CGEBox in a nutshell

Wolfgang Britz, August 2026

CGEBox is a flexible and modular modelling framework for Computable General Equilibrium (CGE) building on the GTAP Data base and auxiliary data from the GTAP center, distributed with some additional open-access data. It comprises the full workflow of working with GTAP-based CGE models: aggregation, filtering and splitting the data base, benchmarking of the model, model runs, sensitivity analysis and tools for results analysis.

CGEBox can configure differently structured single country and global CGE models in comparative-static or recursive dynamic mode, supporting flexible aggregation by region, product, sector and factor from the GTAP Data Base. It is realized in **GAMS** such that all equations are coded in levels. It applies different approaches to allow for a fast solution of highly detailed model instances (Britz and van der Mensbrugghe 2016, Britz 2025) such as a filter routine to skim transactions related to tiny cost and expenditure shares from the GTAP data base, variable substitutions, pre-solves over single country models and using CONOPT4 (based on Drud 1994) in parallel mode. Different options to construct a starting point for the solver with a low number and sum of infeasibilities from the last solution and the shock are available (Britz 2025). It also supports a MCP setup, for instance, to model production or emission quotas. CGEBox comes with a **Graphical User Interface** based on GGIG (GAMS Graphical User Interface Generator, Britz 2014) which lets the user build a suitable data base, configure and run the model, and compare scenarios based on tables, graphs, maps and statistics (cf. Britz et al. 2015). It is **open source and open access**, its code and the freely available data bases feeding in the different modules can be downloaded from a SVN server[1]. A GAMS license is required for model runs plus a license to the GTAP data base for models beyond a 10 region and 10 sector size. CGEBox is currently mostly maintained by a team at the University Bonn and Roberto Roson (Unversity Venice and Loloya University Seville); Dominique van der Mensbrugghe contributes the GTAP Standard model in GAMS (van der Mensbrugghe 2018). CGEBox is fully documented (Britz 2019[2]) and peer-reviewed (Britz and van der Mensbrugghe 2018).

CGEBox departs from the GTAP Standard model (Hertel and Tsigas 1997) as coded by van der Mensbrugghe 2018 which is based on set of canonical assumptions found in most CGE models: competitive markets for products and factors; industries are modelled through representative, cost-minimizing firms operating under constant returns to scale and zero profits; households maximize utility under a budget constraint; revenues are obtained by selling services of primary factors; all macroeconomic identities hold. The GTAP Standard Model adds the so-called “global bank” that collects all savings globally and distributes them as investments to the regions based on expected returns to capital. Accordingly, both foreign and regional savings finance investments such that the BOT in each region is endogenous. Alternative closures to the BOT are supported by the code. Similar to the global bank, a central agent provides global transport services to bi-lateral trade flows differentiated by transport mode. Model and data base depict taxes in rich detail.

[1] https://svn1.agp.uni-bonn.de/svn/cgebox/, password and userid is *cgebox*
[2] https://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox_GUI.pdf

In the GTAP standard model, the so-called regional household approach collects in each region all income and distributes it to regional savings, private and public consumption, depicting a social welfare maximization problem. The alternative **myGTAP module** of CGEBox allows to introduce one or **multiple private households** which manage their factor stocks and distribute household income (from factor income after taxes, private and public transfers, remittances) to consumption, saving, private transfers to other households and remittances. The myGTAP module also introduces a separate account for the government with debt dynamics. It supports asymmetric set-ups where the number of households can differ from region to region. Household surveys for 19 developing countries (Squarcina 2017) distributed with the model code[3] allow to flexibly define household types for these countries directly from the Graphical User Interface. The household survey data can also be used with a post-model micro-simulation extension (see for instance Wilts and Britz 2024). Alternatively, users can provide their own split shares to define households. Households can also be linked to sub-national regions in factor supply and final demand; the latter allows to depict subsistence production[4]. Adding the **GMIG** module (Walmsley et al. 2011) and data allows simulating bi-lateral global labour and population migration along with remittances, reporting both head counts of workers and wages per worker.

