Economic Modelling 59 (2016) 463–472

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

Reducing unwanted consequences of aggregation in large-scale economic models - A systematic empirical evaluation with the GTAP model

Wolfgang Britz [a,*], Dominique van der Mensbrugghe [b]

[a] Institute for Food and Resource Economics, University Bonn, Germany
[b] Center for Global Trade Analysis (GTAP), Purdue University, West Lafayette, IN, United States

ARTICLE INFO

Article history:
Received 15 February 2016
Received in revised form 29 June 2016
Accepted 27 July 2016
Available online 27 August 2016

Keywords:
Computable general equilibrium analysis
Aggregation bias

ABSTRACT

We discuss how to avoid aggregation bias in large-scale global Computable General Equilibrium (CGE) models by reducing the need of pre-model aggregation, based on the combination of algorithmic improvements and a filtering approach which removes small transactions. Using large-scale sensitivity analysis, we show the impact of pre-aggregation and filtering on model size, model solution time and simulated welfare impacts, using a multi-lateral partial trade liberalization simulated with the standard GTAP model as the test case. We conclude that pre-model aggregation should be avoided as far as possible, and that our filtering approach and algorithmic improvements allow global CGE analysis even with highly disaggregated data sets at moderate solution times.

1. Background and introduction

Despite considerable progress in soft- and hardware, computational issues such as solution time, memory requirements or numerical stability continue to constrain the use of economic simulation models. Therefore, the resolution of available data bases is typically not fully exploited in the application of economic models such as computable general equilibrium (CGE), supply side and partial equilibrium models. Rather, researchers aggregate data bases e.g. with regard to commodity and regional detail in order to yield models which can be solved in a fast and stable manner. The widely used GTAP data base (Narayanan et al., 2012; Aguiar et al., 2016), to give an example, is shipped with the GTAPAgg aggregation tool (Horridge, 2006), and almost any published application of the GTAP model (Hertel, 1997) uses a study-specific pre-model aggregation of sectors and regions.[1] At the same time, it is well known that key model results such as welfare changes depend on the chosen aggregation level (cf. Ko and Britz, 2013). However, almost no publication presenting and analyzing model results comprises a section with a sensitivity analysis with regard to alternative pre-model aggregations.

Aggregation issues have clearly gained interest in the community of economic modelers, also as a consequence of availability of more detailed data. A growing body of literature deals e.g. with the aggregation of trade policy instruments in economic modeling (e.g. Pelikan and Brockmeier, 2008) both with examples of combining different model types (Grant et al., 2007) or using welfare-consistent aggregators (cf. Himics and Britz, 2015). So far, implementation of these approaches in larger modeling exercises is still scarce. And clearly, better aggregation of trade policy instruments can only improve one of the many possible sources of aggregation bias in modeling exercises. Our paper therefore adds a different perspective, beyond aggregation of policy wedges, by looking at impacts of pre-model aggregation. Additionally, it also presents and analyzes a complementary approach to pre-model aggregation based on filtering out small transactions while maintaining data consistency and important economic totals. We use large-scale sensitivity analysis to check how pre-model aggregation and our proposed filtering approach impact simulated results, but also model size and solution behavior. Besides complementing the discussion on appropriate aggregation in economic model, a further aim is to arrive at suggestions in applied CGE modeling helping practitioners in their daily work.

Our paper is structured as follows. The first section presents reasons for pre-aggregation, before we discuss possible approaches to reduce the need to do so. Our filtering algorithm is presented in Section 3, followed by an empirical evaluation. Finally, we summarize and conclude.

2. Reasons for pre-model aggregation of data bases and empirical evidence of its impacts

Pre-aggregation is a step before an actual application of the economic model which affects both its data base and parameterization. In

* Corresponding author.
E-mail address: wolfgang.britz@ilr.uni-bonn.de (W. Britz).

[1] This paper is focused on global economic databases and models where aggregation is most relevant. Single country databases and models typically have a resolution that can forego aggregation. Exceptions might be countries with highly detailed activities and commodities data bases such as the 500+ input-output table of the United States. Aggregation may also be useful for sub-accounts, for example a country database that contains a large household survey.

http://dx.doi.org/10.1016/j.econmod.2016.07.021
0264-9993/

applied trade modeling, pre-aggregation typically relates both to the commodity/sector and the regional dimension. Pre-aggregation can be understood as a multi-stage process. Most data are sampled originally at the level of individual agents and compiled by statistical offices to arrive at published totals. Typically, only these compiled totals are available to the research community; they suffer potentially from reporting and processing errors and might comprise systematic bias: observations might have been excluded due to data confidentiality issues or due to sampling thresholds, e.g. with regard to minimum firm size.

In order to make published totals suitable for modeling purposes, specialized teams next process them further, e.g. to render different data sources consistent with each other, to exclude outliers or to fill gaps. The efforts to release the GTAP data base provide many examples for such work such as the compilation of global bilateral trade flows and related data on trade protection levels. Most of the data sources used in that process need to be aggregated in order to arrive at the agreed upon regional and sector resolution of the final GTAP data base, covering 57 sectors and 134 regions in version 8 as used by us.[2]

The costs to duplicate such efforts in order to generate a differently detailed data set would be tremendous. The vast majority of economic modelers therefore rely on ready-to-use data sets such as the GTAP data base which are already pre-processed specifically for the purpose of economic modeling, but imply at the same time a specific maximal disaggregation of analysis. Generally, it would be desirable to exploit the full information of these data sets with regard to economic transactions, policy instruments and behavioral parameters, and increasingly also with regard to so-called satellites accounts with e.g. (bio)-physical information. Full information exploitation, however, requires running the model without any further aggregation which is often considered impossible for computational reasons – model size, solution time and stability – such that some form of pre-model aggregation is typically deemed unavoidable.

Besides reducing model size, which eases model solution and resulting analysis, pre-model aggregation has additional advantages. If errors in the data base are not or even negatively correlated across regions and sectors, aggregation will systematically improve data quality. Peaks in policy wedges as in cost and trade shares will be flattened which helps model solution. Aggregation will also systematically remove small entries such that the overall scaling of the model improves. It reduces the dimensionality of the simulated impacts which eases model analysis. If additional data are needed in a modeling exercise, e.g. environmental accounts, or different types of models are linked, pre-aggregation might also be necessary to arrive at common regional and product definitions.

Ideally, a selected pre-model aggregation should lead at the same time to an acceptable solution behavior and small aggregation bias, i.e. differences in simulated results compared to the model used in conjunction with the data base in full resolution. Unfortunately, as the full-fledged model can typically not be solved–at least not at acceptable cost–, the aggregation bias is not known. Accordingly, the researcher has to make an educated guess which aggregation by regions and sectors will cause a still acceptable bias, e.g. by combining sectors and regions with a similar cost structure or protection level. However, aggregation clearly offers quite limited degrees of freedom in that respect. To give an example: a set of regions belonging to one aggregate might have comparable protection rates in certain sectors and highly divergent ones in others. The flexibility in selecting the aggregation is further reduced if pre-model aggregation serves additional purposes such as to condense model results for easier analysis and presentation, e.g. by aggregating regions to continents or political blocks.[3]

Clearly, study-specific, flexible pre-aggregation is only possible if the model is based on template equations which are structurally identical, i.e. do not vary in structure across regions and sectors, such that differences are solely depicted by parameterization. Accordingly, differently aggregated data sets change only the number of equations and the model's parameterization, but not its structure. The remainder of the paper thus relates only to that rather common type of economic model. Basically all CGE models are based on equation templates, but also Multi-Commodity models such as IMPACT (Rosegrant et al., 2008), the COSIMO part of AGLINK-COSIMO (OECD, 2007), CAPRI (Britz and Witzke, 2014) or ESIM (Banse et al., 2005) from the field of agricultural economics. That does not imply that all these models could be easily run at a more aggregated or dis-aggregated regional and or commodity resolution of their current data bases. Aggregation e.g. of policy instruments or behavioral parameters might be a demanding task, and computer code to do so are not typically available. Furthermore, the model's computer code might not be easily changed to work with different lists of products and regions.

