Reconciling the GTAP Data Base: Where are the Big Changes?[1]

Terrie Walmsley[2] Badri Narayanan, Angel Aguiar and Robert McDougall[3]

Draft paper

This is a work in progress please do not quote.

1 The authors would like to Thank Patrick Jomini for his encouragement in producing this paper.
2 Terrie Walmsley, Director and Senior Economist, ImpactECON, LLC. twalmsley@impactecon.com
3 Badri Narayanan, Angel Aguiar and Robert McDougall are researchers at the Center for Global Trade Analysis, Purdue University, West Lafayette, IN, 47906.

Abstract

Inaccuracies in the collection and compiling of data mean that data supplied by one country is rarely consistent with data supplied by another. Even within the same country, the same data collected from two alternative sources could vary due to differences in how the data is being collected, interpreted, classified, and valued, notwithstanding the differences caused by simple errors and omissions.

Global economic analysis however, requires consistent and reconciled global data, and this reconciliation process is laden with judgments about the quality of the alternative data sets being reconciled. In the case of the GTAP Data Base (Narayanan, Aguiar et al. 2012) the externally collected trade, macro, protection and energy data supplied by international sources are considered superior to individual country data because it has typically gone through a standardization and balancing process. The extent to which this reconciliation process alters the country data however is not known.

In this paper, we hope to shed some light on the extent to which the country data changes as a result of the GTAP construction process. In particular, we are interested in identifying where the largest changes occur as a result of this reconciliation process. Do changes mostly occur in particular countries where data are poorer; or do changes mostly occur in particular data across all countries. The answers to these questions can be used to help ascertain where resources might best be utilized to further improve the quality of global data. We find that there is some evidence that data from developing countries with weaker IO tables and less sectors do undergo more changes than those with more robust IO tables, the largest differences however occur in the sales shares due to differences between the trade data in the IO tables and the balanced trade dataset used in the GTAP database.

1 Introduction

Data collected by one country is rarely consistent with data collected by another due to inaccuracies in the collection and compiling of data. An example of this is in the trade data, where the amount or value of exports one country claims to sell to another rarely matches what the other country claims to import or buy from them, even after transportation and other costs are taken into account. Even within the same country, the same data collected from two alternative sources may differ due to differences in how the data is being collected, interpreted, classified, and valued, notwithstanding the differences caused by errors and omissions. For example, GDP collected from the expenditure side never equals GDP collected from the production side, there is always a statistical discrepancy.

Global economic analysis however, requires consistent and reconciled global data. The reconciliation of global datasets is laden with judgments about the quality of the alternative data sets being reconciled. In the case of the GTAP Data Base (Narayanan, Aguiar et al. 2012) the externally collected trade (Gehlhar, Wang et al. 2008), macroeconomic (World Development Indicators, World Bank), protection (Boumellassa, Laborde et al. 2009) and energy data (Organisation for Economic Cooperation and Development and International Energy Agency 2003) are contributed based on international sources and are considered superior to the individual country data because they have gone through this standardization and reconciliation process.

The choice of which international datasets to include and how to rank them has evolved over the last twenty years as new data has become available and new policy issues have become increasingly important to policymakers and their constituents. For instance, the growing concern for climate change has globally raised the importance of modelling energy and greenhouse gas emissions, leading to significant changes in the way the GTAP Data Base dealt with energy data.

More recently concerns have been raised about the decision by the Center to rank the international datasets above the country data, and as a result other global datasets, namely WIOD developed by Marcel, Erumban et al. (2012), have been produced that give the country data preference. While country data (particularly supply and use tables) has improved considerably over time, and concerted efforts (European Commission, IMF et al. 2008) continue to bring welcome improvements in quality, country data continues to be a source of concern as the country coverage of GTAP expands – the GTAP 8.1 database contained 134 countries, up from 66 in version 5 (Dimaranan and McDougall, 2002). On the other hand, data sources (e.g., Marcel, Erumban et al. (2012) and Tukker, Poliakov et al. (2009)) that start with the intention of giving preferential treatment to country data must ultimately make decisions about what to do about negative residuals in trade.

The extent to which the GTAP reconciliation process (nicknamed *FIT*) alters the country data has not been widely examined, and there is often a great deal of speculation and debate about the extent and necessity of all or some of these changes, particularly as the country data is believed to have improved over time. In this paper we hope to shed some light on this issue by examining the extent to which the country data changes as a result of the GTAP construction process. In particular we are interested in:

Which countries change the most? Are poorer countries affected more by the reconciliation process than richer countries, because of the less reliable data collection methods at their disposal? Do the countries which change the most have older data or more missing data?

Are there particular inputs or uses that stand out across all countries? Are these related to particular external data sources?