For final demand, the user has the choice between a **CDE, LES, CD or MAIDADS demand system**. For the latter, econometrically estimated parameters are available (Britz 2021); the other systems are parameterized based on the CDE parameters available with the GTAP data base. A flexible nesting approach allows introducing CES-nests under these top-level demand systems to better depict cross-price effects between closely related products such as, for instance, different types of meat. Pre-defined nests are available, for instance, for product groups comprised in the MAIDADS demand systems, different food and feed categories, for different energy carriers and for the electricity mix if the GTAP-Power data base is used. The equation system supports the use of volume-preserving CES nests (Van der Mensbrugghe and Peters 2016).

The default set-up based on the GTAP Standard model depicts international **trade** as a **two-stage Armington** approach where the different production sectors, savings as well as private and government consumption each have their own top-level nest which determines the expenditure share of the domestic and imported origin. A second nest with identical shares across all agents determines bi-lateral import shares. Alternatively, all firm demand, all demand, or private and government demand can be aggregated in the top-level nest. The Armington approach can be complemented with a Constant-Elasticity-of-Transformation (CET) function to distribute domestic supply, moreover, commitment terms can be added to the CES Armington function to address the small-shares stays small problem. A heterogenous firm model a la **Melitz** with love of variety, fix cost for industry entry, fix costs on trade links, price mark-ups and thus monopolistic competition can be used instead of the Armington approach, leaving the user the choice which sectors to depict a la Melitz (cf. Yafari and Britz 2018). Products can be depicted as perfect substitutes according to a Spatial Equilibrium set-up (**SPE**) where per unit transport margins change depending on bi-lateral trade quantities. These options can be combined with a **MRIO** module where bi-lateral trade shares and tariffs differ for intermediate, private, public and savings demand. Further, a module working at the single **tariff-line** is available (Jafari et al. 2021). It comprises an extension where TRQs rents are

[3] We thank the FAO for letting us use and distribute this data set with the model.

[4] This extension was developed with support of the EU Commission’s Joint Research Center. See Ferrari, E., Roson, R. B., Britz, W., & Dudu, H. (2019). An extented myGTAP model to address subsistence production and sub-national households as a module in CGEBox, https://ageconsearch.umn.edu/record/333059/files/9236.pdf

optimized under TRQs open to multiple tariff-lines, importer and/or exporter nations by Döbeling (2022). Trade specifications can be applied to specific product markets, only.

The default for the **production functions**, depicted by CES nests which also support the volume preserving additive variant, is the GTAP Standard layout, i.e. substitution only takes place between primary factors. By changing the parameterisation, substitution between the value and the intermediate composite and inside the intermediate composite can be introduced. A **flexible nesting approach** allows introducing multiple nests which each combine primary factors, intermediates and other nests, also in shares. Factor and intermediates demand can be driven in shares by multiple nests, both the standard CES and the volume-preserving variants are supported. That is for instance used to mimic the nesting of the GTAP-E (Burniaux et al. 2002) model. Further nests are pre-defined, for instance, to reflect the structure of the GTAP-CE and GTAP-CM data bases. The activities can be broken down to two **vintages**, an old one using depreciated capital, assumed to be sector specific, and a new one using newly formed capital. The new one is assumed to have higher substitution elasticities between capital and other factors of production. The share parameters of the two old are updated dynamically in recursive-dynamic analysis to reflect the weighted average of old and new technology from the previous period.

A similar flexible nesting approach can be applied to supply, used for instance to replicate features of segmented factor markets from GTAP-AGR (Keeney and Hertel 2005). The code supports non-diagonal make structures based on flexible CET nesting structures, including the volume-preserving variant, applied for instance, to depict the composition of electricity generation following GTAP-Power (Peters 2016).

Production and factor use decision can be depicted at **sub-regional level** for selected regions with a CET/CES approach to distribute factors besides land to the activities and source total domestic demand from the regions. Matching data for that module are shipped for 280 administrative regions across Europe but the user can also supply shares – at the minimum for output by sector - to dis-aggregate other regions. Market clearing is still depicted at the national level and is bi-lateral trade between sub-national units.

