           | draws017               | 806.36        |
|                           | draws018               | 988.29        |
|                           | draws019               | 866.05        |
|                           | draws020               | 800.01        |
|                           | draws021               | 858.21        |
|                           | draws022               | 853.76        |
|                           | draws023               | 953.02        |
|                           | draws024               | 807.46        |
|                           | draws025               | 871.74        |
|                           | draws026               | 809.68        |
|                           | draws027               | 859.39        |
|                           | draws028               | 862.80        |
|                           | draws029               | 6656.11       |

That allows tyour to analyze each draw, if wished. Alternatively, the statistics built in the interface can be used. In order to do this, open with a right click in the table, the pop-up menu and select statistics:

Select the statistics you want to be shown (such as mean and stdDev) and select “Only show outliers” to only show outliers and statistics:

| Income [0]          |        | RES_BASE_SCENARIOS_FUTURE |           |                |            |
|---------------------|--------|---------------------------|-----------|----------------|------------|
|                     |        | Product                   |           |                |            |
|                     |        | Total                     |           |                |            |
|                     |        | RES_BASE_SCENARIOS_FUTURE |           |                |            |
|                     |        | Australia, New Zealand    | East Asia | Southeast Asia | South Asia |
| Consumption index   | Mean   | 1064.58                   | 8790.18   | 1267.71        |            |
|                     | StdDev | 144.51                    | 1437.99   | 340.46         |            |
| Private consumption | Mean   | 788.21                    | 6588.45   | 1083.75        |            |
|                     | StdDev | 144.51                    | 1437.99   | 340.46         |            |
| Public consumption  | Mean   | 276.37                    | 2201.73   | 183.95         |            |
|                     | StdDev | Nan                       | 2.07      | Nan            |            |
| Total income        | Mean   | 1011.97                   | 9587.35   | 1547.67        |            |
|                     | StdDev | 706.44                    | 9601.15   | 1457.58        |            |
| Factor income       | Mean   | 830.71                    | 8670.79   | 1540.18        |            |
|                     | StdDev | 600.08                    | 8442.18   | 1356.41        |            |
| Current account     | Mean   | 21.77                     | -351.98   | -104.45        |            |
|                     | StdDev | 2.74                      | 44.34     | 13.16          |            |
| Tax income          | Mean   | 159.50                    | 1268.54   | 111.94         |            |
|                     | StdDev | 106.74                    | 1139.17   | 100.42         |            |
| Population size     | Mean   | 52.11                     | 2326.15   | 851.49         |            |
|                     | StdDev | 0.04                      | Nan       | Nan            |            |

A view with the core income results is defined which generates histograms:



Note: If several scenarios are analyzed, the distributions for several scenarios will be shown together.

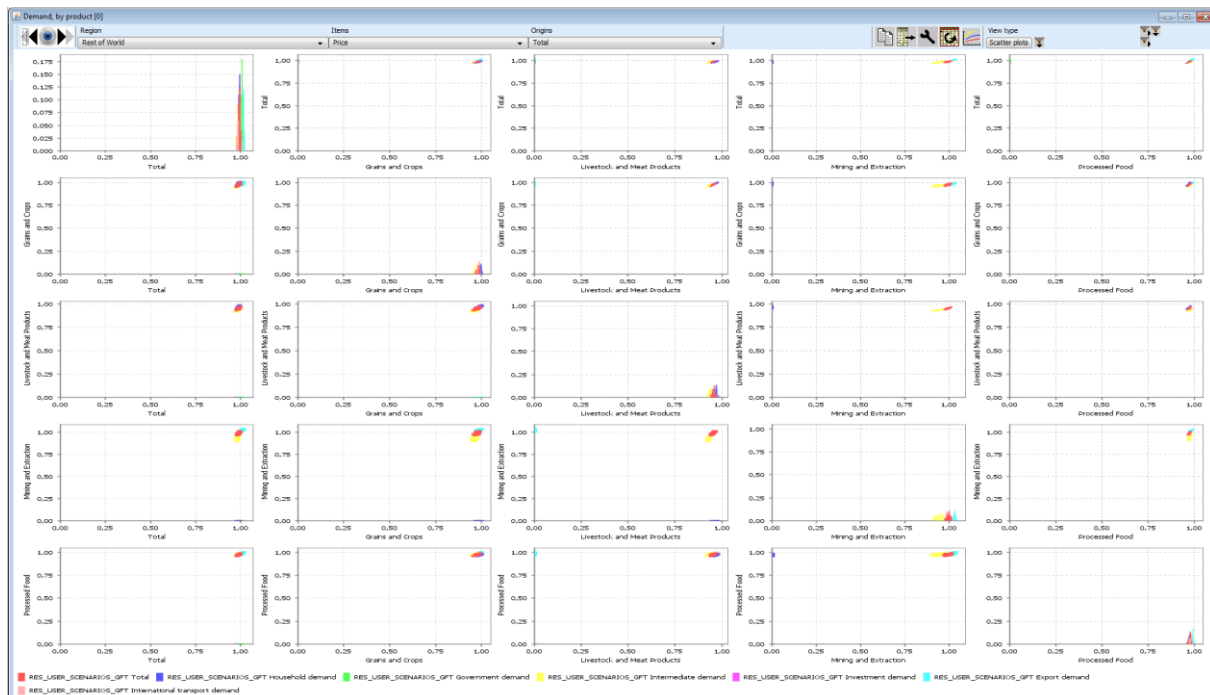
Alternatively, the “Scatter plot” type graphics can be used to plot additionally the correlation between the variables depicted:



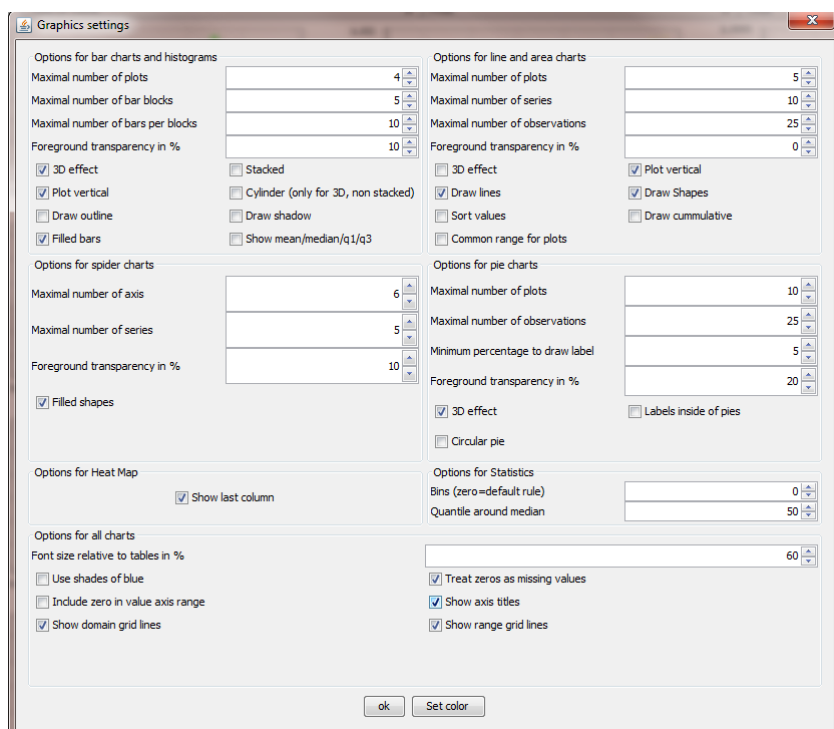
## Parallel solves and analyzing the results



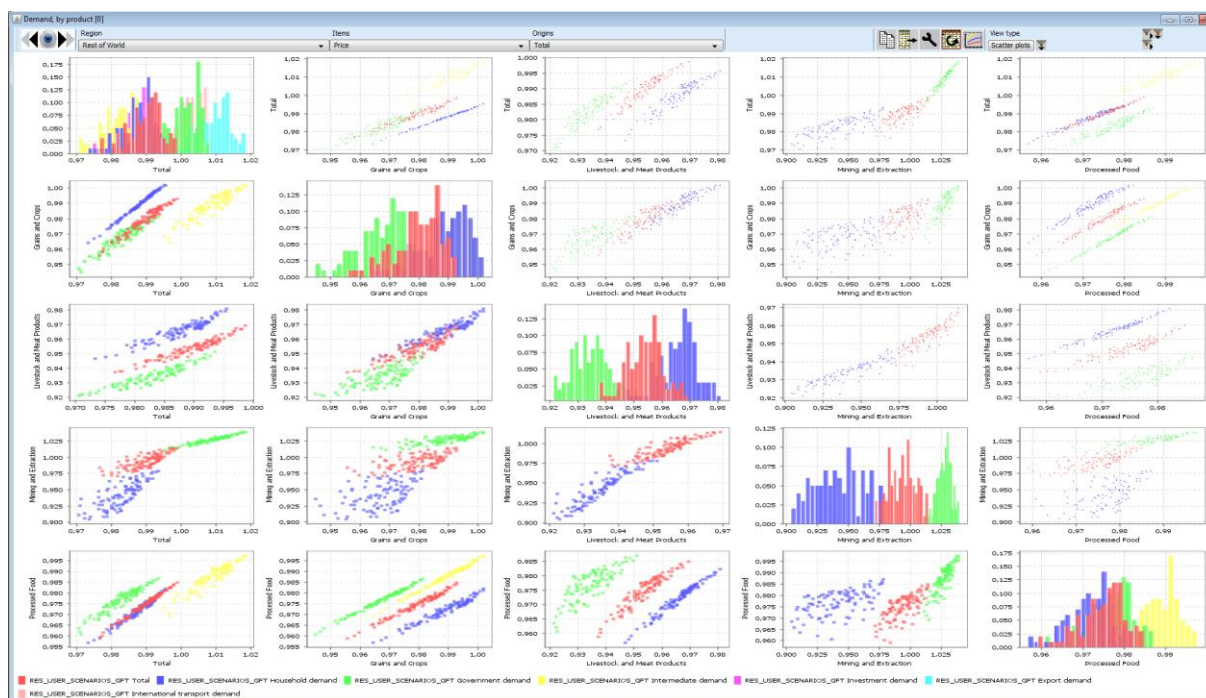
If the view type from existing tables is changed to “Scatter plots”, the result might look as below (the example shows different types of prices in each plot, and different sectors in the rows/columns):



The visualization of histograms and scatter plots can be improved, if the graphics settings are changed such that (1) zeros are treated as missing values, and (2) zero is not automatically included in the axis range:



The resulting scatter plots and histograms look as follows:



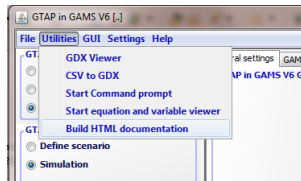
## 9. Getting an overview on the code

The GUI offers two possibilities to ease working with the code:

1. Generation of a set of static HTML pages
2. An equation, parameter and variable viewer linked to a specific model instance, which also allows looking at the GAMS code.