It is hoped that the answers to these questions can be used to help ascertain where resources might best be utilized to further improve the quality of global data. In section 2 we briefly review the GTAP Data Base construction process; and in section 3, we outline the methods used to compare the pre- and post-*FIT* country data. Section 4 then provides the results from the comparison, by country and by input/use and attempts to gain some insights into the quality of the data based on other information we have about the contributed data, e.g., how old the data is, missing information etc. Section 5, then concludes the paper.

2 The GTAP Data Base construction

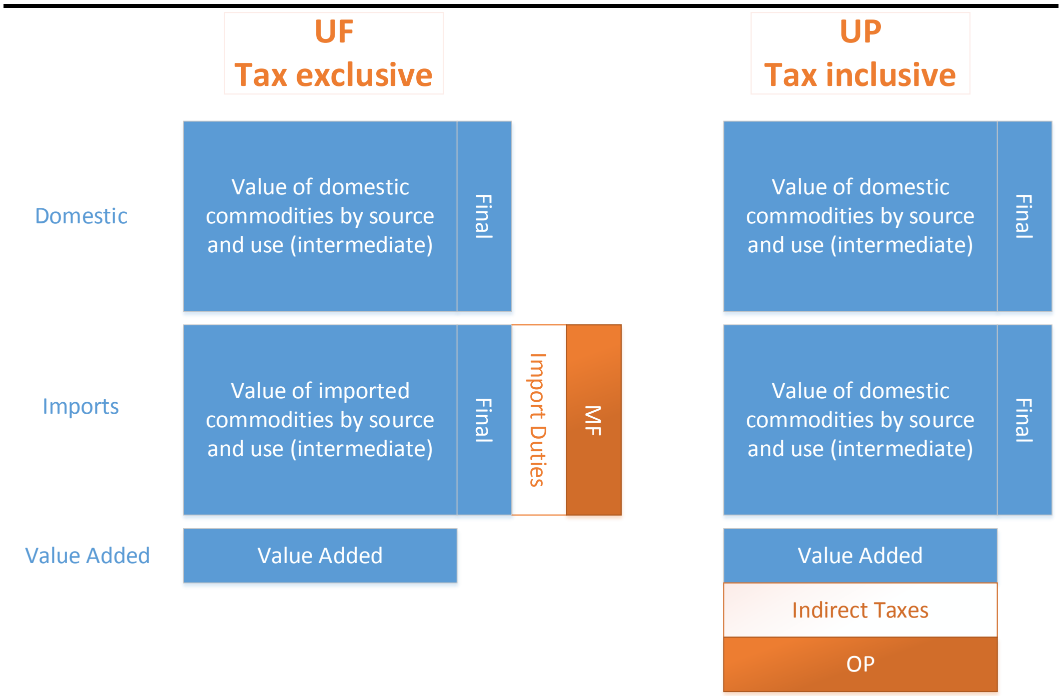
In this section we provide an overview of the data construction process. First we review the IO table contribution process and potential issues with the IO tables faced by contributors and the Center after contribution. Next we examine the GTAP construction process, including any preliminary adjustments made to the IO tables to disaggregate or target production, the reconciliation of the IO tables with international data sources (*FIT*) and the final assembly.

2.1 IO Table Contributions

Economic data for each country including the value of inputs and uses of production is provided as an IO table by a global network of contributors. Full details of what data is required by GTAP is provided in Huff, McDougall et al. (1998). Figure 1 provides a pictorial view of the four arrays (UP, UF, MF and OP) that are needed. The four arrays represent the IO table inclusive and exclusive of sales taxes, import duties and indirect taxes.

The starting point for most contributors is an I-O table, supply and use tables, or social accounting matrix developed by the country's statistical office. Most tables require revisions to be made by the contributor to get the original table into a format ready for GTAP. Below we outline some of the issues that are dealt with by contributors, with assistance from the Center.

Figure 1: Format for IO table Contributions to the GTAP Data Base



Source: Huff, McDougall et al. (1998)

Sectoral Coverage and Concordances: The GTAP Data Base currently includes 57 sectors (https://www.gtap.agecon.purdue.edu/databases/v8/v8_sectors.asp). Contributed tables must distinguish at least 30 sectors, and these sectors must be mapped directly to the 57 GTAP sectors. Moreover they must ensure that agriculture and food processing, and energy are separated from other sectors. Unfortunately not all countries use the standard international classifications and therefore contributors must often determine their own concordance; ambiguity is common. Moreover, any disaggregation performed by the contributor usually involves applying shares equally to the disaggregated sector, unless they have additional data and the skills to apply balancing techniques.

Sign: With the exception of changes in stocks, all values must be positive. A common problem in IO tables is that they reflect a given year and it is possible that returns to capital are negative, due to a bad year. The contributor will need to adjust these returns to reflect a more typical return in a normal year.