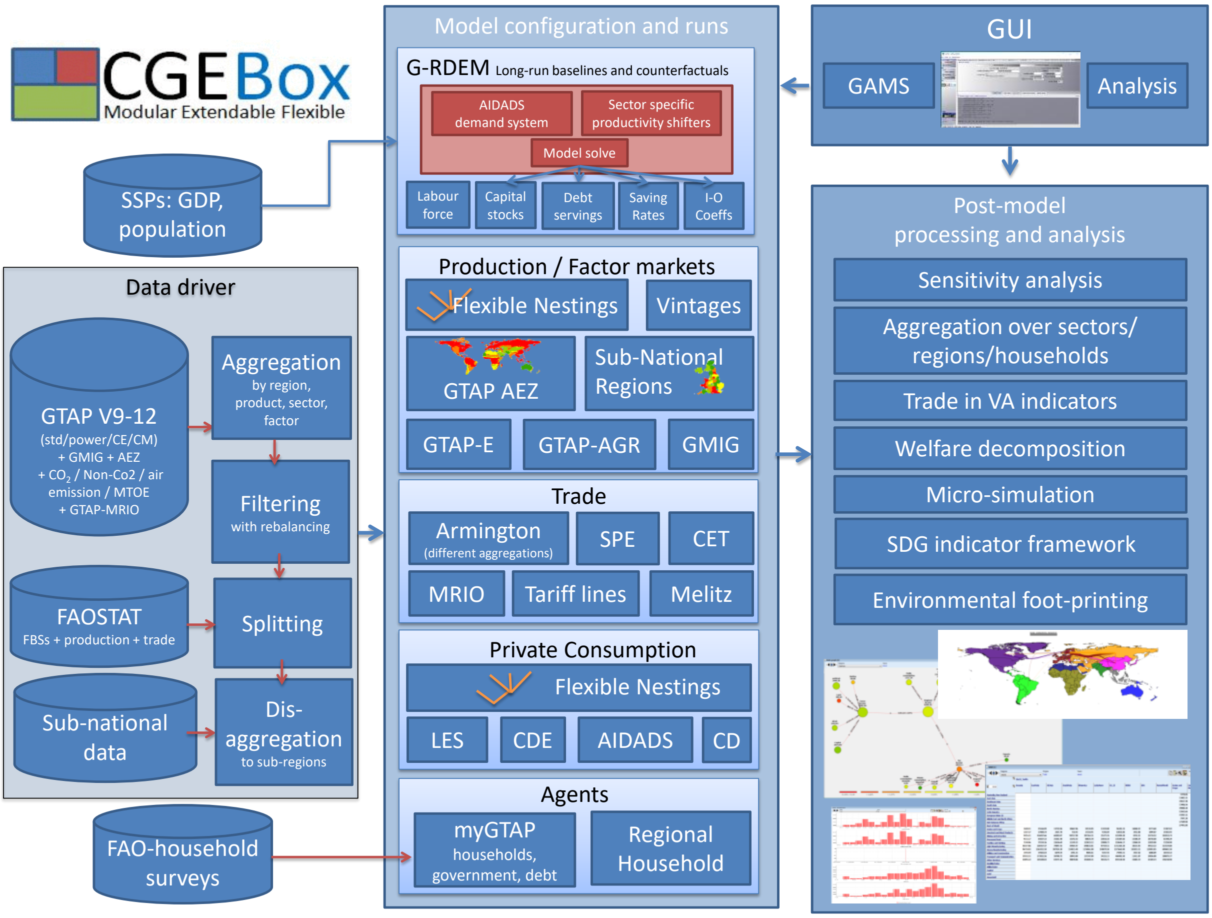
Land use can be modelled based on **GTAP-AEZ** (Lee 2005), i.e. total land in a region is divided into up to 18 so-called Agro-Ecological Zones in which a multi-level CET distributes available land in economic use to the production activities. For the top-level nests in the GTAP-AEZ module, volume-preserving CET functions are used to keep a physical land balance and to support consistent carbon stock accounting. Land supply from different types of unmanaged land can be added. The data driver and model can also introduce water as an additional factor of productions and its distinction between irrigated and non-irrigated agriculture[5], similar to the **GTAP-Water** data base by Haqiqi et al. 2016. The carbon accounting in GTAP-AEZ complements the **CO2 and non-CO2 emission** (Gibbs 2014) accounting framework developed by the GTAP center which can be used in CGEBox. Taxing schemes for CO2 equivalent emissions can be defined via the interface and the code supports the cap-and-trade system at national level or for a group of countries, and can render **process emissions factor endogenous based on marginal abatement cost curves**. CGEBox can also report **air pollution emissions** and **energy use in oil equivalents**, drawing on GTAP data.