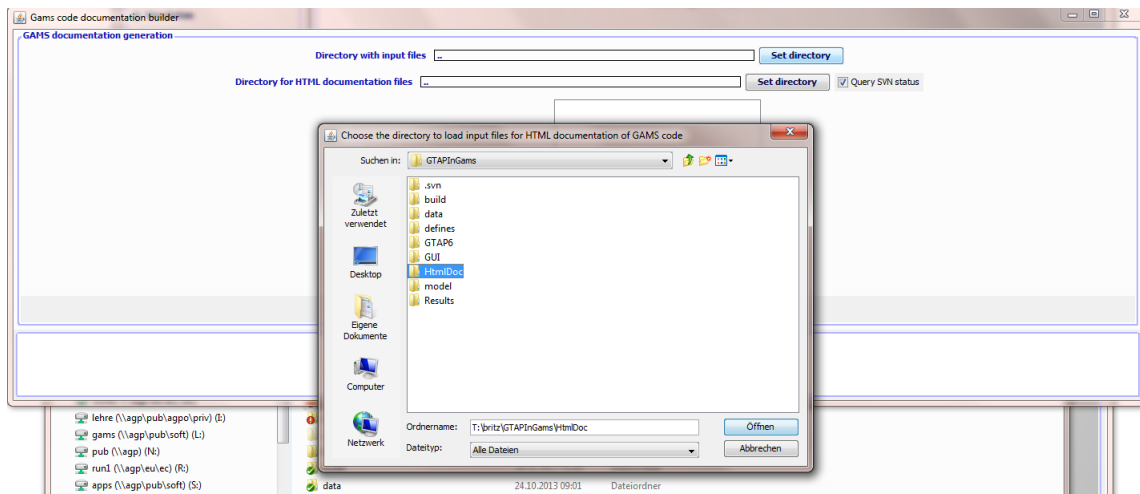
### 9.1 HTML documentation

A HTML documentation of the code can be generated via “**Utilities/Build HTML documentation**”:



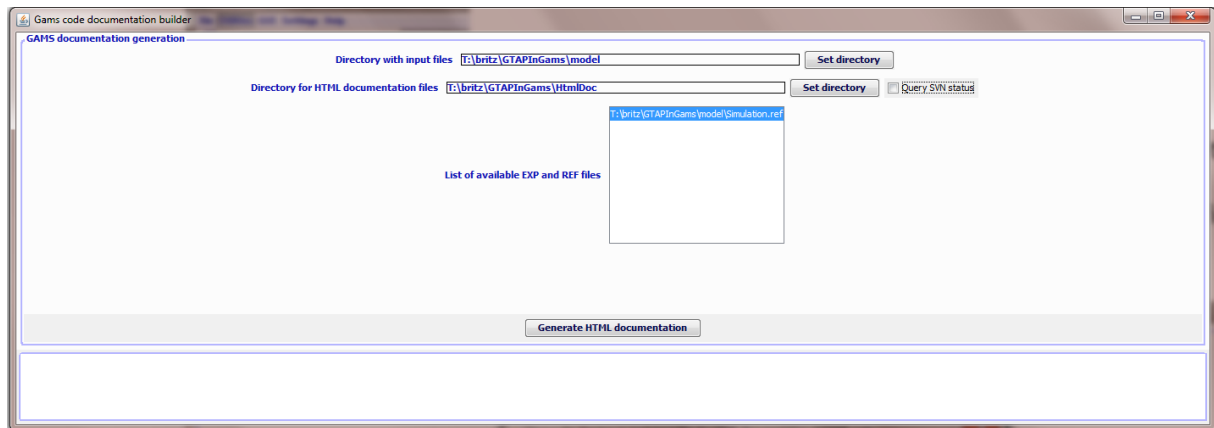
The documentation will include any changes you introduced to the project’s GAMS code.

First choose the directory where the files should be stored (best use the “HtmlDoc” directory as shown below; beware: the utility will generate many files, do not use a directory where other files are already stored). The input files are found in the folder where GAMS was run, in the case of standard GTAP model in GAMS, typically in the “model” directory.



You should be able to find a file named “simulation.ref”. Make sure that “Query SVN Status” is switched off and next press the “Generate HTML documentation” button.

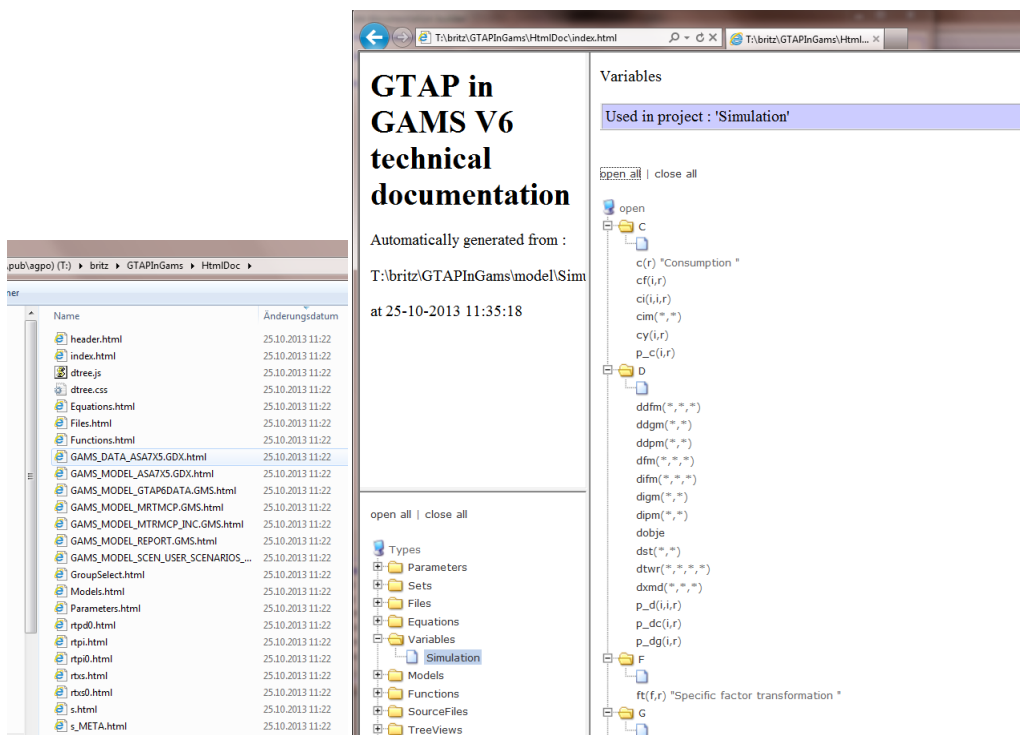




The program will work for a while and should end with “HTML documentation is ready” as shown below.

```
HTML documentation is ready
parsing T:\britz\GTAPInGams\model\mrtmcp.gms done
parsing T:\britz\GTAPInGams\model\report.gms done
parsing T:\britz\GTAPInGams\model\scen\user_scenarios\noshock.gm
Generate HTML page for Gams symbols
Generate HTML page for Gams source files
Generate HTML pages for project : Simulation
HTML pages are generated
```

Afterwards, you find in the chosen output directory a list of files, “index.html” is the starting point.



You can e.g. open all “Variables” used in the task “simulation”, clicking with the mouse opens a symbol page:

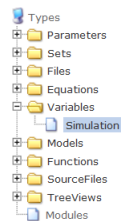
## GTAP in GAMS V6 technical documentation

Automatically generated from :

T:\britz\GTAPInGams\model\Sim

at 25-10-2013 11:35:18

open all | close all



### Variable ddfm

|                      |              |
|----------------------|--------------|
| Name                 | ddfmm(*,*,*) |
| Type                 | Variable     |
| Number of dimensions | 3            |
| Used in project(s)   | Simulation   |

ddfmm is declared in:

[Top](#) | [Declarations](#) | [Definitions](#) | [Assignments](#) | [References](#) | [Elements](#)

File

Projects

[GAMS\MODEL\MRTMCP.GMS](#) Simulation [Edit](#)

ddfmm is assigned in:

[Top](#) | [Declarations](#) | [Definitions](#) | [Assignments](#) | [References](#) | [Elements](#)

File

Projects

[GAMS\MODEL\MRTMCP.GMS](#) Simulation

ddfmm is referenced in:

[Top](#) | [Declarations](#) | [Definitions](#) | [Assignments](#) | [References](#) | [Elements](#)

File

Projects

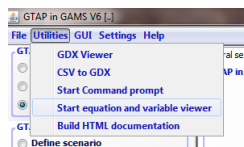
[GAMS\MODEL\MRTMCP.GMS](#) Simulation

[GAMS\MODEL\REPORT.GMS](#) Simulation

[Top](#) | [Declarations](#) | [Definitions](#) | [Assignments](#) | [References](#) | [Elements](#)

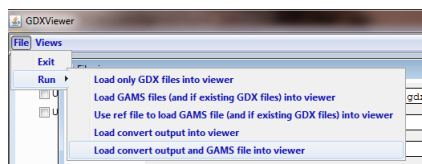
## 9.2 Equation and variable viewer (similar AnalyzeGE)

Whereas the HTML pages document the project, the equation and variable viewer helps you to analyze a specific model instance. The necessary input files are automatically generated with each run in comparative static mode if “CONVERT” output is chosen.