There is ample evidence of unwanted impacts of using less detailed data bases in CGE modeling. Caron, 2012 finds larger differences in estimated international carbon leakage effects based on same CGE model structure using a standard GTAP data base compared to one with more detail for industrial sectors. Alexeeva-Talebi et al. (2012) in a similar exercise focus more on the potential loss of detailed information when using more aggregated data bases, but also report aggregation bias at the aggregate level. Antimiani et al., 2015 focus on the sensitivity of results to energy-related elasticities in CGEs which also depend on the aggregation level. Comparing two different sector aggregations and partial against general equilibrium closures, Brockmeier and Bektasoglu, 2014 conclude that sectoral detail can matter substantially and turns out more important in their analysis than using a GE layout. Ko and Britz, 2013 report larger changes in welfare gains from a EU-Korean free trade agreement when only the regional detail for the European Union is changed. All these findings suggest that at least certain types of pre-model aggregation can indeed systematically affect simulated results.

When looking at the reasons to pre-aggregate, only availability of additional data needed for a modeling exercise or model linkage might be a strict constraint, while computational barriers might be overcome. Indeed, given the impact of pre-aggregation on simulated results (see Section 5.3), studies could be challenged for choosing an aggregation which delivers desired results, for instances larger impacts for specific regions and sectors.

3. Model setup and algorithmic improvements

Model solution behavior can clearly be improved by algorithmic changes. We discuss here some of the steps taken by us in that respect. But before, we briefly present our test framework.

In the following, we use a GAMS (Brooke et al., 1988) implementation of the well-known GTAP Standard Model (Hertel, 1997). GAMS is a widely used language in economic numerical simulation (cf. Britz and Kallrath, 2012), and the implementation by Van der Mensbrugghe and Britz (2015) used here offers those researchers who prefer GAMS over GEMPACK (Harrison and Pearson, 1996) a fully compatible and tested version of the Standard GTAP Model.

The standard GTAP model and extensions thereof have been extensively used since the early 1990s for a wide range of different research and policy questions. The default in its application is a pre-model aggregation of the GTAP data base. Using the combination of this model and

[2] Issues of creating a globally consistent snapshot of the global economy go beyond simply reconciling official data from various sources (e.g. national statistical agencies, COMTRADE, IMF's Balance of Payments, etc.), but also reconciling national IO tables that are provided with different base years and in different currencies. To give an example, most recent benchmark input-output account of the U.S. economy for the year 2007 (released in Dec. 2013) has 389 industries, http://www.bea.gov/industry/io_annual.htm). All are adjusted to a common year and denominated in a common currency.

[3] The GAMS code of the model used by us comprises the possibility for post-model aggregation of the simulation results to user defined list of sectors and regional aggregates such that pre-model aggregation whose main purpose is to ease analysis of simulated results is not necessary.

data base is therefore an obvious candidate for a systematic analysis of the impacts of pre-model aggregation. Furthermore, the standard GTAP model's basic features are quite common in global CGE analysis. On the production side, a Leontief technology for intermediate input demand and total value added demand is combined with a CES function substituting between primary factors under the assumption of constant returns to scale. On the demand side, substitution between domestic and imported goods in demand is modeled by a two-stage Armington structure. Markets are assumed to be competitive, and policies are expressed by relative price wedges. Different assumptions with regard to factor mobility are possible; we opt here for full factor mobility with the exemption of natural resources which are considered immobile. Further detail can be found in Hertel (1997); a documentation of the individual equations as implemented in the GAMS version provide van der Mensbrugghe and Britz (2015).

We solve the model as a constrained system of non-linear equations (CNS), using CONOPT (Drud, 1994) as the solver. A solution as a Mixed Complementarity Problem (MCP) is also supported by our GAMS code, but has shown to somewhat slow down model solution times. As we are not exploiting MCP features in our modeling exercise, we therefore only analyze the more standard CNS case where endogenous variables stay away from their bounds and equalities hold in the final solution. The model's equations are completely written in levels, and not as in the widely used GEMPACK version of the standard GTAP model as a mix of log-linearized and level equations.

In order to speed up solution time, we have firstly developed a sequential pre-solve algorithm which solves all single region models several times at iteratively updated international prices before the full globally linked model. During these pre-solves, only one region is considered at a time with bi-lateral import prices fixed. Equally, the single country faces fixed aggregate import demand from their trading partners at fixed import prices of all other regions. Accordingly, bilateral export demand for the region under consideration is reacting to changes in that country's fob prices due to updates of the import shares at the export destinations. Once all regional models had been processed several times in parallel using the grid solve algorithm from GAMS[4] at updated international prices, the full model is solved. That procedure speeds up overall solution time dramatically for highly dis-aggregated model versions with many sectors and regions as it exploits knowledge about strong and weak relations between blocks of equations and variables in the overall global model which cannot be easily detected by the solver. It typically provides a very good starting point for the full model with very little infeasibilities left. For models with limited sectoral and regional detail, the pre-solves generate unnecessary overhead and can therefore be switched off.

Equally, we substitute out basically all linear definitional relations between variables in the model using the macro facility of GAMS.[5] That reduces especially model set-up time and memory use, but might have a limited impact on time needed by the solver depending on the solution algorithm used. CONOPT, to give an example, will detect definitional equations and treat them differently. Unfortunately, substitution of equations might affect solution behavior negatively, as the solver might introduce tiny infeasibilities at or below the feasibility tolerance in equations otherwise substituted out which relaxes the solution space. In our tests, a negative impact on solution time or stability from substitution could, however, not be observed. In order to reduce the impact of differently scaled economic transactions depicted by the elements of the aggregated SAMs and represented by variables in the model, we programmed code which defines scaling factors for all equations in the model, using the bigger of the value of the related transactions and a rather tiny threshold. That should ensure that smaller regions and sectors are solved towards a similar relative accuracy as larger ones, and render the model better suited to work with more disaggregated data bases where differences in variable levels tend to be larger.

Another reason why researchers pre-aggregate is to ease results analysis (cf. Britz et al., 2015). Here, our code implementation covers additional flexible post-model aggregation (Britz and van der Mensbrugghe, 2015) which aggregates over regions and sectors such that result analysis can be conducted either at the detail with which the actual model run is performed, or at a more aggregate, user chosen level. The post-model aggregation definitions are entered via the GTAPAgg interface of which the steering file is read by the GUI. As such, the same tool used to pre-aggregate the data is applied to select an appropriate post-model aggregation.

Finally, we use an implementation of CONOPT called CONOPT4 which has been updated to take advantage of the larger amount of memory on modern computers such that memory is no longer a practical limitation. In addition, CONOPT4 can use parallel execution. On demand, the inversion of the constraint matrix, the most expensive part of solving large CNS models, is to a large extent parallelized. Furthermore, the evaluation of the equations and their derivatives can be done in parallel using a new feature in GAMS. It should be mentioned that the parallel feature in CONOPT4 is only used for the large and expensive overall model (where it is most needed); the smaller regional models in the pre-solve algorithm use each one thread only as they are already run in parallel using the grid solve approach.