Besides the usual capital accumulation in recursive-dynamic CGEs, the **G-RDEM** module (Britz and Roson 2019) implements features relevant for long-term baseline construction and counterfactual analysis: (1) the MAIDADS demand system mentioned above, (2) endogenous macro-saving rates,

[5] This data base was underlying the study on Food Waste by Britz et al. 2019a.

(3) sector specific productivity changes, (4) GDP per capita depending cost shares, (5) debt servings from past foreign savings. The four first features are based on econometric work; debt servings build on the global bank mechanism mentioned above. The freely available IIASA data base with projections of GDP, population by age class and education levels for the five Socio-Economic Pathways[6] is shipped ready-to-use with the G-RDEM module and can be combined with FAO projections on yields and crop land changes in baseline construction. It is complemented with a scenario driver that quantifies exogenous assumptions such as on carbon taxes, energy efficiencies, electrification, dietary changes or meat taxation (Goebel and Britz 2025).

It can be extended for base-line construction at the level of NUTS2 regions for Europe (Britz et al. 2019b). Equally, parameters in the demand system can be endogenously adjust to let total calorie intake for each household follow an empirically estimated relation to per capita income. The caloric content of food products can also be rendered endogenous based on an empirical relation to GDP per capita. A set a **Climate Change damages** from Roson and Sartori 2016 can be used assess macro-impact of such damages against given baselines.



The **data driver** of CGEBox combines the different data bases – a GTAP data base aggregated according to the user's choice, the land use data base, CO2 and NON-CO2, air emissions data, data on sub-national regions, gender differentiated labor data, households etc. – into a mutually consistent data set with a global SAM at its core. It supports the GTAP standard, power, CE (circular economy) and CM (critical mineral) data bases Version 9 – 12, and can add the AEZ and MRIO data. A separate download of GTAP-E is not necessary, the additional parameters for GTAP-E are comprised in the

[6] http://www.iiasa.ac.at/web/home/research/researchPrograms/Energy/SSP_Scenario_Database.html

code. The GTAP data can be filtered to remove small cost and expenditure shares. Before filtering, the user can introduce corrections to the data. A statistical outlier detection be used to remove suspicious costs share from the data. CGEBox comes along with a split-utility based on Linear Programming to dis-aggregate the global SAM and auxiliary data to higher detail for sectors and products (see Britz 2021). Data from FAO shipped with CGEBox provide the necessary data on global production, use and bi-lateral trade to split agri-food sectors to high detail (see Britz 2022). As the code supports non-diagonal make matrices, the split routine can also introduce multi-output activities such as oil-seed crushing to cake and oil or multiple activities producing one output, used for instance for irrigated and non-irrigated crop activities. Based on a FAO Food Balance Sheets data, **micro and macro nutrient contents** are estimated.

Post-model analysis

The reporting back-end provides a larger set of pre-defined reports as tables, maps or graphs for model analysis. Its concept is summarized by the following points[7]: (1) Each report combines a block of information helping to analyze specific aspects of the model solution. (2) This often implies looking simultaneously at different model variables, say quantities, pre- and post-tax prices or values which is cumbersome with the model listing or GAMS Studio. (2) The back-end therefore removes the technical disaggregation of the model's solution space into GAMS variables, parameters and equations. (3) Each report comprises at least the following information: (a) Which dimension (for instance, SAM row or column, scenario, year) is shown in columns and rows of the user's viewport, and which are moved to drop down lists. This information is called the pivot. (b) The selection of labels from each dimension*,* for instance if regions, sectors or products are shown from SAM's column dimension. (c) The view type (table, map, type of graph).