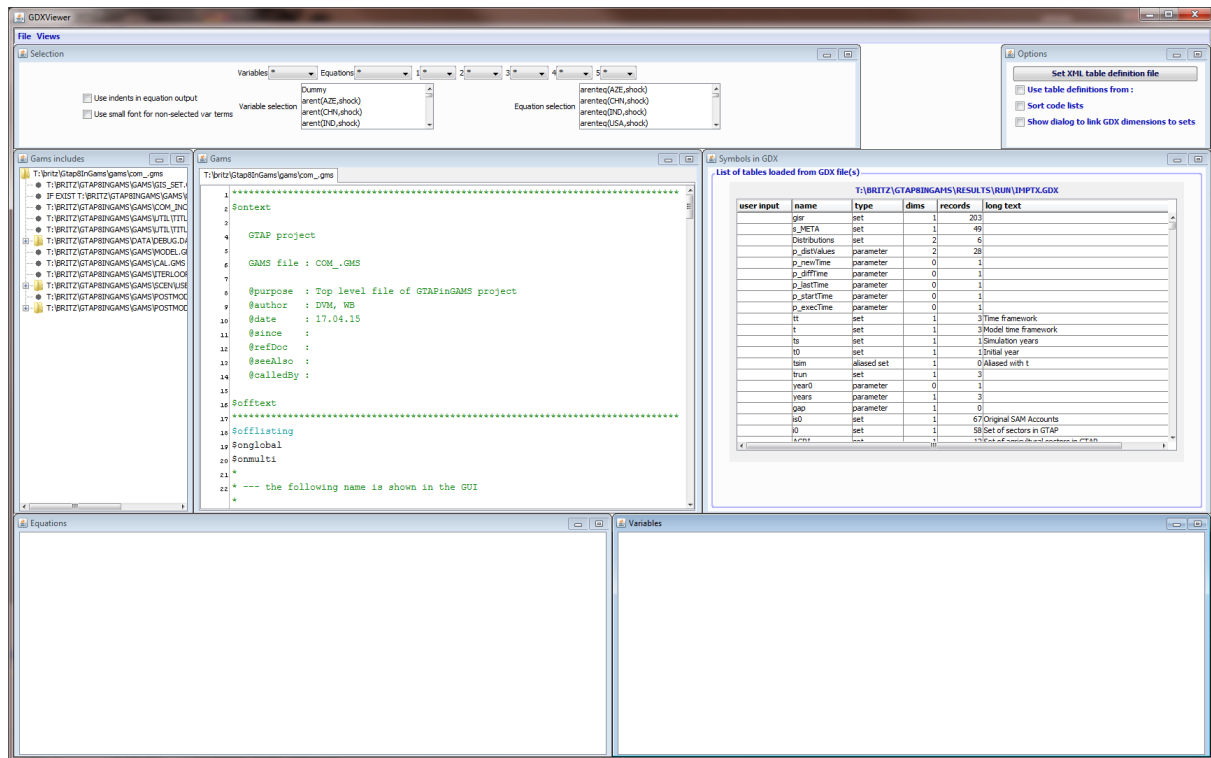


(note: one might want to change the first GDX files to the last experiment you have run)

Select “Load convert output and GAMS file into viewer” from the file menu:



You get now view like below:



The viewer allows you to see:

- Linearized views on the equations, which you can select via the selection button above. The information in green behind each equation provides, separated by commata:
  1. The absolute change in the variable
  2. The absolute change multiplied with the Jacobian, i.e. the approximate change in the equation due to the simulated change
  3. The contribution of the variable to the change in the LHS in percent, using information in the absolute changes in the LHS and the variable and the related Jacobian entries. The variable which the code analyzer has detected as the LHS is indicated with a \* - in the example below that is xd. For the LHS, the relative change is report. In example below, the change is almost entirely coming from the change in nd (100%), the contribution of the other changes is very small.

```
xdeg(SouthAsia,MeatLstk-c,ProcFood-a,shock) "Agents demand for domestic goods " ..
-0.00902671001153436*(0.9999999486802*(0.995354810037336*(1.0000000513198*ps(SouthAsia,MeatLstk-c,shock)
{-0.005,-7.2E-5,1.0%})**(-2.13050842285156)
+0.00464540767085857*(1.00000000249911*pmt(SouthAsia,MeatLstk-c,shock){0.0033,-4.7E-5,0.8%})**(-2.130
50842285156))**(-0.469371530886303)/ps(SouthAsia,MeatLstk-c,shock){-0.005,-7.2E-5,1.0%})**3.130508422
85156*nd(SouthAsia,ProcFood-a,shock){0.67,-0.006,100.0%}
+0.183822553739227*xd(SouthAsia,MeatLstk-c,ProcFood-a,shock){0.034,0.0062,0.6%*})=E=0;
```

- A GDX viewer which allows merging of symbols (from one or several files) for combined analysis.
- A quick view on any symbol: click on any symbol in the GAMS code shown in blue, and it will be loaded in the “Symbol from GDX” view. You can directly compare your counterfactual against the base (or any scenarios against each other which you loaded as GDX file).

## Equation and variable viewer (similar AnalyzeGE)

Gams

T:\britz\GTAPInGams\model\vrmtcp.gms

```

62: *
63: *   Profit function for firms:
64: *
65: *   Value shares for firm inputs of factors and goods:
66: *
67: parameter      thetad(f,j,z)   Factor share of value added,
68:                  thetad(i,j,z)   Domestic share of intermediate input,
69:                  thetai(i,j,z)   Import share of intermediate input,
70:                  theta_f(j,z)    Value added share of sectoral output;
71:
72: alias (f,ff);
73: thetad(f,j,z)$sum(ff,vfm(ff,j,z)*(1+rtf0(ff,j,z)))
74:   = vfm(f,j,z)*(1+rtf0(f,j,z)) / sum(ff,vfm(ff,j,z)*(1+rtf0(ff,j,z)));
75: thetad(i,j,z)$ (vdfm(i,j,z)*(1+rtfd0(i,j,z)) + vifm(i,j,z)*(1+rtfi0(i,j,z)))
76:   = vdfm(i,j,z)*(1+rtfd0(i,j,z)) /
77:     (vdfm(i,j,z)*(1+rtfd0(i,j,z)) + vifm(i,j,z)*(1+rtfi0(i,j,z)));
78:
79: thetai(i,j,z)$vom(j,z)
80:   = (vdfm(i,j,z)*(1+rtfd0(i,j,z)) + vifm(i,j,z)*(1+rtfi0(i,j,z))) / vom(j,
81: z);
82: theta_f(j,z)$vom(j,z) = sum(ff,vfm(ff,j,z)*(1+rtf0(ff,j,z))) / vom(j,z);
83: *
84: *   User input: define for factors, domestic and imported

```

Symbol from GDX

Dim2 (noShock.gdx,GFT.gdx,...)

noShock.gdx

View type

Table

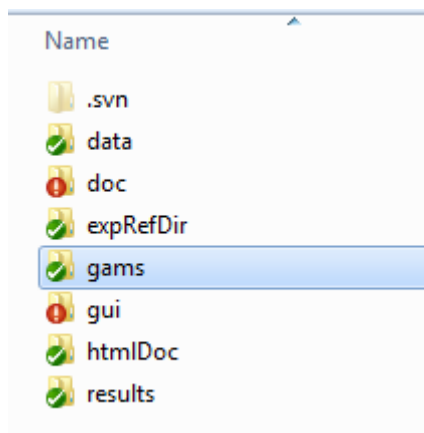
|     | SAF  | RSA  | RSS  | EUR  | ROW  |  |
|-----|------|------|------|------|------|--|
| agr | 0.37 | 0.55 | 0.68 | 0.45 | 0.51 |  |
| foo | 0.28 | 0.39 | 0.29 | 0.32 | 0.29 |  |
| ext | 0.32 | 0.56 | 0.55 | 0.31 | 0.34 |  |
| lgt | 0.38 | 0.38 | 0.36 | 0.34 | 0.34 |  |
| hvy | 0.29 | 0.25 | 0.22 | 0.21 | 0.24 |  |
| tch | 0.36 | 0.41 | 0.33 | 0.34 | 0.33 |  |
| ser | 0.48 | 0.46 | 0.41 | 0.49 | 0.50 |  |

## 10. Technical documentation

### 10.1 Overview on directory structure

The screenshot below shows the directory structure of a “CGEBox” installation. The information is only provided for those who have an interest into technical detail.

The GUI with the Java libraries is found in GUI. The results directory comprises the output from model runs. The data directory comprises also the GAMS readable versions of available data bases. All GAMS files can be found under “gams”, in the subdirectory “scen”, scenario templates and user defined scenarios are stored. Results of model runs are found under “results/run”. The directory “expRefDir” and “HtmlDoc” are used to generate a HTML based documentation of the GAMS project.