4. Filtering as a data driven alternative to pre-aggregation

4.1. Overview and motivation

An alternative or complement to an informed, research driven decision on a specific pre-model regional and sector aggregation is a data driven approach. We propose and present here an algorithm which aggregates small transactions in global SAMs to larger ones in a flexible manner. It combines three elements:

1. Filtering out small economic transactions from the data base
2. Rebalancing of the thinned out regional SAMs and bi-lateral trade flows
3. Maintaining closely the sum of important transactions such as GDP

This approach[6] can be applied to global SAMs with different regional and sectoral detail and thus can be combined with pre-model aggregation as discussed above. A filtering algorithm based on (1) and (2) has been distributed for some years to complement the GTAPinGAMS model (Rutherford and Harbor, 2005), but we add here additional elements.[7]

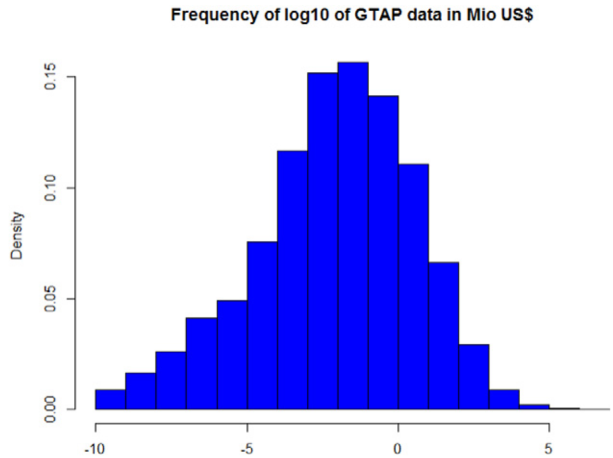
Graph 1 below renders it obvious why it might pay off to remove small transactions: a large share of the about 2.2 million non-zero transactions found in the GTAP 8 data base is quite small. As most users of the data base, we cannot tell if these tiny values are originally observed in

[4] We use the release 24.7 of GAMS which allows for in-memory parallel execution of several model instances as the central part of its grid solve approach. Earlier versions handled the instances via scratch directories which slows downs parallel processing compared to the new approach especially in the case that solution time for each instance has a small share in the overall handling of the model instances.

[5] This substitution option likewise is available in the GEMPACK software (Harrison et al., 1994), which allows for the user to identify both the variables to be substituted out and the equation that can be used to as a substitute for the variable. In our GAMS based model, the substitution of variables based on macros can be switched off from the interface to yield an equation listing comprising all variables. If substitution is used, we calculate post-model the solution values for the substituted out variables to ease result analysis. Using macros has the advantage to maintain readability: the macro's name is chosen to be almost identical to the original variable's name, for instance m_pefob(r,i,rp,t) instead of pefob(r,i,rp,t), such that readability of the equations is maintained.

[6] While the filtering approach is seamlessly integrated in the data processing step of our package, its results can also be used easily by other models drawing on the GTAP data base as the filter algorithm works on the matrices as comprised in the output from GTAPAGG and thus in the original data base. As such, even a conversion back into GEMPACK format is straightforward.

[7] The authors would like to thank Tom Rutherford not only for letting them use his filtering code as a starting point for the discussed extensions, but also for helpful feedback.



Source: Own compilation

Graph 1. Histogram with distribution of size of transactions in original SAMS/trade flows.

official statistics, based on expert assumptions such as using a IO-table from a neighboring country when constructing a SAM or for instance introduced by some balancing algorithm. Changes in really tiny transactions such as accounting only for a fraction of a dollar should not have a significant impact e.g. on simulated welfare impacts, but might affect solution behavior negatively, be it by simply driving up solution time and memory needs or even worse, by affecting solution stability. We quantify these impacts below in Section 4 and 5 for a wider range of test simulations.

The data driven filtering approach will remove potentially a large share of these tiny values. Consequently, less aggressive pre-model aggregation is needed in order to arrive at models which can be solved in an acceptable time. That should systematically reduce the aggregation bias especially resulting from flattening out the distribution of policy wedges, as we will show in the empirical part of the paper. Equally, removing tiny values can reduce computational difficulties. However, we would remind the reader that the functional forms typically used in CGE modeling imply that zeros remain zeros, i.e. removing a value in the benchmark will also remove it under any shock. If a country or sector is small, simply deleting entries based on absolute size thresholds runs the risk of removing valuable information. As discussed next, the filtering algorithm therefore removes data based on thresholds relative to totals chosen to avoid e.g. bias towards small countries. Furthermore, we should note here that the user interface available with the GTAP model in GAMS (Britz and van der Mensbrugghe, 2015) which also steers the filtering algorithm allows setting specific thresholds for selected regions and sectors during filtering. That permits limiting or eliminating completely the filtering process for exemptions which are of special interest in the analysis, not at least to avoid for them the "zero stays zero" effect.

4.2. Algorithmic details of the data driven aggregation process

As indicated above, the algorithm consists of three major parts: (1) filtering, (2) re-balancing and (3) data driven aggregation, now discussed in some detail.

The filtering process tries to remove small items from the regional SAMs and from transactions relating to bi-lateral trade. The definition of what is small is based on a comparison relative to specific regional totals. In particular, domestic or imported intermediate demands of a commodity by a sector in a region are measured against total regional intermediate demand of that sector. Demand by the government, for investment or by households for domestic produce or imports of a commodity in a region are compared against the respective total domestic or import demand. Trade flows of a product are dropped if both shares on total exports of that product by the exporter and shares on imports of that product by the importer are below the relative threshold. Production is flagged as to be removed if net production of a commodity in a region, i.e. after intermediate use of that commodity in its own production is deducted, is below the relative threshold with regard to total regional net production, or if both domestic and export demand have been flagged as deleted.

After introducing the flags which indicate data entries to remove, the regional SAMs are re-balanced, one region at the time. In order to ensure that bi-lateral imports and exports match, import and export trade flows have to be fixed during the re-balancing step of the individual regional SAMs. The balancing constraints for each regional SAM comprise three types of equalities. The first one ensures revenue exhaustion for each sector, i.e. the value of domestic and export sales net of output taxes of that sector must be equal to its production cost, i.e. the value of domestic and imported intermediates plus primary factor costs, taking into account any taxes. The second constraint refers to the domestic market: total domestic sales are equal to final, export and intermediate demand. The last of the balancing constraints is equally structured and relates to the import market. Additional constraints ensure that factor demands and total intermediate demand are not dropped if production remains non-zero.

Besides the fixed exports and imports, all other discussed elements are endogenous variables in the re-balancing problem which enter its objective function. Elements flagged as deleted receive penalty terms with a high weight to favor sparsity. The remaining elements are drawn towards the original data based on relative quadratic difference compared to the initial value. Using relative differences instead of absolute ones has shown to give similar results for larger transactions, but to improve the outcome for smaller ones: cost, consumption and tax shares tend to stay closer to the initial ones.

Without further modification, the discussed algorithm would have very little to do with aggregation issues per se. The process removes small data entries, such that the regional economies and international trade would systematically shrink. The algorithm is therefore expanded by two further elements. The first element relates to trade flows. Total trade is scaled to maintain the original world totals, and import and export tariffs are scaled to recover each country's original protection. That is achieved by first deriving a bi-lateral scaling factor which takes the change in total imports and exports for each country into account. It is applied to the individual trade flows. Afterwards, the corrected trade flows for each commodity are scaled to match the original world totals and tax rates on imports and exports are re-adjusted to match the original bi-lateral country totals of import and export tax revenues which will tend to flatten the distribution of related tax rates. Finally, transport margins are updated to recover the original transport demand per commodity and region. That process is applied before the actual SAM re-balancing, as trade related variables are fixed during the re-balancing step of the individual regional SAMs. The second element which contributes to maintain totals comprises constraints defining important totals such as GDP and total use of each primary factor. Deviations from these totals enter the objective function with a high weight.

These two additional elements imply a kind of data driven aggregation: the economic value of the dropped small transactions is not lost, but added to the remaining, larger transactions of the same type. The relative quadratic deviations in the objective function should distribute them rather equally in relative terms to these larger ones. As one of the totals in the objective function is total government consumption, the algorithm will also maintain closely total tax income which implies a more even distribution of tax rates compared to the unfiltered case. The algorithm is implemented in GAMS and uses CONOPT as the solver.