Further properties of reports can be added, such as the number of digits, if percentages changes should be additionally shown, if empty rows and columns should be hidden, etc. The user can change the pivot, view type, add statistics or filter out additional labels.

To enable this concept, variables, parameters and slacks in equations are copied into a n-dimensional parameters (region, item, SAM row, SAM column, version, time), where the item dimension covers inter alia: Q=quantity, P=Price, V=Value, T=tax rate, G=tax income, plus emissions (set em). The SAM row dimension also comprises additional variables not found in a global SAM constructed from the results, such as macro totals, cohorts, debt related variables, emissions, nests in the production or demand functions, parameters etc. The SAM column also comprises additional detail on such variables, such as for instance, emission sources.

When working with the reporting back-end, no knowledge of GAMS, exact equation structures or variable names is needed, as report use fewer technical terms that should be easily understood (such as supply, demand, price, imports, exports). The short labels used for regions, sectors or products are equally not needed as results are shown with their long-text descriptions in the reports. Analysis is supported by optional post-model aggregation to group of regions, products and sectors and **a de-composition of welfare impacts**. Various model results quantify an **SDG indicator framework** (Wilts and Britz 2022). Accordingly, courses with the model can analysis results even from complex model set-ups and scenarios as long as students understand the underlying model mechanisms (behavioral and other core assumptions, closures, properties of functional forms used etc.) without the need to memorize variable names or data base labels.

[7] See Britz et al., (2015).

Quality assurance

CGEBox is a large-scale software project and applies approaches from software quality management (Britz et al. 2026) to reduce as far as possible errors in the code. These approaches include the use of software versioning system to track changes in the code and allow a continuous integration of code improvements and corrections, adherence to coding guidelines, a larger test suite applying shocks to various model configurations and options to solve the model to reduce as far as possible errors in the current code base and manyfold in-built assertions in the code to trap errors and avoid error propagation. The release strategy is based on continuous integration.

User community and networking

CGEBox is a living project where new modules are added or existing ones refined on a more or less regular basis. Care is taken that the different modules are as far as possible mutually compatible. This modularity offers the user quite some freedom to structure a suitable CGE model for a specific application. CGEBox is an open project which invites others to contribute and share. The model is used, for instance, by the World Bank and the French government.

The team in Bonn organizes two-weekly on-line meetings outside the summer break with a group of users, mostly European PhD students and their supervisors, where projects with CGEBox are discussed, features demonstrated and CGE modelling papers discussed. A yearly block course of one week in spring offers PhD students and other interested parties an introduction in the model. Its content alternates bi-yearly between a focus on different modules under comparative-static analysis and recursive-dynamic analysis with a focus on G-RDEM.

References

This overview only provides some selected references which relate mostly to CGEBox itself. A more complete set of references can be found in the CGEBox model documentation.

Britz W. (2019): CGEBox–a flexible and modular toolkit for CGE Modelling with a GUI. Institute for Food Resource economics, version November 2019, see: http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox_GUI.pdf

Britz W. (2025): Britz, W. (2025). How to solve large recursive-dynamic Computable General Equilibrium models, https://ideas.repec.org/p/ags/ubfred/352149.html

Britz, W. (2014): A New Graphical User Interface Generator for Economic Models and its Comparison to Existing Approaches, German Journal of Agricultural Economics 63(4): 271-285

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

Britz, W. (2022): Disaggregating Agro-Food Sectors in the GTAP Data Base, Journal for Global Economic Analysis 7(1): 44-75

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

Britz, W., Dudu, H., Roson, R., Fusacchi, I., Jafari, Y., Salvatici, L., Sartori, M. (2019): Economy-wide analysis of food waste reductions and related costs, in: JRC technical reports, JRC113395