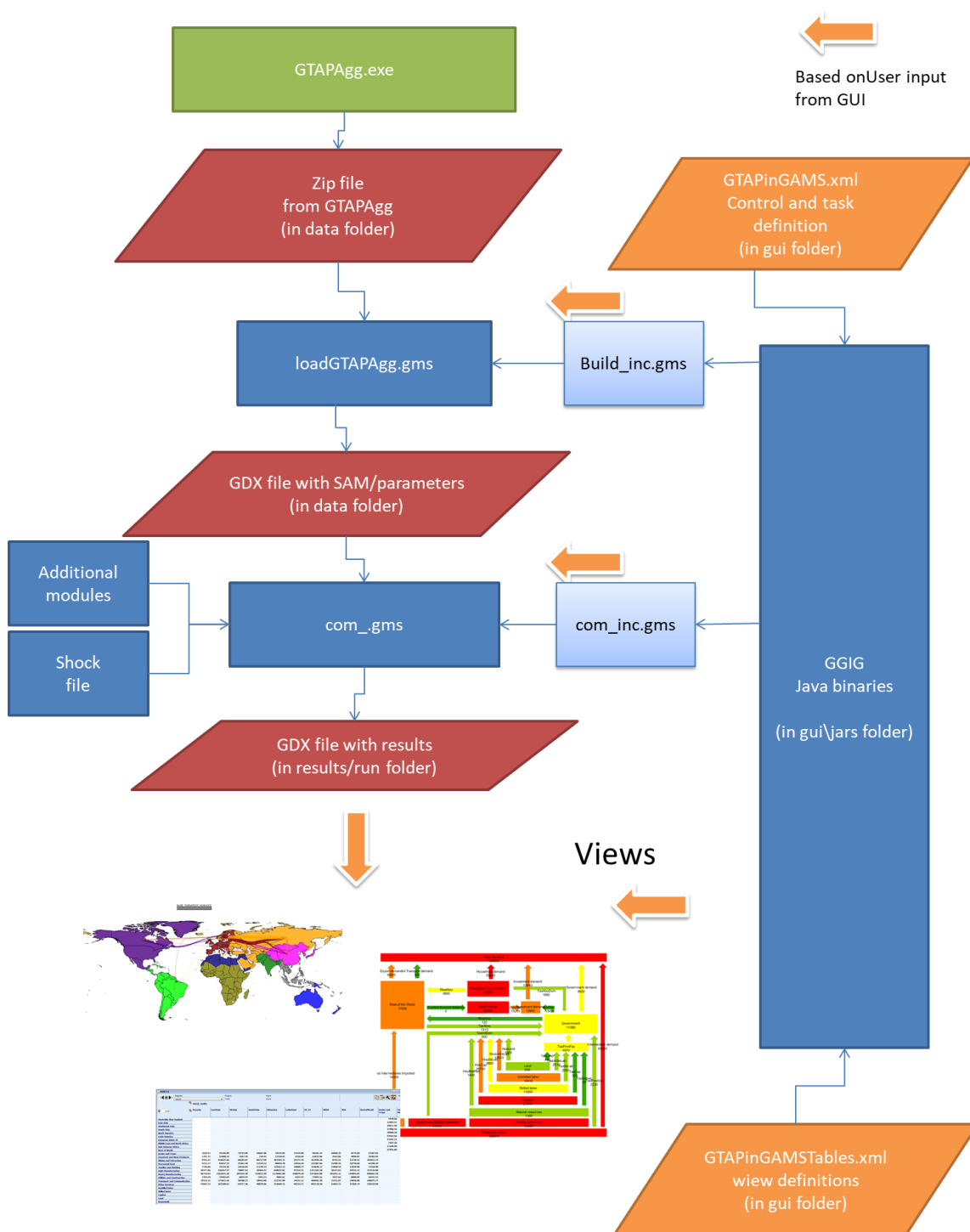


### 10.2 GUI overview

The GUI is programmed in Java and based on GGIG ([Gams Graphical Interface Generator](http://www.gams.com/docmain/reference/ggig/ggig.html)), which is also used for other projects, e.g. CAPRI ([www.capri-model.org](http://www.capri-model.org)). While binaries are distributed for free, the Java code itself is currently not open source. It uses third-party libraries provided under licenses which allow for further distribution, and often as well code changes. The main jar to start is “jars\ggig.jar”. In order to modify the GUI, no Java programming is necessary, as the Java code reads an XML file (“gui\cgebox.xml”) which defines tasks, work steps and the individual GUI controls. Equally, an XML file (“gui\cgebox\_tables.xml”) defines the different views. The link to GDX files is based on the Java API distributed by GAMS.com as part of any GAMS installation, and uses dynamic link libraries. As both Java and GAMS are available also for non-Windows operation systems, it would be possible to also port the standard GTAP model in GAMS to other software platforms. Indeed, students have worked with CGEBox under native MAC.

The interaction between the GUI and the GAMS processes is based on includes files which are generated anew for each run and capture the state of the user operated control.

### Graphic: Overview on interaction between user, GUI and GAMS



### 10.3 Processes

In order to work with the model, a zip files with GDX container needs to be downloaded from the GTAP center's website and converted with

Next, the "GtapAgg.gms" programs reads the GAMS symbols in the different containers and aggregates them to the user defined regions, products, sectors and factors. On demand, GTAPAgg introduces manual correction from user provided files, uses a filter processes to remove small values

and rebalance the global data, adds gender differentiated labor data and applies a SAM split. The result from these steps is stored in the “data” directory as on GDX container.

The GDX container generated by “GTAPAgg.gms” provides the input to actual model runs with “com\_.gms”. The GAMS code initializes all variables to benchmark values and starts the solver. On demand, it will introduce multiple households based on user provided data. The benchmarking will consider the modular set-up. That model start should basically need no iterations and lead to zero infeasibilities, it should prove that the model’s parameters are correctly set up to replicate the benchmark. Next, the user choasen shock file is read and changes to parameters introduced. Afterwards, the model is solved for the shock.

## 10.4 Post-model processing

While the processes described above are more or less identical to the original code of Dominique and Van der Mensbrugghe and Tom Rutherford, the post-model processing was added to allow using the exploitation part of GGIG. The file “report.gms” stores the results back into a SAM like structure, e.g. consumer prices

```
p_results(r,"P",i,"hou","dom") = p_dc.l(i,r);
```

The “p\_results” parameter is stored in a gdx file and later read by the interface for exploitation. It is generally organized as follows:

- First dimension: regions, including a world aggregated stored under “wor”
- Second dimension: tables for values “V”, quantities “Q”, prices “P”, tax rate “T” and tax income “G”
- Third dimension: commodities, factors
- Fourth dimension: institutions, sectors
- Fifth dimension: origins / destinations
- Sixth dimension: base, check or shock

Based on the flexible aggregation used in GTAP, the sets used for regions, sectors and factors are run specific and depend on the data set loaded.

The information about how the output is logically structured is also inputted in the GGIG definition file “gtapInGams.xml”:

```
<task>
  <name>Simulation</name>
  <gamsFile>mrtmcp</gamsFile>
  <incFile>model\mtrmcp_inc</incFile>
  <curDir>model</curDir>

  <regionDim>8</regionDim>
  <dim5Dim>1,Items</dim5Dim>
  <productDim>2</productDim>
  <activityDim>3,Sectors and institutions</activityDim>
  <dim6Dim>4,Origins</dim6Dim>
  <scenDim>5,Scenarios</scenDim>

  <resdir>run</resdir>
  <gdxSymbol>p_results</gdxSymbol>
  <Filemask>.*gdx$</filemask>
</task>
```

The “report.gms” program also performs aggregations, e.g. to total use

```

*
* --- aggregate over origins (domestic and imported) to totals
*
set agg(PQU) / Q,U,G /;

*
* --- aggregate to total use
*

set useCols / set.i,hou,gov,intTrs,set.s,inv /;

p_results(r,agg,i,"use",oris) = sum(useCols, p_results(r,agg,i,useCols,oris));

```

Average prices and tax rates are calculated afterwards:

```

p_results(r,"P",i,"use",oris) $ p_results(r,"Q",i,"use",oris)
= p_results(r,"U",i,"use",oris) / p_results(r,"Q",i,"use",oris);

p_results(r,"T",i,"use",oris) $ p_results(r,"U",i,"use",oris)
= p_results(r,"G",i,"use",oris) / p_results(r,"U",i,"use",oris);

```

The “report.gms” programs also stores meta-information on the run on the cube, such as the number of sectors, regions and factors; model size and solution status, and about how factor mobility is modeled:

```

* -----
*
* Meta information on model structure and solve
*
* -----
*

p_results("Wor","Meta","# of sectors","tots","dom") = card(i);
p_results("Wor","Meta","# of factors","tots","dom") = card(f);
p_results("Wor","Meta","# of regions","tots","dom") = card(r);

acronym MCP,CNS;
$iftheni %modeltype% == MCP

p_results("Wor","Meta","Model type","tots","dom") = MCP;
$else

p_results("Wor","Meta","Model type","tots","dom") = CNS;
$endif

$iftheni NOT %modeltype% == NONE

p_results("Wor","Meta","# of equations","tots","dom") = %model%.numequ;
p_results("Wor","Meta","# of variables","tots","dom") = %model%.numvar;
p_results("Wor","Meta","SumInfes","tots","dom") = %model%.sumInfes $ (%model%.sumInfes ne inf);
p_results("Wor","Meta","NumInfes","tots","dom") = %model%.numInfes $ (%model%.numInfes ne inf);
p_results("Wor","Meta","# Iterations","tots","dom") = %model%.iterusd;
p_results("Wor","Meta","# seconds solution time","tots","dom") = %model%.resusd;

acronym sluggish, mobile, fixed priced;

p_results("Wor","Meta","Lab Market","tots","dom") = %labmarket% ;
p_results("Wor","Meta","skl Market","tots","dom") = %sklMarket% ;
p_results("Wor","Meta","capital Market","tots","dom") = %capitalMarket% ;
p_results("Wor","Meta","res Market","tots","dom") = %resMarket% ;
p_results("Wor","Meta","lnd Market","tots","dom") = %lndMarket% ;

$endif

```

That approach differs considerably from the way “runGTAP” allows to exploit results. For a formal discussion on these differences, see Britz 2014 and Britz et al. 2014. The “equation and variable” viewer, discussed above, allows views on the variables, equations and related parameters, more similar to the “runGTAP” exploitation tools.

## 10.5 Substitution of variables based on macros

With highly detailed SAMs, the number of variables and equations could reach several millions. Beside filtering, the possible to aggregate the Armington first nests, variable substitutions are used to reduce model size. We will in here briefly discuss some of these macros to ease reading the code.



Generally, macros are named as the related variables with the prefix “m\_”. The macro m\_pa hence defines the variable p\_a in the model, i.e. the Armington prices for the agents. As these is a complex case, it will be used here as the example.

```
*
* --- same as above, difference is each sector has its own Armington nest for intermediate demand
*
$$macro m_pa(r,i,aa,t) { [ pa(r,i,aa,t) $ (not a(aa)) \
+ m_padef(r,i,aa,t) $ a(aa) ] $ (alphan(r,i,aa,t) $ alphad(r,i,aa,t)) \
+ (m_pdp(r,i,aa,t)*alphad(r,i,aa,t)) $ ((not alphan(r,i,aa,t)) $ alphad(r,i,aa,t)) \
+ (m_pmp(r,i,aa,t)*alphan(r,i,aa,t)) $ ((not alphad(r,i,aa,t)) $ alphan(r,i,aa,t)) \
+ pa(r,i,aa,t) $ ( not (alphan(r,i,aa,t)+alphad(r,i,aa,t)) ) \
}
```

The macro considers four different cases:

- (1) Both domestic and import demand is present (alphan and alphad is given). In that case, for agents which are not production sectors, the variable pa itself is used and defined in an equation (first line). Otherwise, another macro m\_padef is used.
- (2) Only domestic demand is present. In that case, the price is equal to the domestic price, defined by the macro m\_pdp.
- (3) Only import demand is present, in which case only the import price defined by the macro m\_pmp is used.
- (4) No import and domestic shares are defined, a case found when the Melitz model is used.