The relative thresholds which determine the transactions to be dropped are introduced step-wise over iterations for two reasons. Firstly, as non-dropped transactions are scaled upwards, some transaction might become larger than the thresholds over these iterations. And secondly, perhaps more importantly, it allows checking for a desired number of non-zero transactions and thus to control more easily the resulting model size.

4.3. Example findings from filtering

This section shows the impact of filtering on the GTAP data using a selected pre-aggregated data base. We opted to maintain the full sectoral resolution (i.e. 57 sectors) and tested different regional aggregations in the following. Generally, the largest reduction in the number of data entries results from removing tiny intermediate demand transactions and bi-lateral trade flows.

Table 1 below gives an illustrative example from filtering a 57 sector × 56 region aggregation, i.e. the full sectoral detail of the GTAP data base is maintained. The unfiltered data based has around 800,000 non-zero transactions. It is first worthwhile to note that the regional SAMs and even the bilateral trade flows are almost completely dense in the starting point, i.e. after pre-aggregation of the GTAP data: the non-zero final demands are equal to number of sectors times the number of regions. There is an additional product (the trade margin) which enters the calculation, which explains why the amount of non-zero domestic output entries is somewhat higher. Also the intermediate demand matrices are completely dense. Even the trade matrix has with about 175,000 non-zero entries almost the full potential size of about 180,000.

Already a quite tiny threshold of 0.01% (itr2) relative to the respective total (see Section 4.2 above for their definitions) removes about one third of all transactions, with the largest absolute reduction achieved by thinning out intermediate demand. That alone removes about 200,000 entries, another 40,000 entries are taken from the bilateral trade flows. Government demand entries are the most affected in relative terms. The final threshold of 0.1% used in our example removes about half of the transactions. As discussed below, the impact on simulated welfare changes is often quite small.

The user can run the filter program once with a rather aggressive threshold, say 5% and many steps (up to 30 are supported) to analyze the relation between the filtering threshold and the number of non-zero transaction left, but also to check which agents are affected most. The program also generates a table with details for individual sectors/commodities, not shown above. That allows for an informed decision which relative and/or absolute threshold might be deemed acceptable. The empirical part following next shows impacts on model solution and solving time both from pre-aggregating and filtering, but also on welfare impacts in for selected shocks. These findings can further help to decide on a suitable combination of a specific study. The table produced by the utility also comprises information on changes per region. As the algorithm defines the thresholds for transactions to remove relative to regional totals, differences between regions with regard to the relative reduction in the number of non-zero transactions are typically quite small and not reported here.

5. Empirical evaluation

Our empirical evaluation is based on a typical counterfactual analysis with global CGEs, a partial multi-lateral trade liberalization, and runs it against a benchmark based on aggregating and filtering from the same global GTAP 8 data set. The analysis of the outcomes is split into two sections. First, we show how data-driven filtering impacts model size and solution behavior under different shocks, while the Section 2 analyzes how pre-model aggregation and filtering impacts simulated welfare changes.

5.1. The test framework

We used for a test a two-step aggregation/filtering setup: first, the original data base is aggregated to a pre-defined level of regional detail using GTAPAgg, and next the filtering/aggregation algorithm is applied. Specifically, we generated data sets at the full 57 sector resolution with regional aggregates of 10, 16, 24, 36, 45, 57, 68, 82 and the full size of the data base with 134 regions, using mostly population size as a guideline to select countries to single out in more dis-aggregated versions. We used filtering on these different aggregated versions of the GTAP data base to derive data sets between 60,000 and 400,000 transactions in steps of 20,000, plus four sets with around 500,000, 600,000, 700,000 and 800,000 transactions. Clearly, not all data sets are available in all resolutions: some data sets are already quite small to start with, and others would require very aggressive filtering thresholds to reduce the number of transactions below a certain size. For the full data base, we considered data sets with around 2 million and 1.6 million transactions beside the unfiltered case with around 2.2 million non-zero entries.

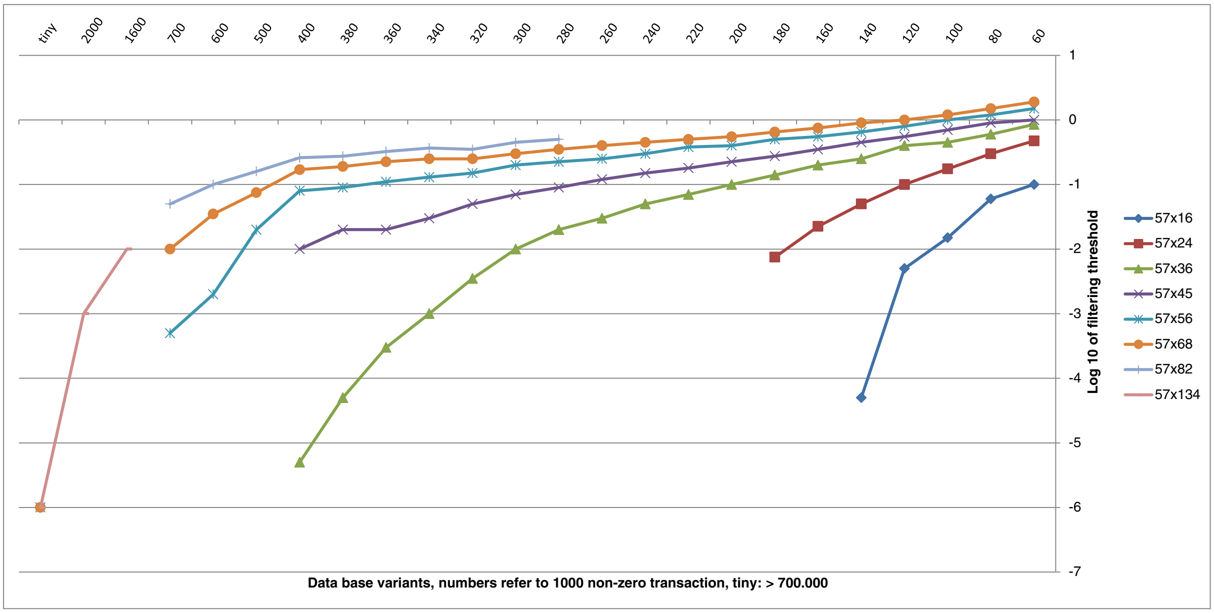
The actual active thresholds can be found in Graph 2. Each colored curve represents one level of pre-aggregation. As the relative filtering threshold is increased (the vertical axis), the size of the resulting dataset in thousands shrinks (the horizontal axis). To give an example: the purple line for the 57 × 45 case shows that a 0.1% filtering threshold (log10 is −1) reduces the size of the global data base from around 400,000 to about 280,000 non-empty transactions. Using 1% instead shrinks the data base to 60,000 non-zero transactions.

In order to generate our test data sets, we changed step-wise the desired maximal size of the non-zero transactions and let the program for each desired maximum iteratively increase the relative threshold until a data set was left with not more than the desired number of non-zeros. The relative filtering thresholds shown in the graphic below have hence been determined automatically by our algorithm. We reduced however the search time for the algorithm by first running a rough scan with a large relative threshold for each data set (an example is shown in Table 1) to have an indication about maximum relative thresholds matching our different desired data base sizes in order to avoid too many iterations. Even if generating each data set requires maximally some minutes, repeating the exercise for our 90 examples made that preliminary check worthwhile. At the same time, restricting the spread of the relative search range leads to more fine grained

Table 1
Example for impact of filtering on the number of transaction.