Balance: The table must balance. That is total sales of each sector must equal to total costs. There may be cases were the original IO table does not balance or, more likely, that changes made elsewhere by the contributor cause am imbalance elsewhere that must be rectified.

Taxes: Taxes may be missing in the original supply and use or I-O table. Contributors must therefore chose to add or exclude them from the contributed IO table.

Dwellings: Dwellings are defined as imputed rents and are calculated based on ownership of dwellings. The main cost of dwelling is usually capital, and it is usually sold directly to households. Despite the fact that dwellings estimates are required under the SNA, many countries do not report dwellings or aggregate them with rent. Adjustments must be made, either by contributors or by the GTAP Center to include or disaggregate them.

Government: The government should purchase primarily from the government services sector. IO tables may treat government differently and adjustments must be made, either by contributors or by the GTAP Center.

Trade and transport margins: Since domestic margins are not treated separately in GTAP, they must be brought to account as direct purchases of trade and transport services by the consuming sector. Statistical offices often account for these domestic margins separately in the IO or supply and use table. Contributors must therefore make adjustments to include them as direct purchases.

Re-exports: Re-exports are not permitted in the GTAP Data Base. If the input-output table includes re-exports then they need to be removed from imports and exports.

Domestic and imported use matrix: GTAP requires that the table distinguishes between domestic and imported uses; this data is often not available in the original data obtained from the statistical office and adjustments must be made by the contributor.

Additional row/column constructs: I-O tables obtained from the statistical office sometimes contain additional rows/columns to represented imputed items that are not required by GTAP. An example of this is FISIM (financial intermediation) which represents imputed banking services; these need to be removed.

Value-added: Contributors are asked to provide data on inputs of land, labor and capital by sector. Land in particular is usually missing and no account is specifically taken of self-employed labor.

2.2 The Construction of the GTAP Data Base

The GTAP construction process, described in Figure 2, can be divided into three components: IO tables pre-processing, FIT, and assembly. Each of these is discussed in turn below.[4]

2.2.1 PRE-PROCESSING OF CONTRIBUTED IO TABLES

The contributed input-output tables which have already satisfied the guidelines for contributors, then go through a cleaning procedure and are pre-processed to produce consistent 57 sector tables. This pre-processing includes:

Minor Cleaning: Any small imbalances or small negatives are removed.

Disaggregation: Those I-O tables which do not have full agricultural and/or non-agricultural disaggregation then go through the I-O disaggregation procedure. Agricultural IO tables based on FAO data and contributed by Everett Peterson are used to disaggregate agriculture. This ensures that the cost shares of the disaggregated agricultural commodities look reasonable. Non-agricultural disaggregation is done using the shares obtained from the representative table. Note that disaggregation does not alter the aggregated totals, it merely apportions the cost and use structures while keeping the same totals.

Composite Regions: Input-output tables are constructed for all composite regions – GTAP regions for which there are no contributed tables. This is done by matching each country within the composite region with a country for which we have an IO table. This country's IO table then act as a proxy for the missing country. Later all the proxy IO tables for countries in each composite region are aggregated to form the composite regions.