The macro m\_padef is the actual dual price aggregator:

```
*
* ---- definition of Armington dual price aggregator from domestic and import price
*
$$macro m_padef(r,i,aa,t) { ( (alphad(r,i,aa,t)*m_pdp(r,i,aa,t)**(1-sigman(r,i))) $ alphad(r,i,aa,t) \
+ (alphan(r,i,aa,t)*m_pmp(r,i,aa,t)**(1-sigman(r,i))) $ alphan(r,i,aa,t))** (1/(1-sigman(r,i))) }
```

The domestic price macro m\_pdp is defined as follows, i.e. it applies taxes to the domestic price, again defined via the macro m\_domPrice. As taxes are usually fixed, it will in most cases simply remove a linear equation by substitution:

```
*
* --- the macro m_pdp substitutes out the domestic sales prices: with an infinite CET for exports, use the supply
* price ps, otherwise, used the domestic prices
*
$$macro m_pdp(r,i,aa,t) { m_domPrice(r,i,aa,t) \
* (1 + dintx(r,i,aa,t) + itxshft(r,aa,t) $ iTaxShift(i)) \
*
* --- add emission taxes for domestic use
* $if1 %ModulesCO2_Emissions%==on + emispr(r,t) * (emid0(r,i,aa)/xda0(r,i,aa)) $ xda0(r,i,aa) \
}
```

The macro for domestic price reflects if a CET approach is used on the supply side. If the CET is infinite, the supply price ps is used directly, otherwise the domestic sales prices pd is used (we refrain from a discussion of the details of the Melitz model):

```
$$macro m_domPrice(r,i,aa,t) [ ( pd(r,i,t) $ (omegax(r,i) ne inf) + ps(r,i,t) $ (omegax(r,i) eq inf)) $ (not iMel(i)) \
+ ( pd(r,i,t) $ (not mFlag(r,aa,i)) + ps(r,i,t) $ mFlag(r,aa,i)) $ iMel(i) ]
```

The macro for the import price m\_pmp is defined as follows, as for the domestic prices, it considers taxes on the import price pmt:

```

*
* --- the macro m_pmp substitutes out the imports sales prices
*
* $$macro m_pmp(r,i,aa,t) {pmt(r,i,t)*(1 + mintx(r,i,aa,t) + itxshft(r,aa,t)$ iTaxShift(i)) \
*
* --- add emission taxes for domestic use
* $$ifi %ModulesCO2_Emissions%==on + emisp(r,t)* (emii0(r,i,aa)/xma0(r,i,aa)) $ xma0(r,i,aa) \
* }

```

A large set of equations is dropped by substituting out prices relating to bi-lateral trade flows. The basis free-on-board price for products not treated a la Melitz are defined via a macro mm\_pecfob as follows:

```

*
* --- free on board price (f.o.b.):
*
* $$macro mm_pecfob(r,i,rp,t) { \
*
* --- export taxes
*
* [(1 + exptx(r,i,rp,t) + etax(r,i,t)) \
*
* --- the following might substitutes out pe depending on the CETs
* between domestic sales and export resp. between exports
*
* * ( \
* --- use bilateral import prices if a CET between export destination is active
* pe(r,i,rp,t) $ psFlag(r,i,"pe") \
* --- use the export prices if there is only a CET between total exports and domestic use
* + pet(r,i,t) $ psFlag(r,i,"pet") \
* --- otherwise, use supply price
* + ps(r,i,t) $ psFlag(r,i,"ps") )] $ (pecfob.range(r,i,rp,t) ne 0) \
* --- case where export price is fixed
* (used in pre-solves with only one country in model)
* + pecfob.1(r,i,rp,t) $ (pecfob.range(r,i,rp,t) eq 0) }

```

As seen, beside considering export taxes, the macro reflects which offer price to use depending whether a CET is present or not.

The final macro used in the model considers if the product is treated a la Melitz or not:

```

*
* $$macro m_pecfob(r,i,rp,t) [ m_pFirm(r,i,rp,t) * (1 + exptx(r,i,rp,t) + etax(r,i,t)) $ iMel(i) \
* + mm_pecfob(r,i,rp,t) $ (not iMel(i))]

```

That macro is subsequently used to define based on the macro m\_pmcif the cif price, taking the transport margins and exchange rates into account:

```

*
* --- costs, insurance and freight price (c.i.f): add transport margins to f.o.b. price
*
* $$macro m_pmcif(r,i,rp,t) { [ m_pecfob(r,i,rp,t)/lcu(r,t) \
* + sum(n $ angn(n,r,i,rp), angn(n,r,i,rp)*ptng(n,t)/lanbdang(n,r,i,rp,t))*tnarg(r,i,rp,t)]*lcu(rp,t) }

```

## 10.6 Exploitation and flexible aggregation

The exploitation is based on views in the multi-dimensional cube defined in p\_results. Several such cubes representing different counter-factual runs (plus typically a benchmark) can be loaded simultaneously in the viewer. The views are defined in “CgeboxTables.xml”. As the set of sectors, regions and factors can differ from run to run, “report.gms” generates a XML file “generated.xml” which that information, e.g. for the products:

```
loop(i,  
  put "<product>" /;  
  put "  <itemName>", i.te(i),"</itemName>" /;  
  put "  <key>", i.tl,"</key>" /;  
  put "  <sel>all,products,sectorView</sel>" /;  
  put "</product>" /;  
);
```

That file is included into the view definitions at run time.

In order to allow the viewer to work with maps, a co-ordinate set of individual countries is stored in “GUI\world.zip”. The standard case of GGIG applications are fixed lists of regional code and matching coordinate set which link each regional id to a list of polygons. The flexible regional aggregation in GTAP required a more flexible approach. The XML-definition of a view can register for a region a list of components under the tag “<disagg>”:

```
<region>  
  <itemName>Australia, New Zealand</itemName>  
  <key>Oceania</key>  
  <sel>all</sel>  
  <disagg>[AUS,FJI,PYF,KIR,NCL,NZL,MNP,PNG,SLB,TON,UUT,WSM,]</disagg>  
</region>
```

The individual code listed under “<disagg>” indicate the regional Ids used in the coordinate set. In the case of GTAP, these are codes for individual countries. The actual mapping between the aggregates used in the current model instance and the GTAP regions is read from an “agg”-file; the mapping between the GTAP regions code in the data base and the individual country ids is defined in “model\map\_regions.gms”, e.g.

```
rrmap("BRN","XSE") = YES;  
rrmap("MMR","XSE") = YES;
```

That file should be currently set-up to work with GTAP version 8. Other versions require an update of the mappings, as the list of GTAP regions might differ.

A first view reports the meta-information on the model. It is here used as a first example for the XML based definition of views. Basically, each view can define filters in the different dimension. The filters end with “sel” and start with the logical name of a dimension. To give an example: where normally information on a sector or institution would be stored (fourth dimension, the “activityDim”), the label “tots” is used. It is chosen in the filter below based on a regex expression:

```
<activitySel>tots_REGEX</activitySel>
```

```

<table>
  <theme>Model overview</theme>
  <name>Model properties</name>
  <regionSel>Wor_REGEX</regionSel>
  <regionText>hide</regionText>
  <productSel>.*_REGEX</productSel>
  <productText>hide</productText>
  <activitySel>tots_REGEX</activitySel>
  <activityText>hide</activityText>
  <dim5Sel>Meta</dim5Sel>
  <dim5Text>hide</dim5Text>
  <hideEmptyRows>YES</hideEmptyRows>

  <dim6Text>hide</dim6Text>

  <defview>Table</defview>
  <defpivot>ØP ØS</defpivot>

  <itemDim>dim6</itemDim>
    <item>
      <itemName>Dummy</itemName>
      <key>dom</key>
    </item>
  </table>

```

Another example provides the table for intermediate demand. Here all products and sectors are selected; instead of a regex expression, the pre-defined lists found in “generated.xml” are used:

```

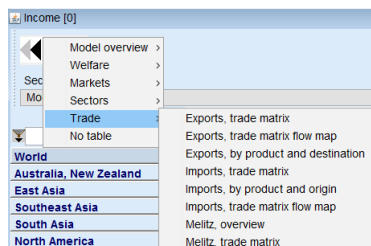
<table>
  <name>Intermediate demand</name>
  <theme>Markets</theme>
  <regionSel>all</regionSel>
  <productSel>products</productSel>
  <productText>Inputs</productText>
  <activitySel>sectors</activitySel>
  <activityText>sectors</activityText>
  <dim5Sel>all</dim5Sel>
  <dim6Text>Origin</dim6Text>
  <dim6Sel>origins</dim6Sel>

  <defview>Table</defview>
  <defpivot>ØPSØ</defpivot>
  <itemDim>dim6</itemDim>
    <item>
      <itemName>Total</itemName>
      <key>tot</key>
    </item>
    <item>
      <itemName>Imported</itemName>
      <key>Imp</key>
    </item>
    <item>
      <itemName>Domestic</itemName>
      <key>Dom</key>
    </item>
  </table>

```

Details on how views can be defined can be found in GGIG programming guide, Britz 2010b.

Detail on simulated value of variables found in the Melitz module can be found in two tables under the “Trade group” as shown below:



Meltz, overview [0]