	Start	itr2	itr4	itr6	itr8	Done	Delta	Delta (%)
All transactions	816,152	547,377	479,927	438,547	405,561	390,097	−426,055	−52,2
Domestic output	3248	3045	2947	2867	2785	2743	−505	−15,55
Factor demand	14,390	13,651	13,302	13,024	12,724	12,576	−1814	−12,61
Bilateral trade flows	175,672	137,990	121,088	111,482	103,885	100,013	−75,659	−43,07
Trade margins	239,802	221,940	204,039	189,384	177,150	171,117	−68,685	−28,64
Import demand sectors	185,136	68,134	52,177	44,233	38,520	36,139	−148,997	−80,48
Domestic demand sectors	185,136	96,129	80,586	72,085	65,246	62,367	−122,769	−66,31
Import demand households	3192	2335	2177	2085	2023	1995	−1197	−37,5
Domestic demand households	3192	2529	2335	2241	2160	2123	−1069	−33,49
Import demand government	3192	958	753	676	627	602	−2590	−81,14
Domestic demand government	3192	666	523	470	441	422	−2770	−86,78
Relative tolerance (%)		0,01	0,03	0,05	0,07	0,08		

Source: Own compilation. Note: uneven iterations removed to increase readability.



Source: Own compilation

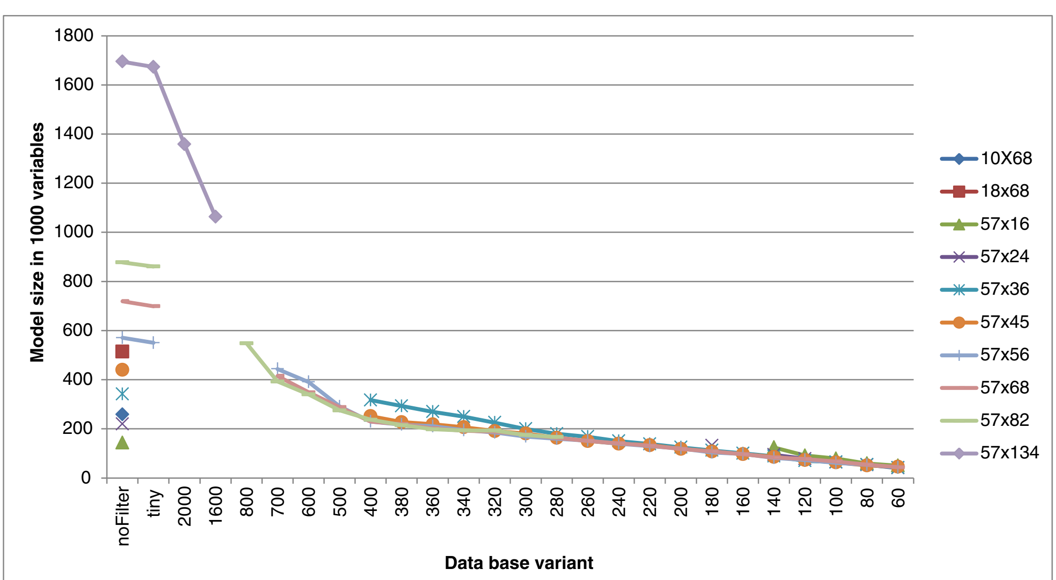
Graph 2. Data base size in 1000 and related filtering thresholds in log10%.

changes in the number of non-zeros and thus to data bases sizes closer to our desired number of non-zeros.

Where possible, also an unfiltered version of the pre-aggregated data base was used, and additionally one with a very tiny filtering threshold of 1.E-6% labelled with “tiny” in the graphs. In total, we considered about 90 different variants of the same structural model layout which solely differ in the data bases, and almost all have the full sectoral detail of the GTAP data base.

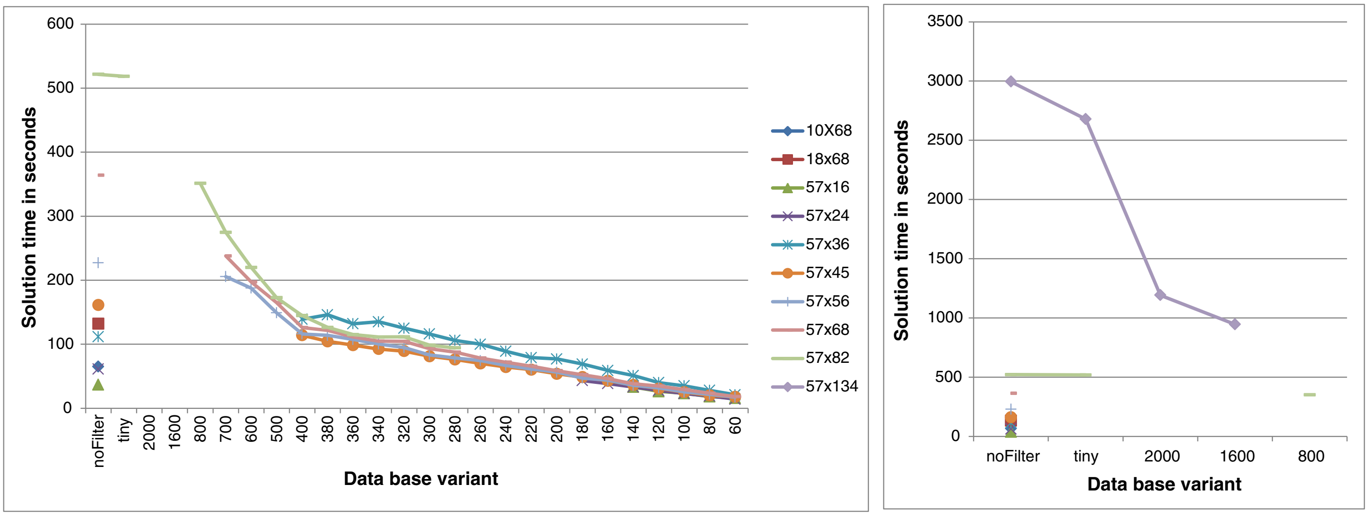
The reader should note that the pre-aggregation process in GTAPAgg also automatically aggregates the behavioral parameters. Maintaining the full sector breakdown of the GTAP data base in most experiments has the advantage of limiting aggregation bias due to sector-specific parameter aggregation, as key parameters such as the substitution elasticities used on the CES production functions and the Armington nests are sector-specific, but identical across regions. For comparison, we also run tests with two exemptions: 10 and 18 sector resolutions combined with 68 regions respectively regional aggregates. These two additional data sets are not filtered.

Graph 3 above shows the model size - measured as the number of non-fixed variables in the model which is equal to the number of equations - as a function of the chosen level of pre-aggregation and filtering. The first interesting observation is that the number of variables in the model is generally smaller than the number of non-zero transactions in the data base: the variants with 160,000 non-zero transactions (horizontal axis) lead to models with about 100,000 variables (vertical axis). That can be explained by the fact, as mentioned above, that purely linear definitional equations such as those defining bi-lateral import prices, import and domestic prices for the agents and the Armington demand quantities for intermediates are substituted out of the model. Furthermore, transactions linked to taxing economic activities are typically substituted out



Source: Own compilation

Graph 3. Relation between model size (in 1000 equations), pre-aggregation and filtering.



Source: Own compilation

Graph 4. Solution time in seconds with differently detailed data bases and model configurations.

from the model. To give an example: instead of having a variable which captures the value of taxing individual bi-lateral imports, that transaction is depicted in the model equations by multiplying the bi-lateral trade flow quantity variable with bi-lateral import price variable plus the fixed tax rate. These substitutions are obviously offsetting the effect from splitting the transaction from the data base into a quantity and price index plus introducing additional variables, e.g. to define total intermediate inputs and total value added and related prices for each sector.