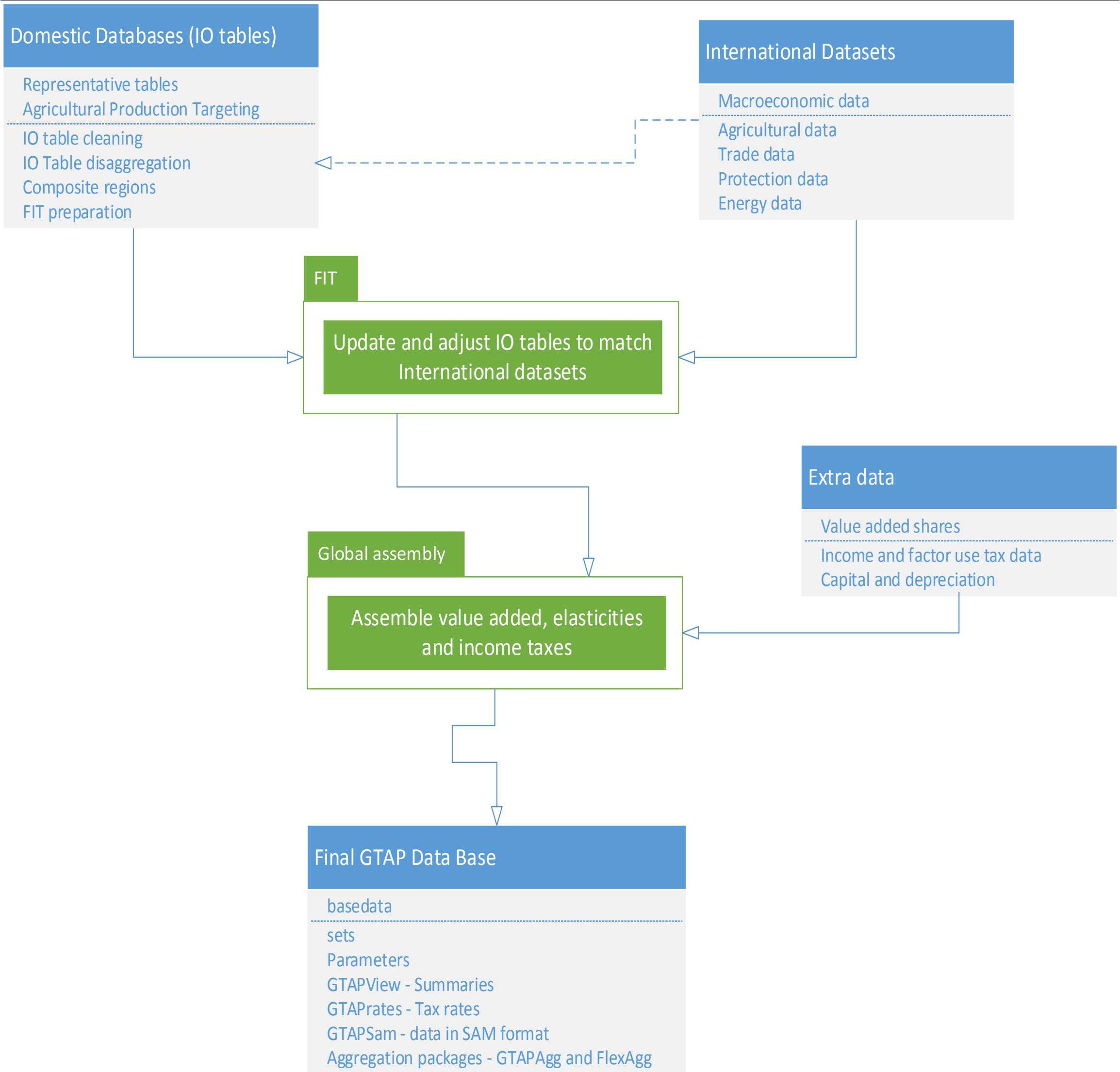
Agricultural Production Targeting: Agricultural output in several countries, especially the OECD members, is targeted to match agricultural production statistics by sector.

In this paper, our initial data is taken from this stage in the construction process. It is the contributed IO table after disaggregation and cleaning, but after agricultural production targeting. By taking the data at this stage we can circumvent aggregating the final data back to the original aggregation, which will differ by country and make country rankings more difficult. The disaggregated initial data we use aggregates back to the original contributed IO table. Table A 5 shows which contributed tables were aggregated and hence had to be disaggregated. Composite regions are considered separately.

---

[4] For those looking for a more basic understanding of the GTAP Data Base we refer you to Walmsley, Aguiar and Narayanan (2013), Harslett (2013) and Narayanan, Aguiar and McDougall (2012).

Figure 2: GTAP Global Data Base Construction Procedure



Source: Authors' construction

Table 1: International Datasets used in GTAP 8.1 Data Base

Data Set	Details
Reference year	2007
Standard Countries (# of countries international data are collected for)	244
Macroeconomic Data	World Development Indicators
Govt. Consumption	International Financial Statistics
Goods Trade Data	COMTRADE, processed by Mark Gehlhar
Services Trade data	OECD and CPB
Domestic Support	OECD and Hans Jensen
Export Subsidies	David Laborde
MFA Export Tax Equivalent	Joe Francois
Agricultural Tariffs	MacMAPv3 from ITC/CEPII
Merchandise Tariffs	MacMAPv3 from ITC/CEPII
Energy Data	IEA energy price & volumes data
Agricultural Factor Split	FAO, processed by Peterson
Income & Factor Taxes	IMF data
Population Data	World Bank

Source: Authors' construction

2.2.2 FIT

A single procedure is used to achieve the next three objectives:

- The IO tables are updated to the reference year – Version 8.1, 2007.[5]
- The IO tables are adjusted to match the trade, protection, energy, and macro-economic variables in the global datasets (Table 1).
- Changes in stocks are eliminated.

We call the adjustment procedure fitting the I-O tables, after the program FIT that implements it. FIT applies entropy-theoretic methods (Theil (1967) and Bacharach (1970)) to adjust an I-O table to various external constraints derived from the international data sets. It was originally developed as part of the SALTER project at the (Australian) Industry Commission (James and McDougall 1993), and has since been extended for GTAP.