Products: Total Years: shock Percentage diff. to Scenario: noShock

ttp

|                          | All firms entered | All firms operating | Variable cost price | Fix cost price | General fixed cost | Bi-lateral fixed cost | Variable cost | General fixed cost per unit | Bi-lateral fixed cost per unit | Variable cost per unit | All firms entered | All firms operating |
|--------------------------|-------------------|---------------------|---------------------|----------------|--------------------|-----------------------|---------------|-----------------------------|--------------------------------|------------------------|-------------------|---------------------|
| World                    |                   | 0,00                | 1,00                | 1,00           | 5217,96            | 1637,32               | 24002,60      | 0,14                        | 0,04                           | 0,66                   |                   |                     |
|                          |                   | 0,11%               | -0,01%              | -0,00%         | 0,00%              | 0,00%                 | 0,00%         | -0,00%                      | 0,00%                          | -0,00%                 |                   |                     |
| European Union 28        | 1,00              | 0,00                | 1,00                | 1,00           | 1654,72            | 457,21                | 7611,71       | 0,15                        | 0,04                           | 0,69                   | 1,00              |                     |
|                          | 0,01%             | 0,41%               | -0,02%              | -0,01%         | -0,01%             | -0,01%                | -0,01%        | -0,02%                      | -0,02%                         | -0,02%                 |                   |                     |
| Norway, Switzerland      | 1,00              | 0,03                | 1,00                | 1,00           | 63,72              | 15,01                 | 293,13        | 0,16                        | 0,04                           | 0,72                   | 1,00              |                     |
|                          | -0,02%            | 0,05%               | -0,01%              | -0,01%         | -0,03%             | -0,06%                | -0,03%        | -0,01%                      | -0,03%                         | -0,01%                 |                   |                     |
| United States of America | 1,00              | 0,00                | 1,00                | 1,00           | 929,68             | 315,31                | 4276,51       | 0,15                        | 0,05                           | 0,67                   | 1,00              |                     |
|                          | 0,01%             | 0,71%               | 0,00%               | 0,02%          | 0,03%              | 0,06%                 | 0,03%         | 0,05%                       | 0,07%                          | 0,05%                  |                   |                     |
| North America            | 1,00              | 0,02                | 1,00                | 1,00           | 105,15             | 35,58                 | 483,70        | 0,15                        | 0,05                           | 0,68                   | 1,00              |                     |
|                          | 0,01%             | -0,04%              | 0,00%               | 0,01%          | 0,03%              | -0,01%                | 0,03%         | 0,02%                       | -0,02%                         | 0,02%                  |                   |                     |
| Mercosur                 | 1,00              | 0,01                | 1,00                | 1,00           | 157,07             | 63,75                 | 722,54        | 0,15                        | 0,06                           | 0,67                   | 1,00              |                     |
|                          | 0,00%             | 0,01%               | -0,01%              | -0,01%         | -0,01%             | -0,01%                | -0,01%        | -0,01%                      | -0,02%                         | -0,01%                 |                   |                     |
| China                    | 1,00              | 0,00                | 1,00                | 1,00           | 694,68             | 237,00                | 3195,53       | 0,13                        | 0,04                           | 0,59                   | 1,00              |                     |
|                          | 0,00%             | -0,01%              | -0,01%              | -0,00%         | -0,00%             | -0,00%                | -0,00%        | -0,00%                      | -0,01%                         | -0,00%                 |                   |                     |
| ASEAN 10                 | 1,00              | 0,01                | 1,00                | 1,00           | 198,15             | 54,63                 | 911,50        | 0,15                        | 0,04                           | 0,68                   | 1,00              |                     |
|                          | 0,01%             | -0,01%              | -0,01%              | 0,00%          | 0,01%              | -0,03%                | 0,01%         | -0,00%                      | -0,04%                         | -0,00%                 |                   |                     |
| OtherOECD                | 1,00              | 0,00                | 1,00                | 1,00           | 694,27             | 231,10                | 3193,65       | 0,13                        | 0,04                           | 0,62                   | 1,00              |                     |
|                          | 0,00%             | -0,01%              | -0,00%              | -0,00%         | -0,00%             | -0,01%                | -0,01%        | -0,01%                      | -0,01%                         | -0,01%                 |                   |                     |
| EU Mediterrean Partners  | 1,00              | 0,01                | 1,00                | 1,00           | 123,51             | 38,28                 | 568,13        | 0,15                        | 0,05                           | 0,70                   | 1,00              |                     |
|                          | -0,00%            | 0,04%               | -0,01%              | -0,01%         | -0,01%             | -0,03%                | -0,01%        | -0,01%                      | -0,02%                         | -0,01%                 |                   |                     |
| Low Income               | 1,00              | 0,10                | 1,00                | 1,00           | 19,82              | 5,97                  | 91,17         | 0,16                        | 0,05                           | 0,72                   | 1,00              |                     |
|                          | -0,00%            | 0,07%               | -0,02%              | -0,02%         | -0,02%             | -0,06%                | -0,02%        | -0,01%                      | -0,06%                         | -0,01%                 |                   |                     |
| Rest of World            | 1,00              | 0,00                | 1,00                | 1,00           | 577,18             | 183,48                | 2655,02       | 0,15                        | 0,05                           | 0,69                   | 1,00              |                     |
|                          | -0,00%            | 0,01%               | -0,01%              | -0,01%         | -0,01%             | -0,02%                | -0,01%        | -0,01%                      | -0,02%                         | -0,01%                 |                   |                     |
| World                    |                   | 0,00                | 1,00                | 1,00           | 5217,96            | 1637,32               | 24002,60      | 0,14                        | 0,04                           | 0,66                   |                   |                     |
|                          |                   | 0,11%               | -0,01%              | -0,00%         | 0,00%              | 0,00%                 | 0,00%         | -0,00%                      | 0,00%                          | -0,00%                 |                   |                     |
| European Union 28        | 1,00              | 0,00                | 1,00                | 1,00           | 1654,72            | 457,21                | 7611,71       | 0,15                        | 0,04                           | 0,69                   | 1,00              |                     |
|                          | 0,01%             | 0,41%               | -0,02%              | -0,01%         | -0,01%             | -0,01%                | -0,01%        | -0,02%                      | -0,02%                         | -0,02%                 |                   |                     |

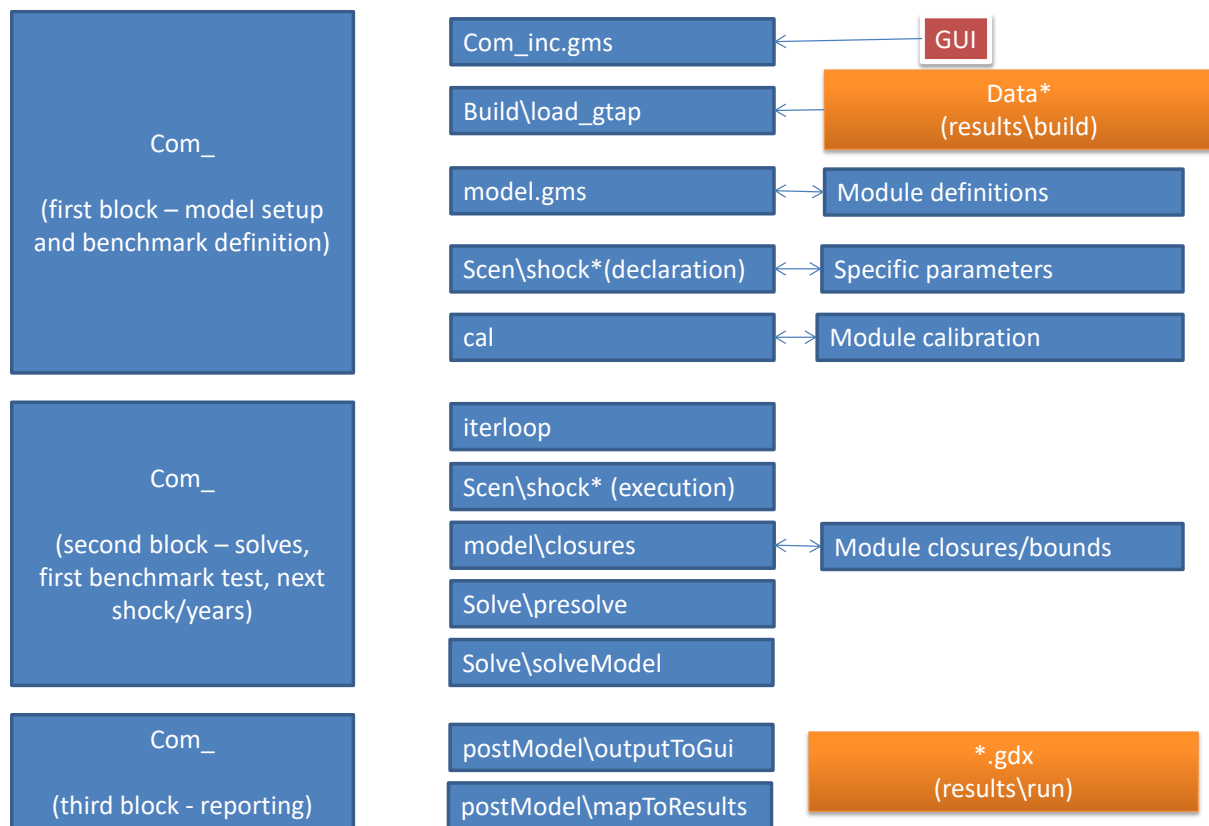
Meltz, trade matrix [0]

Items: Firm price Products: Total Years: shock Percentage diff. to Scenario: noShock

ttp

|                          | Domestic sales | European Union 28 | Norway, Switzerland | United States of America | North America | Mercosur | China  | ASEAN 10 | OtherOECD | EU Mediterrean Partners | Low Income | Rest of World | Domestic sales | European Union 28 |
|--------------------------|----------------|-------------------|---------------------|--------------------------|---------------|----------|--------|----------|-----------|-------------------------|------------|---------------|----------------|-------------------|
| World                    | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,03%         | 0,02%             | -0,01%              | 0,10%                    | 0,01%         | -0,01%   | -0,01% | 0,00%    | -0,00%    | -0,01%                  | -0,01%     | -0,01%        | -0,01%         | -0,01%            |
| European Union 28        | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,06%         | -0,02%            | -0,01%              | 0,49%                    | 0,02%         | -0,01%   | -0,00% | 0,00%    | 0,00%     | -0,01%                  | -0,01%     | -0,01%        | -0,01%         | -0,01%            |
| Norway, Switzerland      | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,01%         | -0,03%            | -0,00%              | -0,00%                   | 0,01%         | -0,01%   | -0,00% | 0,01%    | 0,00%     | -0,00%                  | -0,01%     | -0,00%        | -0,00%         | -0,00%            |
| United States of America | 1,00           | 1,01              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,04%         | 0,76%             | -0,03%              | 0,01%                    | -0,02%        | -0,01%   | -0,01% | -0,00%   | -0,01%    | -0,02%                  | -0,02%     | -0,02%        | -0,01%         | -0,01%            |
| North America            | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | 0,01%          | -0,03%            | -0,03%              | -0,01%                   | -0,02%        | -0,01%   | -0,01% | -0,00%   | -0,01%    | -0,02%                  | -0,02%     | -0,01%        | -0,01%         | -0,01%            |
| Mercosur                 | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,01%         | -0,06%            | -0,02%              | -0,02%                   | 0,02%         | -0,01%   | -0,00% | -0,02%   | -0,00%    | -0,01%                  | 0,02%      | -0,01%        | -0,01%         | -0,01%            |
| China                    | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,01%         | -0,02%            | -0,02%              | 0,01%                    | -0,01%        | -0,01%   | -0,00% | 0,00%    | -0,00%    | -0,02%                  | -0,02%     | -0,01%        | -0,01%         | -0,01%            |
| ASEAN 10                 | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,00%         | -0,03%            | -0,02%              | -0,02%                   | -0,00%        | -0,02%   | -0,01% | -0,00%   | -0,01%    | -0,02%                  | -0,02%     | -0,01%        | -0,01%         | -0,01%            |
| OtherOECD                | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,00%         | -0,03%            | -0,02%              | 0,00%                    | 0,02%         | -0,01%   | -0,01% | 0,00%    | -0,00%    | -0,02%                  | -0,02%     | -0,01%        | -0,01%         | -0,01%            |
| EU Mediterrean Partners  | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,01%         | -0,02%            | -0,02%              | 0,01%                    | -0,01%        | -0,01%   | -0,00% | 0,00%    | -0,00%    | -0,01%                  | -0,01%     | -0,01%        | -0,01%         | -0,01%            |
| Low Income               | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,01%         | -0,11%            | -0,02%              | 0,00%                    | 0,02%         | -0,01%   | -0,00% | 0,01%    | -0,00%    | -0,00%                  | -0,00%     | -0,00%        | -0,00%         | -0,00%            |
| Rest of World            | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,01%         | -0,03%            | -0,02%              | -0,00%                   | 0,02%         | -0,01%   | -0,00% | 0,00%    | -0,01%    | -0,01%                  | -0,01%     | -0,01%        | -0,01%         | -0,01%            |
| World                    | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,03%         | 0,02%             | -0,01%              | 0,10%                    | 0,01%         | -0,01%   | -0,01% | 0,00%    | -0,00%    | -0,01%                  | -0,01%     | -0,01%        | -0,01%         | -0,01%            |
| European Union 28        | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,06%         | -0,02%            | -0,01%              | 0,49%                    | 0,02%         | -0,01%   | -0,00% | 0,00%    | 0,00%     | -0,01%                  | -0,01%     | -0,01%        | -0,01%         | -0,01%            |
| European Union 28        | 1,00           | 1,00              | 1,00                | 1,00                     | 1,00          | 1,00     | 1,00   | 1,00     | 1,00      | 1,00                    | 1,00       | 1,00          | 1,00           | 1,00              |
|                          | -0,06%         | -0,02%            | -0,01%              | 0,49%                    | 0,02%         | -0,01%   | -0,00% | 0,00%    | 0,00%     | -0,01%                  | -0,01%     | -0,01%        | -0,01%         | -0,01%            |