Graph 3 suggests a more or less linear relation between the number of SAM entries and model size, rather independent from the chosen pre-aggregation. For the larger regional aggregations, quite aggressive relative thresholds between 0.5% and 2% are necessary to yield smaller models such that a huge share of the original transactions needs to be deleted (see Graph 2).

To give an example: in order to solve the 57 × 82 case in less than 2 min for our test shock (see next section), it is necessary to generate a data set with less than 300,000 non-zero entries. As the unfiltered 57 × 82 case has close to 900,000 equations, that would require removing a lot of detail from the data base and might lead to clearly unwanted structural changes e.g. with regard to important cost and trade shares. Such aggressive filtering is clearly not recommended for any serious model application, and as discussed in the next section, also not necessary.

5.2. Model solution behavior depending on filtering and aggregation

In order to test model behavior, we have selected a typical simulation exercise for global CGE models: a multi-lateral trade liberalization scenario which cuts all export subsidies and import taxes by 50%.

Graph 4 below shows the solution times in seconds[8] for our multi-lateral trade shock. Overall, model solution times are quite modest even for rather large models. Results without using the pre-solves and the older CONOPT3 version are discussed below. As seen, the maximal time needed in our test was around 1 h for the unfiltered full data base, dropping to around 20 min once the full data base was filtered down to 2 million transactions and to 15 min with 1.6 million transactions. However, we found by testing different shocks that solution behavior on models of that size is still unstable and numerical convergence might fail.

All other models could be solved in less than 10 min. An interesting observation is the fact that once the filtered versions come close to the unfiltered ones, the almost linear relation between the number of transactions and model solution time is broken. That effect can be seen for the 57 × 36 case at around 340,000 transactions, and for the 57 × 45 variant even around 320,000 transactions.

The most important observation is clearly that quite moderate filtering on the biggest models pushes solution time below 5 min. We also tested various other shocks (50% in consumption taxes, factor taxes, direct taxes; 20% increases in factor endowment and 10% increases in total factor productivity) on a subset of our input data, and found roughly similar solution times as for the multi-lateral trade shock. We thus tend to conclude that there are no computational reasons for high pre-model aggregations in analysis with the GTAP standard model or variants thereof[9] as long as not a very large set of runs is performed. Clearly, if one performs large-scale sensitivity experiments e.g. with behavioral parameters, solution times for larger models might still be considered too long.

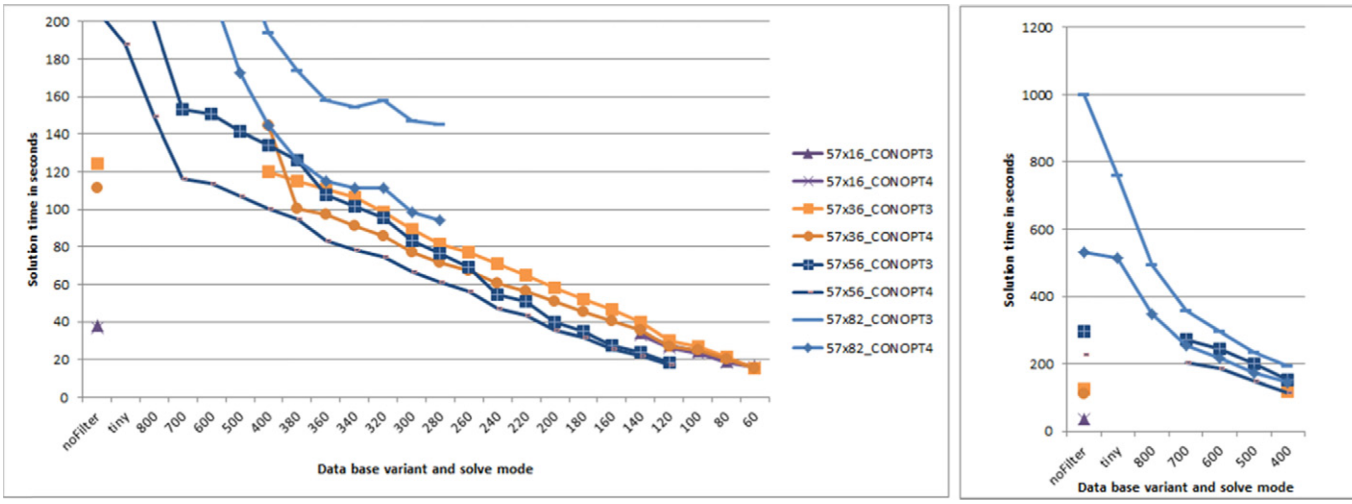
In order to shed light on the contribution of the improved CONOPT version, we show in Graph 5 above the results obtained if the pre-solve algorithm is switched on and the two solvers are compared for selected cases. The share of the full model solution time on total run time increases the larger the model. Thus, the impact of a more performing solver becomes more visible the larger the model. Whereas differences are negligible for smaller models, the improved CONOPT4 version can reduce overall run time by up to 40% for larger models, despite the fact CONOPT3 was used for the pre-solves. The reader should note that solution times larger than 1000 s are cut off.

Next, we turn to the impact of the pre-solve algorithm. For these tests, CONOPT4 was used for full model solves and CONOPT3 for pre-solves. Graph 6 below indicates that the pre-solve algorithm about

---

[8] The experiments were run independent from each other on a multi-core computing server with Xenon 3 GHz cores and fast hard drives. Times reported refer to cold start with GAMS, i.e. not using any restart files, and encompasses all steps until variables and equations were stored back to disk, including post-model processing to re-calculate a SAM, but not the more demanding data transformation required for GUI exploitation tools which encompass for instance aggregations over sector and commodities and welfare decompositions. Each run always includes a trial solve of the benchmark case to check that the model is correctly set-up. The same experiment with a 10 × 10 model without pre-solves takes less than 2 s on that machine. Run times on a modern laptop should be around two to four times as much, as long as the machine features enough memory and disk space. We used three iterations with the pre-solve algorithm and relaxed somewhat the feasibilities tolerances in CONOPT, but still above what seemed the default in the GEMPACK solver. The reader should note again that all equations are scaled with equation specific scaling factors, typically depending on the LHS in the benchmark such that the infeasibility tolerances reflect the logical structure of the model.

[9] We also ran similar tests where we added simultaneously an implementation of GTAP-AGR and GTAP-AEZ.



Source: Own compilation

Graph 5. Solution time in seconds with differently detailed data bases and model configurations for CONOPT3 and CONOPT4.