We apply the fitting procedure after disaggregating the primary I-O tables and constructing composite tables for each country in a composite region. Thus the inputs into the procedure are a complete set of fully disaggregated regional I-O tables for 244 countries, and a set of international data sets (listed in Table 1) and the outputs are the fitted I-O tables and an adjusted energy volumes data set.

[5] The reference year of the IO table does not need to be the same reference year in the final GTAP Data Base. Bringing all IO tables, from different base years, to a common reference year is the first adjustment performed to the country data.

The following targets are applied:

i. from constraints imposed by the GTAP data base structure: zero values for changes in stocks of domestic product and imports, by commodity;

ii. from the macroeconomic data set: values at purchasers' prices for GDP, aggregate private consumption, government consumption, and investment;

iii. from the trade data set, modified according to the energy data set: border values of exports and imports, by commodity;

iv. from the protection data set, modified and supplemented from the energy data set: import duty rates, by commodity; export subsidy rates, by commodity; non-commodity indirect tax rates, by industry; commodity tax rates on intermediate usage, by industry and commodity; and rates of tax on private consumption of energy, by commodity; and

v. from the energy data set: basic values for intermediate usage of energy, by energy industry and energy commodity; basic values for private consumption of energy, by energy commodity (the energy industries and energy commodities are aggregations of standard GTAP sectors).

The FIT program incorporates an I-O quantity model, an I-O price model, and an entropy-theoretic balancing procedure. Broadly speaking, the I-O quantity model serves to remove changes in stocks and adjust exports, consumption, and investment. It feeds these final demand changes backward through the I-O structure to determine new levels for intermediate usage and primary factor employment. The I-O price model feeds tax rate changes forward through the I-O structure to adjust basic and post-tax prices for intermediate usage and final demands. The entropy procedure adjusts taste and technology variables to meet the import and energy usage targets. The general rule in the fitting procedure is to adjust the regional I-O tables to the international data sets, rather than the other way around, with some exceptions in agricultural domestic support and energy data.

It should be noted that since the international data sets match the data base reference year, adjusting the I-O tables to the international data sets is also the method used to update the year to the base year. This also converts the IO table into the correct units and currency.

2.2.3 Assembly

The data assembly module is where:

i. adjustments to value added take place;

ii. the factor payments data are adjusted to incorporate land- and capital-based payments;

iii. the various international data sets and domestic data bases are put together and final checks are made;

iv. additional data such as population, capital stocks, depreciation, and savings are included; and

v. additional datasets used in the standard GTAP model or its variants (e.g., sets, elasticities, energy volumes) and summary (e.g., time series trade data, tax rates) datasets are produced.

In terms of value added, several adjustments are made. First, labor payment data are disaggregated into skilled and unskilled labor payments using payment shares generated in the estimation procedure. Second, factor employment data for primary agriculture and natural resource-based sectors are adjusted using primary factor shares documented in Narayanan, Aguiar et al. (2012). In agriculture, external estimates of factor earnings shares are used; and for natural resource based sectors a proportion of the earnings of labor and capital is reallocated to natural resources to achieve target supply elasticities.

3 Methodology for Comparing Initial and Final IO tables

In order to compare the original contributed IO tables with the final country data in the GTAP Data Base we must first put them both into a comparable format. The format chosen is the four arrays structure (Figure 1) available to contributors (Huff, McDougall et al. 1998). IO table contributions in alternative formats are first converted into this format and then combined into a single file. The GTAP database itself is also converted into this format (see Figure 1 for a schematic of how the GTAP headers match the four array format).

The need for a comparable format also raises the question about which initial data to use, since many of the IO tables were not contributed with the full 57 sectors disaggregated. For this reason we have decided to take the initial IO tables out of the pre-processing stage after minor cleaning and disaggregation. As noted above, at this stage they still aggregate up to the original tables.

Another issue we need to consider is that the initial data may not be in the correct currency/units and hence comparing the initial and final values would result in large differences simply due to differences in currency and units; for this reason we compare shares rather than values. This means that we do not examine how different the contributed table's estimates of GDP is from the World Bank once units and currency are taken into account. We can however examine differences in the shares of final demand to GDP between the initial and final table. We can also look at the absolute differences between initial and final sales shares as well as initial and final cost shares.

Finally, in the interests of being able to diagnose when in the process the changes we are also in the process of seeing if at least some intermediate inputs can be compared. Unfortunately, there are not a lot of intermediate inputs available since much of the processing occurs in one step during the FIT module. Nevertheless, this would be beneficial to examining the impact of agricultural production targeting and the processing of value added.