Britz, W., Jansson, T., & Schäfer, D. (2026). Reflections and recommendations on Software Quality Management in economic models. *Social Sciences & Humanities Open*, *14*, 10318

Britz, W., Pèrez Dominguez, I., Narayanan, G. B. (2015): Analyzing Results from Agricultural Large-scale Economic Simulation Models: Recent Progress and the Way Ahead, German Journal of Agricultural Economics 64(2): 107 – 119

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

Britz, W., Roson, R., Sartori, M. (2019b): SSP Long Run Scenarios for European NUTS2 Regions, in: Working Papers Department of Economics Ca' Foscari University of Venice, No. 22/WP/ 2019

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

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

Burniaux, J. M., Truong, T. P. (2002). GTAP-E: an energy-environmental version of the GTAP model. GTAP Technical Papers, 18. https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=923

Döbeling, T. (2022). Modeling the interactions of complex tariff regimes, https://ageconsearch.umn.edu/record/333462/files/11123.pdf

Drud, A. S. (1994). CONOPT—a large-scale GRG code. *ORSA Journal on computing*, *6*(2), 207-216

Gibbs, H., S. Yui and R. Plevin (2014). New estimates of soil and biomass carbon stocks for global economic models. GTAP Technical Paper No. 33. Centre for Global Trade analysis Project (https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=4344)

Goebel, T., & Britz, W. (2025). Global land use and land cover projections under the shared socio-economic pathways: An integrated computable general equilibrium analysis with sub-national resolution for Europe. *Global Environmental Change Advances*, *5*, 100027

Haqiqi, I., F. Taheripour, J. Liu and D. van der Mensbrugghe (2016). Introducing irrigation water into GTAP data base version 9. Journal of Global Economic Analysis, 1(2):116-155

Hertel, T. W., & Tsigas, M. E. (1997): Structure of GTAP. In: Hertel, T.W. (ed.), Global Trade Analysis: modeling and applications, 13-73, Cambridge University Press.

Jafari, Y., Britz, W. (2018): Modelling heterogeneous firms and non-tariff measures in free trade agreements using Computable General Equilibrium, Economic Modelling 73: 279-294

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

Keeney, R., Hertel, T. (2005): GTAP-AGR: A framework for assessing the implications of multilateral changes in agricultural policies. https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=1869

Lee, Huey-Lin (2005). Towards an integrated land use data base for assessing the potential for greenhouse gas mitigation. GTAP Technical Papers 26. https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=1839

Peters, J. C. (2016). The GTAP-power data base: disaggregating the electricity sector in the GTAP data base. *Journal of Global Economic Analysis*, *1*(1), 209-250

Roson, R., & Sartori, M. (2016). Estimation of climate change damage functions for 140 regions in the GTAP9 database. Journal of Global Economic Analysis 1(2): 78-115

Squarcina M. (2017): Small Famile Farm data Portrait, Basic Information Document, Methodology and Data description. FAO publication available at: http://www.fao.org/fileadmin/user_upload/smallholders_dataportrait/docs/Data_portrait_variables_description_new2.pdf

Van der Mensbrugghe, D. (2018): The Standard GTAP Model in GAMS, Version 7. Journal of Global Economic Analysis: 3(1), 1-83

Van der Mensbrugghe, D., & Peters, J. C. (2016). Volume preserving CES and CET formulations, https://ageconsearch.umn.edu/record/332784/files/8380.pdf

Walmsley, T.L., Winters, A., Ahmed A. (2011). The impact of the movement of labour: results from a model of bilateral migration flows. Global Economy Journal 11.4 (2011): 185-240

Wilts, R., & Britz, W. (2022). Quantifying SDG indicators for multiple SSPs up to 2050 with a focus on selected low and low-middle income countries and the bio-economy based on CGE analysis, GTAP conference paper, available at https://ageconsearch.umn.edu/record/333473/

Wilts, R., & Britz, W. (2024). Distribution matters: Long-term quantification of the Sustainable Development Goals with household detail for different socio-economic pathways. *Global Transitions*, *6*, 173-186