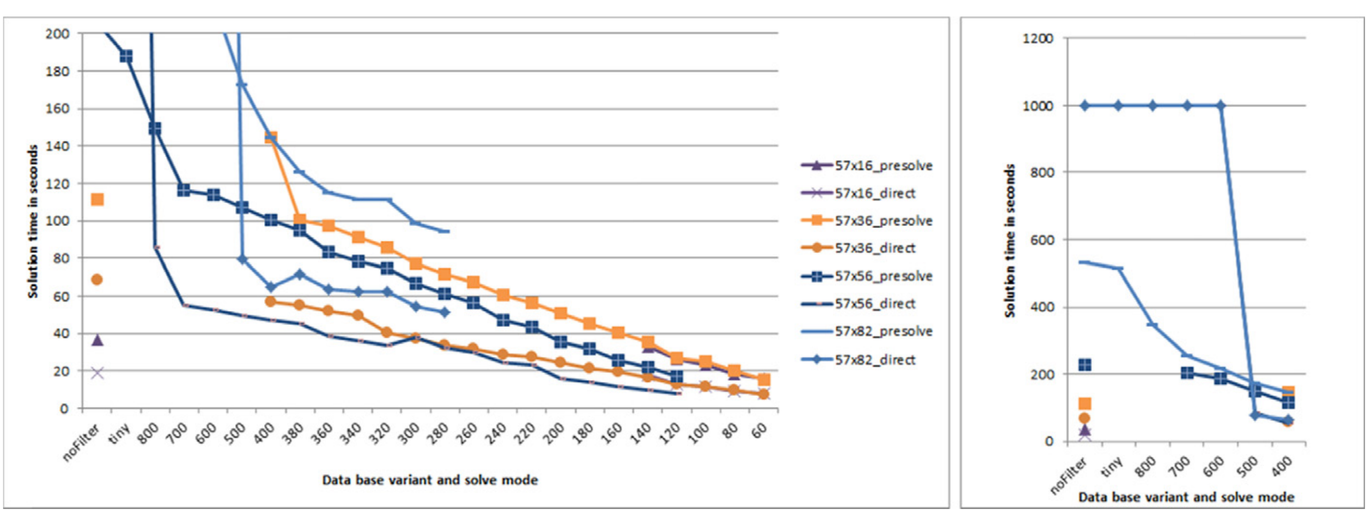
doubles solution times for small to medium sized models. That was the case for all variants up to the 57 × 36 pre-aggregation. Beyond that hard to predict point, dropping the pre-solves can increase solution times dramatically, as found for the 57 × 56 and 57 × 82 case beyond 500,000 non-zeros in the SAM. In some of these cases, the solver was no longer able to solve the model in any reasonable time if the pre-solves are not used. If solution time is not a core issue, the pre-solve algorithm should hence be switched on if the SAM comprise more than around 300,000 non-zero entries. Overall, our findings with regard to solution times underline the tremendous technical progress in hard- and software which shifts frontiers in quantitative modeling.

The interesting question here is naturally to what extent simulated results depend on the aggregation and potential filtering of the data base. Do we really gain a lot by using models which are harder to solve due to more dis-aggregated data bases which might potentially only add a lot of rather tiny values? We will shed some light on that question in the next section.

5.3. Simulated welfare impacts

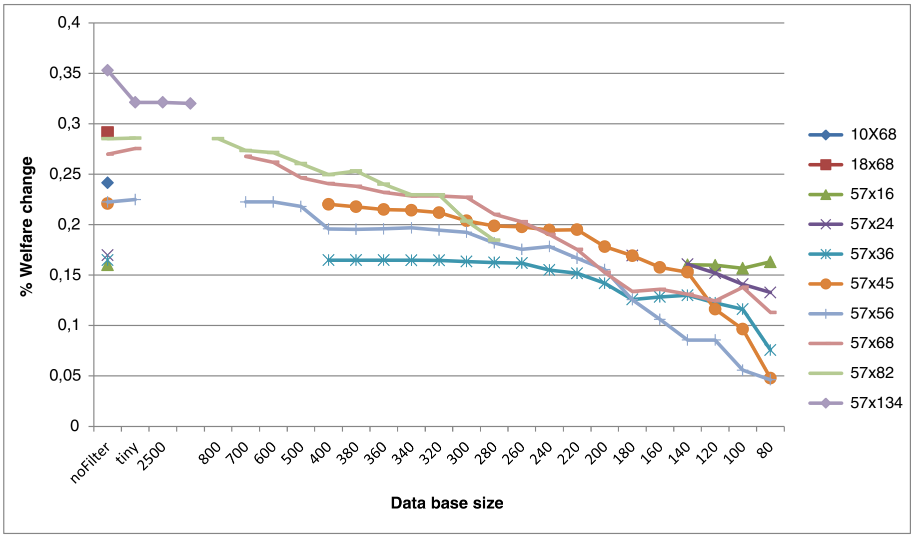
The impact of the combination of pre-model aggregation and filtering on model size and solution behavior discussed so far is clearly of interest for modelers, but what matters most are differences in results. We compare here one unique indicator at continental and global scale which received quite some attention over time: the welfare gain from tariff liberalization. Our simulated gains are well in line with similar studies and subject to the same criticisms (e.g. Ackerman and Gallagher, 2008). But we clearly do not aim at policy analysis or a discussion of potential weaknesses in our or similar simulation exercise. Rather, we use that central metric for a systematic look at aggregation bias.

A first illustrative example provides the results for an African total shown in Graph 7 below, aggregated from differently detailed results for single African countries and country blocks depending on the aggregation level used. An obvious finding is that aggressive filtering has a



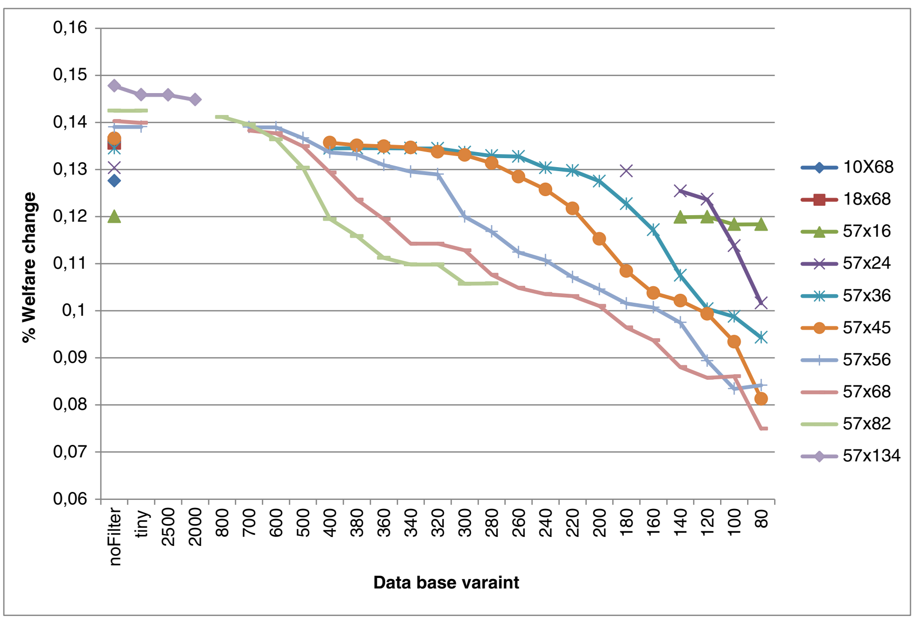
Source. Own compilation

Graph 6. Solution time in seconds with differently detailed data bases and model configurations, w/wo the pre-solves.



Source: Own compilation

Graph 7. Simulated welfare gain for Africa.



Source: Own compilation

Graph 8. Simulated welfare gain at world level.

detrimental effect on the simulated welfare effect. As seen for example from the 57 × 45 case, very aggressive and clearly not recommended filtering such as deleting 75% of the transactions reduces the simulated welfare gain by about the same relative amount.[10] But equally, reducing the number of regions in our analysis by pre-aggregation can more than half the welfare impact on Africa as a whole. The larger welfare gains simulated with more disaggregated regional data sets stem mostly from adding detail for Africa, and one might assume that any study with a focus on Africa would have used a database with similar regional detail.

The differences between simulated welfare change for a world total are less striking, see Graph 8 below, but the same pattern arises: more regional detail increases the welfare gains. The reader is however reminded here that the GTAP aggregation facility does not delete any trade and related taxation during aggregation. The reduction in simulated welfare hence stems form the fact that peaks in protection rates are flattened by averaging across regions, and not from converting former international trade into domestic sales.

Given these results, some patience seems to be recommended. As the impact of a chosen pre-model aggregation on the simulated results might be hard to predict, it seems wise to opt in doubt for a more disaggregated model, even if running it might take somewhat longer. However, moderate filtering in the range of 0.1–0.3% seems to barely impact the simulated results, but allows running models with far more regional and sector detail compared to unfiltered versions.

[10] One might assume that filtering mostly will remove tiny bilateral trade flows, but Table 1 above suggests that the reduction is more evenly spread across the data base. The reader is also reminded that the filtering algorithm re-scales trade volumes and related tax income and maintains important totals such as GDP.