As mentioned above, we examine both sales ($S^I_{i,u,r}$ or $S^F_{i,u,r}$) and cost shares ($C^I_{i,u,r}$ or $C^F_{i,u,r}$) in the initial (I) and final (F) data:

$$S^I_{i,u,r} = \frac{UP_{i,u,r}}{\sum_i UP_{i,u,r}} \tag{1}$$

and

$$C^I_{i,u,r} = \frac{UF_{i,u,r}}{\sum_i UF_{i,u,r}} \tag{2}$$

Both the (percentage point) differences ($D_{i,u,r}$) in shares are examined, as well as an entropy measures ($E_{i,u,r}$) of the differences (Walmsley and McDougall 2004).

$$D_{i,u,r} = S^F_{i,u,r} - S^I_{i,u,r} \tag{3}$$

The entropy method calculates the difference between the shares in the I-O table as an equal weighted average of each share multiplied by the natural log of the ratio of that share to the equivalent share in the comparison I-O table. In this case the example is sales shares, although cost shares can be examined in the same way.

$$E_{i,u,r} = 0.5\left[S^I_{i,u,r}\left[LOG_e(S^I_{i,u,r}/S^F_{i,u,r})\right]\right] + 0.5\left[S^F_{i,u,r}\left[LOG_e(S^F_{i,u,r}/S^F_{i,u,r})\right]\right] \tag{4}$$

where: $E_{i,u,r}$ is the entropy measure of the difference between $S^I_{i,u,r}$ and $S^F_{i,u,r}$.

$S^I_{i,u,r}$ is the adjusted share of input i in used in use u in the initial I-O table.

$S^F_{i,u,r}$ is the adjusted share of input i used in use u in the comparison/final I-O table.

The benefit of using the entropy approach is that the same absolute difference between two large shares is considered to be less important than the same absolute difference between two smaller shares, while in the shares method they are treated equally. The entropy measure itself however, has no meaning and hence we tend to concentrate on the differences and the standard deviation of those differences in shares across uses, inputs or countries.

4 Results from the Comparison

In this section we examine the results of the comparison. In the first sub-section we combine the cost and sales shares differences by country to examine if these differences are larger in particular countries. We calculate the mean and standard deviation of the differences/changes in order to obtain 90 percent confidence intervals of the shares. When looking at countries the mean changes are zero, since positive changes are offset by negative changes in these shares.

In the second section we calculate the mean and standard deviations of the changes in the IO tables by inputs (costs) and uses (sales) in order to examine if there are particular inputs or uses that change more than others. In this case the means may not be zero if there are persistent changes in the shares across all countries. For example, land is generally not include in contributed IO tables and hence the share of land in costs usually rises, giving a mean of greater than zero.

Table A 6 provides a list of sectors and Table A 7 provides a list of other codes. Country codes are taken from GTAP and/or the GTAP ISO classification.

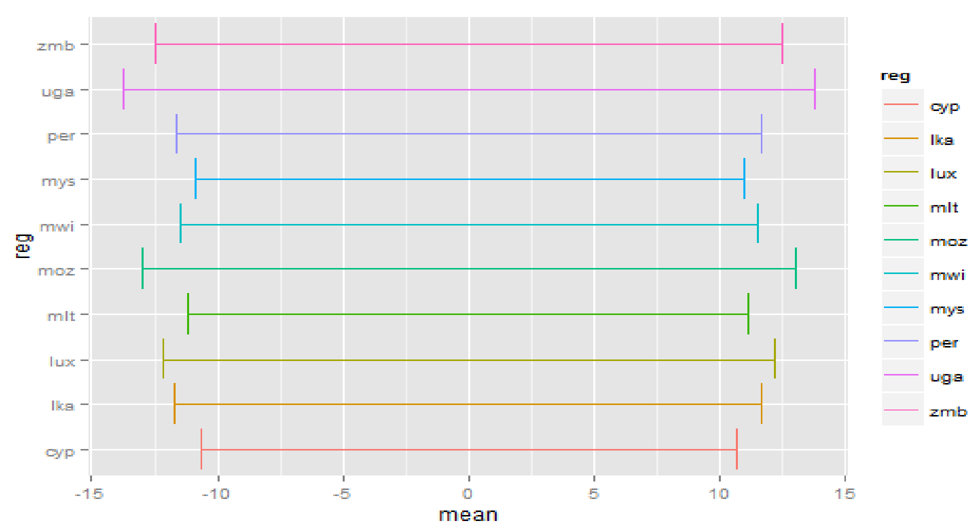
4.1 By Country

Table A 1 in the appendix lists the standard deviation, minimum, maximum and 90% confidence interval of the percentage point differences (Equation 3) in the cost shares by country, as well as the entropy measure (equation 4). The table orders countries according to the standard deviation, which are correlated (not perfectly) with the entropy measure. We find that the standard deviation in cost shares ranges from 5.5 percent differences for Uganda to 1.4 percent in New Zealand. Likewise, Table A 2 shows the results for the sales shares. In general the changes in the sales shares are greater than those seen in the cost shares, with the standard deviation ranging from 9.26 percent in Malawi to 1.64 percent in Argentina.