## 10.7 Main programs when solving a shock and integration of modules



The main driver program “`com_.gms`” can be roughly broken into three major blocks:

- I. **Benchmark set-up:** That first block deals with loading the data and parameters and setting up the model such that it replicates the benchmark.
  - a. *Data input and model setup:* The information about what files, extensions and modules to use and further detail on the model run is read from “`com_inc.gms`”, generated by the GUI. As a next step, the “`build\loadGTAP_agg.gms`” reads the global SAM and related parameters and initialized core sets (products, activities, agents, households etc.). These sets are either directly read from the data set or are defined by run specific file such as the household types present in the myGTAP module.
  - b. Next, the *equation system is set-up*. The core equations for the standard GTAP model plus equations to host extensions are found in “`model.gms`” along with the necessary declarations of variables and parameters. According to the modules in use, additional equations are loaded in from module definitions which might replace equations from the standard GTAP model.
  - c. Based on that information, the *calibration step* in “`cal.gms`” initializes all variables for the benchmark and calculates parameters such as the share parameters for CES functions. The different modules as the GTAP-Melitz extension have their own calibration programs which are included on demand.
- II. The second block comprises the model solve.
  - a. First, all variables are copied from t-1 to t (skipped for benchmark), see “`iterloop.gms`”. Iterloop might shift variables depending on the chosen start point. If G-RDEM is active, it will run.

- b. Next, if not in benchmark check, the shock file is executed which hence can either use the results from t-1 (recursive dynamics) and/or benchmark results.
- c. Followed by that step is the definition of bounds for the variable which can include fixing depending on the closures. Again, module specific includes will initialize bounds e.g. for GTAP-AGR.
- d. On demand, a start point can be constructed based on forecasting and initialization.
- e. Next, the current year is solved. If that is a benchmark solve, it will only build the equation system and does not allow the solver to improve the model (iterlim=0). That step should yield a total sum of infeasibilities  $< 1.E-4$ , otherwise, probably, some error in the benchmarking has occurred. In follow up solves, pre-solve steps for the individual regions might be added. The model can attempt several solves and fix variables at their lower (tiny) bounds in CNS/NLP mode.
- f. Variables for non-reported years are removed in rec-dyn mode (*solve/removeNonRep.gms*).
- g. On demand, bounds, marginal and scaling and all fields for equations are removed (*solve/kill\_bounds.gms*). If the heapsize (= memory use by GAMS) has increased by more than 100 MB, a reorganization is triggered. It is not recommended to switch this off on the interface as doing so can severely reduce executing speed.
- h. Calculate substituted out variables from the model solution that feed into reporting.

### III. Reporting

- a. The reporting block will in the simplest case only store all variables, sets and parameters to a GDX.
- b. If output to the GUI is selected, the variables and parameters are mapped onto one multi-dimensional parameter *p\_results* which can be inspected by the GUI (see above), performed by “*postmodel/mapToResults*”. In that case, an XML file will be generated which comprises the list of products, activities etc. for use with the exploitation tools.
- c. Further output options (building a SAM, ALTERNATE etc.) can be activated on the GUI and are executed accordingly.

## 10.8 Tests and checks

The code comprises a range of inbuilt tests and assertions. This reflects recommendation on software quality management by Britz et al. 2026<sup>130</sup>.

### 10.8.1 Alerting users of compile and run-time errors

1. Almost all files will throw an error at compile time or runtime when they find errors raised by the software. This is implemented by two lines at code at the end of the file:

```
$$if not errorfree $abort Compilation error after file: %system.fn%
```

```
if ( execerror, abort "Run-Time error in file: %system.fn%, line: %system.incline%");
```

Run-time errors are not checked for most of the post-model reporting to avoid that results from an otherwise error-free simulation run are not stored to disk due to errors in post-model reporting. Doing

---

<sup>130</sup> Britz, W., Jansson, T., & Schäfer, D. (2026). Reflections and recommendations on Software Quality Management in economic models. *Social Sciences & Humanities Open*, 14, 103184

can prevent a flood a follow-up errors after the first error but will mean that new errors can emerge in follow-up files once the first ones have been successfully removed.

There are also various checks for runtime errors after statements that haven proven prone to provoke divisions by zero and similar. These checks typically display parameters or variables that might help to find the offending cases, such as:

```
if ( execError, abort "Run-Time error in file: %system.fn%, line: %system.incline%", pa.l;);
```

### 10.8.2 Data base checks

The GTAPAgg routines tests for a closed SAM, see “*gtapAgg/buildSam.gms*”. This check is repeated if a split is applied in “*build\split\split\_reportAndMap.gms*”.

The benchmarking program “*cal/cal.gms*” tests various exhausting conditions. Besides the SAM, also additional data such as the decomposition to the domestic and imported origin are subject to checks. An example is shown below:

```
parameter testit "Help parameter to test various balance conditions";
*
* --- check that output at net-of-tax value is equal to value
*   of intermediate plus value-added for all regions and activities
*
testit(rs,a,t0) $ xp.l(rs,a,t0) = xp.l(rs,a,t0)*px.l(rs,a,t0)-xtNest.l(rs,"va",a,t0)-xtNest.l(rs,"nd",a,t0);
*
* --- errors below a certain threshold are accepted (set by user on interface)
*
testit(rs,a,t0) $ (abs(testit(rs,a,t0)) le 1.E-%imbalaceThreshold%) = 0;
if ( Card(testit),
*
* ---- we found imbalance: remove balanced cases
*
xp.l(rs,a,t0)          $ (not testit(rs,a,t0)) = 0;
xtNest.l(rs,"va",a,t0) $ (not testit(rs,a,t0)) = 0;
xtNest.l(rs,"nd",a,t0) $ (not testit(rs,a,t0)) = 0;
px.l(rs,a,t0)          $ (not testit(rs,a,t0)) = 0;
prdtx.l(rs,a,t0)       $ (not testit(rs,a,t0)) = 0;

sam0(rs,is,js) $ ( (not testit(rs,js,"%t0%")) ) = 0;
*
* --- report variable of interest for imbalanced case to help with debugging
*
abort "error in SAM, production value does not fit va + int, accuracy 1.E-%imbalaceThreshold%",
testit,xp.l,xtNest.l,px.l,prdtx.l,sam0,
```

### 10.8.3 Checks built into the equation system of the model and shock files for tests

The code will check that the model is square, i.e. that the number of non-fixed variables is equal to the number of equations. Note here that GAMS will remove empty equations, they do not count. Errors with non-square models can be typically found by solving the model as an MCP. Equations in the models are paired with a few exemptions where closure swaps are built in the code. Built in MCP checks in GAMS will hence alert the user of non-free variables that are not matched to equations.

The following checks are coded in the equation system

- The equation *e\_walras* checks  $I=S$  for the residual region for which the foreign savings are determined residually. A larger non-zero value for the walras variable implies some logical flaw in the model. This hints typically a violation of some exhaustion condition, i.e. all not all revenues or costs/expenditures are accounted.
- Similarly, the *e\_bopSlack* equation tests for closed B.o.P. for each model region. All model versions comprise the B.o.T. and the foreign savings in the B.o.P. check, depending on the



model set-up additional financial transactions such as remittances or debt payments might be consider in the test.

A test for the correct implementation of homogeneity conditions (no money illusion) for demand functions, cost functions and revenue maximization problems can be run with the so-called numeraire shock (*gams/scen/test\_shocks/num\_shock.gms*). It increases the fixed numeraire price  $pNum$  in the model by 20%. All prices should follow the 20% increase while quantities should stay unchanged.

A no-shock should keep the model solution unchanged in a comparative-static setting. These can be tested by running the shock files “empty” or “noShock” from *scen/testShocks* directory. The code will automatically check for no change in real GDP against the benchmark in this case.

#### 10.8.4 Checks for negative results and exhaustion

The program “*gams/solve/checkZeros.gms*”, called after the model is solved, tests if various variables are below  $-1.E-7$ . Note that  $1.E-7$  is the default infeasibility tolerance chosen for the CONOPT4 of the solver. Accordingly, in case where a variable has a minus infinite lower bound, such small negative value might be found despite a correct implementation of functional relationships. Offending cases are reported but not considered an error:

```
*
* --- report if such cases are found
*
if (card(checkZeros),
    display "Zero or negative results", checkZeros;
    checkZeros("sum", "", "", "", "") = sum( (r,u1,u2,u3,u4) $ checkZeros(u1,r,u2,u3,u4), abs(checkZeros(u1,r,u2,u3,u4)));
    checkZeros("max", "", "", "", "") = smax( (r,u1,u2,u3,u4) $ checkZeros(u1,r,u2,u3,u4), abs(checkZeros(u1,r,u2,u3,u4)));
    put_utility 'msglog' / 'N variables are negative, ' card(checkZeros):0:0 ;
);
```

Afterwards, an additional program “*solve/minLevels.gms*” sets back price and output levels in the positive domain, using minimal values of  $1.E-6$  for prices and  $1.E-10$  for quantities

Another program, “*solve/resetNegatives*”, more cautiously used only after repeated infeasibilities, tests for much more cases and uses more aggressive resets, such as 10% price levels from the previous solution (the benchmark for comparative-static runs).

If the output type “Check” is added, “*postModel/checkRes.gms*” checks for various exhaustion conditions and reports offending cases.