The most interesting finding seems to be that using very small filtering values up to 0.001% had no discernible impact on simulated welfare changes, but can speed up solution time considerably. That type of filtering can hence always be recommended.

6. Summary and conclusions

We presented a new algorithm useful for CGE modeling which filters out tiny and thus from an economic viewpoint hardly important entries from global data bases while maintaining consistency and important economic totals. Due to a sizeable reduction in the number of transactions, the resulting models are far smaller and consequently solve faster and are more stable compared to using the data base in its original unfiltered form. We test the algorithm with a fully compatible implementation of the GTAP standard model in GAMS which adds a pre-solve algorithm based on solving single region modules, helping to speed up considerably solution time for larger models and shocks. Equally, we use a new version of CONOPT which allows for parallel execution.

We show that the approach allows running the model with a far higher number of sectors and regions compared to typical GTAP applications published in the literature, even solving the model without any pre-aggregation is possible, at least for selected shocks if solve times around an hour are accepted. In order to check impacts on modeling outcomes we first construct data sets using the default approach, i.e. aggregating the data base over sectors and regions without filtering any small values. Next, we use our algorithm to obtain filtered data bases of similar sizes which implies more regional or sectoral detail. For these about 90 data base variants, we run a partial multi-lateral trade liberalization scenario and compare simulated welfare gains, including those based on a rather large, filtered data set. This allows us to check how aggregation and filtering impact simulated results, i.e. to quantify the aggregation bias. Our results show the well-known effect that less detail reduces the overall welfare impact. As we could solve models with full sectoral resolution and up to 82 regions even without any filtering in less than 10 min, we conclude that pre-model aggregation by sectors can generally not be recommended and that larger pre-model aggregations by regions should be equally avoided.

Our analysis seems to indicate that, when using models of comparable size, the aggregation bias from the default pre-model aggregation is larger compared to using a more disaggregated data base with subsequent filtering leading to a similar size. That holds as long as the filtering thresholds are not too aggressive. We thus conclude that it is recommended to use our more data driven approach when building a tailored aggregated data base entering a CGE modeling exercises. Our tests seem to indicate that it is possible to run the GTAP model, even with additional modules, at full sector breakdown and high number of regions, potentially even at the full regional resolution of the GTAP data base, when small transactions are removed from the data base even under relatively large shocks.

The GAMS code of the filtering algorithm and to perform model runs along with a Graphical User Interface is available from the GTAP center.

References

Ackerman, F., Gallagher, K.P., 2008. The shrinking gains from global trade liberalization in computable general equilibrium models: a critical assessment. International Journal of Political Economy 37 (1), 50–77.

Aguiar, A., Narayanan, B., McDougall, R., 2016. An Overview of the GTAP 9 Data Base. Journal of Global Economic Analysis. 1(1) pp. 181–208. http://dx.doi.org/10.21642/JGEA.010103AF.

Alexeeva-Talebi, V., Böhringer, C., Löschel, A., Voigt, S., 2012. The value-added of sectoral disaggregation: implications on competitive consequences of climate change policies. Energy Econ. 34, 127–142.

Antimiani, A., Costantini, V., Paglialunga, E., 2015. The sensitivity of climate-economy CGE models to energy-related elasticity parameters. Implications for climate policy design. Econ. Model. 51, 38–52.

Banse, M., Grethe, H., Nolte, S., 2005. European Simulation Model (ESIM) in the General Algebraic Modeling System (GAMS): Model Documentation (Berlin and Göttingen).

Britz, W., Kallrath, J., 2012. Economic simulation models in agricultural economics: the current and possible future role of algebraic modeling languages. In: Kallrath, J. (Ed.), Algebraic Modelling Systems: Modeling and Soving Real World Optimization Problems. Springer, Heidelberg, Germany, pp. 199–212.

Britz, W., van der Mensbrugghe, D., 2015. "Interactive Interface for an Extended Version of the Standard GTAP Model in GAMS," GTAP Technical Paper, Forthcoming.

Britz, W., Witzke, P., 2014. CAPRI Model Documentation 2014. University Bonn, Institute for Food and Resource Economicshttp://www.capri-model.org/docs/capri_documentation.pdf (last accessed 16.09.2015).

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.

Brockmeier, M., Bektasoglu, B., 2014. Model structure or data aggregation level: which leads to greater bias of results? Econ. Model. 38, 238–245.

Brooke, A., Kendrick, D., Meeraus, A., 1988. GAMS: A User's Guide. The Scientific Press, California.

Caron, J., 2012. Estimating carbon leakage and the efficiency of border adjustments in general equilibrium—does sectoral aggregation matter? Energy Econ. 34, 111–126.

Drud, A.S., 1994. CONOPT—a large-scale GRG code. ORSA J. Comput. 6 (2), 207–216.

Grant, J.H., Hertel, T.W., Rutherford, T.F., 2007. Tariff line analysis of US and international dairy protection. Agric. Econ. 37 (1), 271–280.

Harrison, W.J., Pearson, K.R., 1996. Computing solutions for large general equilibrium models using GEMPACK. Comput. Econ. 9 (2), 83–127.

Harrison, W.J., Pearson, K.R., Powell, A.A., Small, E.J., 1994. Solving applied general equilibrium models represented as a mixture of linearized and levels equations. Comput. Econ. 7 (3), 203–223.

Hertel, T.W., 1997. Global Trade Analysis: Modeling and Applications. Cambridge University Press, Massachutes, USA.

Himics, M., Britz, W., 2015. Flexible and welfare-consistent tariff aggregation over exporter regions. Econ. Model. 53, 375–387.

Horridge, M., 2006. GTAPAgg Data Aggregation Program. Global Trade, Assistance, and Production: The GTAP 6 data base.

Ko, J.-H., 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 (last accessed 16.09.2015).

Narayanan, G.B., Aguiar, A., McDougall, R. (Eds.), 2012. Global Trade, Assistance, and Production: The GTAP 8 Data Base. Purdue University, Center for Global Trade Analysis.

OECD, 2007. Documentation of the Aglink-Cosimo Model. Organization for Economic Cooperation and Development, Parishttp://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=AGR/CA/APM(2006)16/FINAL&docLanguage=En (last accessed 16.09.2015).

Pelikan, J., Brockmeier, M., 2008. Methods to aggregate import tariffs and their impacts on modeling results. Journal of Economic Integration 23 (3), 685–708.

Rosegrant, M.W., Msangi, S., Ringler, C., Sulser, T.B., Zhu, T., Cline, S.A., 2008. International Model for Policy Analysis of Agricultural Commodities and Trade (Impact): Model Description (p. 42). International Food Policy Research Institute, Washington, DC, USA.

Rutherford, T., 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 (last accessed 16.09.2015).

Van der Mensbrugghe, D., Britz, W., 2015. The Standard GTAP Model in GAMS Version 6.2, GTAP Technical Paper, Forthcoming.

Dominique van der Mensbrugghe is Research Professor and Director of the Center for Global Trade Analysis (GTAP) at Purdue University. Prior to joining Purdue, he worked at the Food and Agriculture Organization (FAO), the World Bank and the OECD. His work has concentrated on long-term structural change of the global economy and the analysis of global economic policy issues—including agricultural policies, regional and multilateral trade agreements, demographics and international migration, and climate change. He is one of the world's experts on global computable general equilibrium modeling and has a PhD in economics from the University of California, Berkeley.

Wolfgang Britz after finishing his doctoral degree in agricultural economics worked as a consultant inter alia for the European Commission, the FAO and the OECD in the field of economic modeling and impact assessment of policies related to food and agricultural markets. He holds a tenured position as a senior lecturer and researcher at the Institute for Food and Resource Economics, University Bonn, Germany. His research focuses on economic and bioeconomic simulation models, from farm to global scale.