The confidence intervals of the 10 worst countries in terms of cost and sales shares are depicted in Figure 3 and Figure 4, respectively. A comparison of these panels also reveals that sales shares are more likely to change as a result of the GTAP reconciliation process than cost shares, with the confidence interval of ±20 percent as opposed to ±10 percent. This is not surprising since most of the external macro data imposed on the IO tables relate to total final use. Hence changes to private consumption, government consumption, investment and exports will alter their (sales) shares relative to each other and to intermediates, which are not targeted. The allocation of total final demand across commodities is not targeted and hence (cost) shares, the share of final demand by commodity, remain unchanged. External data imposed on the cost structure, on the other hand, tend to be imposed equally across all uses causing slight changes in the cost structure of all industries and final demand, rather than large changes in some industries. For example, total imports by commodity are targeted and will result in a dampened rise or fall of their (cost) share in all intermediate and final demand categories. Value-added is another example. The allocation of value-added across capital, land and labor types is assumed to change, but the share of total value-added is not altered, thereby limiting changes in value-added relative to intermediate demand.

Figure 3: Top 10 Countries with largest confidence intervals around (percentage point) differences in Cost Shares between initial IO tables and Final GTAP Data Base

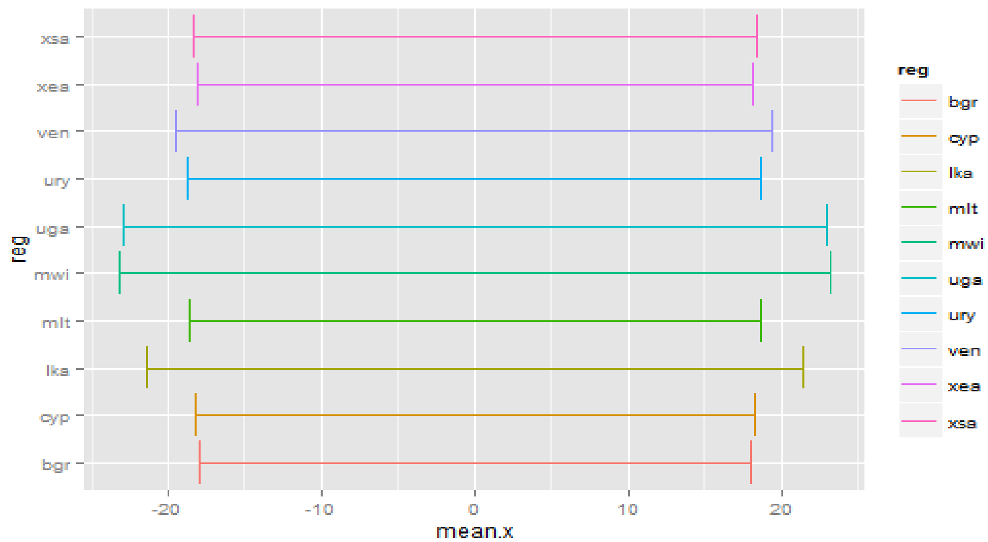
Confidence interval around (percentage point) Change in Cost Shares by Country ordered by Standard Deviation



Source: Authors' calculations

Figure 4: Top 10 Countries with largest confidence intervals around (percentage point) differences in Sales Shares between initial IO tables and Final GTAP Data Base

Confidence interval around (percentage point) Change in Sales Shares by Country ordered by Standard Deviation



Source: Authors' calculations

In general, we also find that countries which have the largest cost shares differences also experience the largest sales share differences (98 percent correlation between the two country lists). This is probably not surprising given that ultimately changes in data are likely to affect both sales and cost shares.

We are interested in the extent to which these differences might be related to the 'quality' of the IO tables, broadly defined. To consider quality we might examine things like the base year, when it was last contributed, how many sectors were contributed etc. The relationship between the changes in the IO table shares and the version number or base year is much weaker than expected; newer tables have slightly lower sales share differences but the correlation is only 0.15. The number of sectors, particularly manufacturing sectors, is a better indicator of changes, with a correlation of almost -0.5 for both sales and cost shares. We might also expect the level of development to indicate quality of the underlying IO table. Again we find only a small correlation between the changes and per capita GDP, although there is a larger correlation between changes and the World Bank development categories (0.36). The larger the economy, measured in terms of GDP or population, the larger the percent differences in both sales and cost shares (-0.3). This suggests that there are other factors involved in explaining these changes.

Finally, Table 2 compares the ranking using standard deviation to those obtained if entropy is used. There is a high correlation between the two measures, 0.88.