The default settings of the GUI will throw an error if after the number of allowed repeated trials the model comprises larger infeasibilities. Excluded are intermediate solves during shock slicing. It is normally not recommended to switch this option off.

#### 10.8.5 Test suite

The modular structure of CGEBox combined with various, partially mutually exclusive options implies that only a part of the overall code will be compiled and run for any simulation run. During code development, working with one setup of the model, only, might not reveal errors triggered if other setups are used. Manual tests with different set-ups are certainly helpful but will likely not systematically explore all relevant combinations of modules and options. Numerical problems might also depend on the detail of the data the model is exposed to. Especially issues with sparsity, i.e. missing data, often become visible only once highly detailed or aggressively filtered data bases are used for benchmarking.

To advance here, the batch utility of GGIG is used. It allows to start in an automated fashion a pre-defined set of runs. The test-suite is called “*gui/tests\_cgebox.txt*” and encompasses test cases for tests data base generation, comparative static shocks, sensitivity analysis as well as long-term baseline generation with G-RDEM.

It constructs three data bases from GTAP-Power V12 uses filtering with a threshold of 1.E-5, adds the gender differentiated labor and generates the MRIO split factors. The three data bases are a default 10x10 case, a 76x24 and a 76x10 case plus Ireland (to check for the NUTS2 extensions) with a further split of this data base to agri-food detail (additional 76 products). The two larger data bases comprise one electricity product, only.

For the three data sets, two simulations are run (“no-shock”, a global 50% tariff liberalization scenario as code in “*gams/scen/policy\_shocks/tariff.gms*”), with zero or three rounds of pre-solves. The start point is set to “*Benchmark plus initialization*” such that also the lengthy code of “*solve/iniVars.gms*” is tested. The tariff liberalization scenario is inviting as test as modules known to provoke numerical problems, such as the Melitz and SPE ones, react sensitive to changes in tariffs.

For each of the 12 combinations (3 data sets x 2 scenarios x 2 options for pre-solves), at least 19 test cases are applied (if not otherwise stated, all other settings at defaults as defined in “*gui/cgebox.xml*”):

- Using the CDE, LES, AIDADS and CD demand systems without further modules. All other tests use the default CDE demand system.
- Restricted module tests, i.e. only one or some modules are active:
  - MyGTAP (one aggregate household)
  - GTAP-AEZ (four nests with all AEZ aggregated on/off, six nests with aggregated off)
  - All env-accounting module switched on (CO-2 emissions, Non-CO2 emissions, Air pollution emissions, MTOE accounting)
  - Armington aggregation: All, government plus household, all intermediate demand
  - GTAP Melitz, with and without fix cost nest (one product)
  - Armington with commitment terms (one product)
  - SPE (one product)
  - GTAP-MRIO (all products)
  - Capital vintages
  - Sub-regions (one country)
  - Extended parameters (labor markets sluggish, CET transformation elasticities finite)
  - Nestings (maidads, food-feed)
  - Different output options combined (GUI, GDX, CONVERT, ALTERNATE)
- Combined test: myGTAP, sub-regions, GTAP-AGR, GTAP-AEZ; all env-accounting modules, GTAP-Melitz, Armington with commitments, SPE, GTAP-MRIO, GTAP-E, Nestings

For the 76x10\_IRL case and the tariff shock solved with up to three pre-solves iterations, also shock slicing, the single region set-up and the partial equilibrium variants with three products are tested. This yields in total 279 tests with the comparative-static version. After these tests, the sensitivity analysis is tested for the 10x10 case with changes in the lower-level Armington elasticities.

Finally, the baseline generation with G-RDEM from 2023 to 2050 in bi-yearly steps is subject to two tests. The first one uses two different model set-ups: one with all environmental accounting options, GTAP-E and the AISADS demand system and a second evolved one that adds the sub-regions,

myGTAP, GTAP-AEZ and GTAP-AGR. Baselines for all SSPs for the two set-ups construct are constructed, with all or no G-RDEM features, yielding ten tests. The second test constructs base-lines for SSP2 using the evolved model set-up with all and all G-RDEM features with the available six different options to construct a start point.

The compile time checks take about 2 minutes<sup>131</sup>. The 10x10 test cases with and without the pre-solves in the no-shock scenario require and most of the trade liberalization scenario require not more than 3 seconds. If all 310 tests are run, they take around 50 minutes.

Checking the outcome of the test suite is relatively straightforward as the batch mode generates a HTML page that documents each run (settings chosen), reports errors and saves the listing and include files. The screen shot below shows the first lines of the page. If users want to apply the test suite, they might need to switch the parallel execution off and increase maximal solutions, depending on the machine the tests are started. Currently, solely two test fails: the generation of base lines without any G-RDEM features become infeasible for SSP1.

#### GGIG batch execution report

| Runs are generated from                                                |            |      |                                                   |                                                                                                                                                                                                                     | C:\cgebox\guitests_cgebox.txt                                                             |                                                                            |                                                   |                                     |  |
|------------------------------------------------------------------------|------------|------|---------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|---------------------------------------------------|-------------------------------------|--|
| at                                                                     |            |      |                                                   |                                                                                                                                                                                                                     | Mi, 17 Sep 25 15:12:40                                                                    |                                                                            |                                                   |                                     |  |
| Worksheet                                                              | Task       | User | Directories:<br>work<br>result<br>data<br>restart | Settings                                                                                                                                                                                                            | Game options                                                                              | Started<br>Ended<br>Time used                                              | RC                                                | Listing<br>OutDir                   |  |
| Scenario file: base_scenarios/noShock; Input file: 10x10; presolves: 0 |            |      |                                                   |                                                                                                                                                                                                                     |                                                                                           |                                                                            |                                                   |                                     |  |
| Simulation                                                             | Simulation |      | C:\cgebox\gams<br>.\results<br>.\data<br>-        | 10x10_Standard_0_presolves<br><a href="#">Updated settings (click to show/hide)</a><br><a href="#">All settings in use (click to show/hide)</a><br><a href="#">Non default settings in use (click to show/hide)</a> | run<br>-scen=com_int<br>-ggp=on procdirPath=c:\scdir threads=12<br>BreadthKey=0 qtdir=opt | Mo, 17 Sep 25<br>15:12:41<br>Mo, 17 Sep 25<br>15:12:43<br>0 min and 1 secs | 0                                                 | C:\cgebox\17.09.2025_15.12.40\1.txt |  |
| Simulation                                                             | Simulation |      | C:\cgebox\gams<br>.\results<br>.\data<br>-        | 10x10_LES_0_presolves<br><a href="#">Updated settings (click to show/hide)</a><br><a href="#">All settings in use (click to show/hide)</a><br><a href="#">Non default settings in use (click to show/hide)</a>      | run<br>-scen=com_int<br>-ggp=on procdirPath=c:\scdir threads=12<br>BreadthKey=0 qtdir=opt | Mo, 17 Sep 25<br>15:12:43<br>Mo, 17 Sep 25<br>15:12:45<br>0 min and 1 secs | 0                                                 | C:\cgebox\17.09.2025_15.12.40\2.txt |  |
| Simulation                                                             | Simulation |      | C:\cgebox\gams<br>.\results<br>.\data<br>-        | 10x10_ADAD0_0_presolves<br><a href="#">Updated settings (click to show/hide)</a><br><a href="#">All settings in use (click to show/hide)</a><br><a href="#">Non default settings in use (click to show/hide)</a>    | run<br>-scen=com_int<br>-ggp=on procdirPath=c:\scdir threads=12<br>BreadthKey=0 qtdir=opt | Mo, 17 Sep 25<br>15:12:45<br>Mo, 17 Sep 25<br>15:12:47<br>0 min and 2 secs | 0                                                 | C:\cgebox\17.09.2025_15.12.40\3.txt |  |
| Simulation                                                             | Simulation |      | C:\cgebox\gams<br>.\results<br>.\data<br>-        | 10x10_CD_0_presolves<br><a href="#">Updated settings (click to show/hide)</a><br><a href="#">All settings in use (click to show/hide)</a><br><a href="#">Non default settings in use (click to show/hide)</a>       | run<br>-scen=com_int<br>-ggp=on procdirPath=c:\scdir threads=12<br>BreadthKey=0 qtdir=opt | Mo, 17 Sep 25<br>15:12:48<br>Mo, 17 Sep 25<br>15:12:49<br>0 min and 1 secs | 0                                                 | C:\cgebox\17.09.2025_15.12.40\4.txt |  |
| Simulation                                                             | Simulation |      | C:\cgebox\gams<br>.\results<br>.\data<br>-        | 10x10_myGTAP_0_presolves<br><a href="#">Updated settings (click to show/hide)</a><br><a href="#">All settings in use (click to show/hide)</a><br><a href="#">Non default settings in use (click to show/hide)</a>   | run<br>-scen=com_int<br>-ggp=on procdirPath=c:\scdir threads=12<br>BreadthKey=0 qtdir=opt | Mo, 17 Sep 25<br>15:12:49<br>Mo, 17 Sep 25<br>15:12:51<br>0 min and 1 secs | 0<br>Errors<br><a href="#">Click to show/hide</a> | C:\cgebox\17.09.2025_15.12.40\5.txt |  |

Failed test runs are shown in red such that they can be easily spotted.

|            |            |  |                                            |                                                                                                                                                                                                                         |                                                                                                            |                                                                               |                                                                                                                                                                               |                                     |
|------------|------------|--|--------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|
| Simulation | Simulation |  | C:\cgebox\gams<br>.\results<br>.\data<br>- | 10x10_OTAP_Multi_3Presolves<br>Updated settings( <a href="#">click to show/hide</a> )<br>All settings in use( <a href="#">click to show/hide</a> )<br>Non default settings in use( <a href="#">click to show/hide</a> ) | optdir=opt<br>run<br>-scen=com_int<br>-ggp=on procdirPath=c:\scdir<br>threads=12 threadsKey=0<br>qtdir=opt | Mo, 17 Sep<br>25 15:15:37<br>Mo, 17 Sep<br>25 15:15:42<br>0 min and 5<br>secs | 3<br>Errors<br>( <a href="#">click to show/hide</a> )<br>Error at line 19918: Execution halted:<br>about "Some relevant infeasibilities in last<br>model solve were observed" | C:\cgebox\17.09.2025_15.12.40\7.txt |
|------------|------------|--|--------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|

<sup>131</sup> The test suite is run by the team in Bonn on a large computing server with 96 cores where blocks of tests (such as same data base, shock and number of pre-solves for the comparative-static case), are run in parallel based on the parallel execution option built into the batch mode of GGIG. The tests take in total around 50 minutes to run on this machine. Note that the compile test will fail if the data bases are not found on disk.