Table 2: Comparison of Entropy v Difference Rankings

Cost Shares			Sales Shares		
I	II	III	IV	V	VI
Top ten Countries (ordered by standard deviation)	Ranking of countries in column I if entropy used to order	Top 10 Countries ordered by Entropy)	Top ten Countries (ordered by standard deviation)	Ranking of countries in column I if entropy used to order	Top 10 Countries ordered by Entropy)
Uganda	4	Mozambique	Malawi	2	Uganda
Mozambique	1	Singapore	Uganda	1	Malawi
Zambia	9	Sri Lanka	Sri Lanka	4	Venezuela
Luxembourg	20	Uganda	Venezuela	3	Sri Lanka
Sri Lanka	3	Venezuela	Uruguay	5	Uruguay
Peru	13	Malawi	Malta	17	Peru
Malawi	6	Tanzania	Rest of South Asia	7	Rest of South Asia
Malta	17	Madagascar	Cyprus	12	Chile
Malaysia	36	Zambia	Rest of East Asia	18	Rest of Former Soviet Union
Cyprus	12	Bangladesh	Bulgaria	19	Malaysia

Source: Authors' calculations

4.2 By Use and Input

In this section we examine differences with respect to the inputs (cost shares) and uses (output shares) across all countries to determine if there are particular inputs or uses that are causing most of the differences across countries. The results for cost shares are shown in Figure 5 and Table A 3; and those for the sales shares are found in Figure 6 and Table A 4.

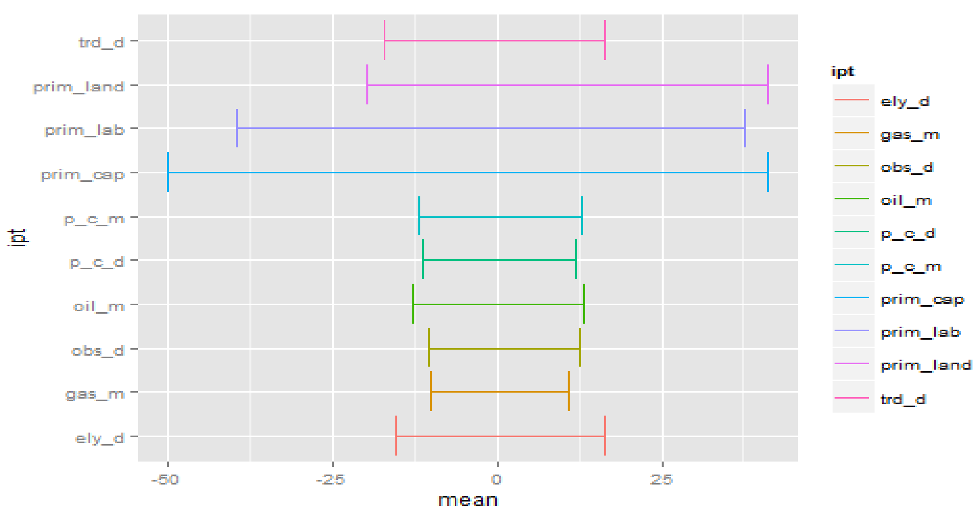
In the case of cost shares the largest differences can be found in two areas, value-added (capital (prim_cap), land (prim_land) and labor (prim_lab)) and energy (domestic electricity (ely_d), imported oil (oil_m), imported petroleum (p_c_m), domestic petroleum (p_c_d) and imported gas (gas_m)).[6] The changes in value added are large, with standard deviations between 12 and to 19 percent, however they are not surprising given that most contributed IO tables do not include land and hence this must be estimated and extracted from other components of value-added. As a result the mean change in the share of land between the initial and final IO table is positive, while the mean change in the share of labor and capital is generally negative.

[6] Not that these tables were obtained after agricultural targeting and hence they underestimate the potential changes that occur to agricultural commodities in cost shares in those countries that undergo agricultural targeting.

The standard deviation of the changes in the cost shares of domestic and imported intermediates range between 6.67 for domestic trade and 1.67 for domestic cattle meat (cmt_d). Energy commodities account for 5 of the top 10 and are likely the result of the inclusion of additional energy data obtained from the IEA data. In general the share of energy in production costs rises due to the inclusion of the additional energy data, although the increases are less than 0.6 percent on average (mean). The more surprising results are the inclusion of changes to the cost share of domestic other business services (obs_d) and trade (trd_d) in the top 10 (Figure 5). Further examination of this suggests these difference come from changes in intermediate use of trade by energy; and changes in intermediate use of other business services by agriculture or services in Europe.

Figure 5: Top 10 largest confidence intervals around (percentage point) differences in Cost Shares between initial IO tables and Final GTAP Data Base (across all countries)

Confidence intervals around (percentage point) differences in Cost Shares by Input ordered by Standard Deviation

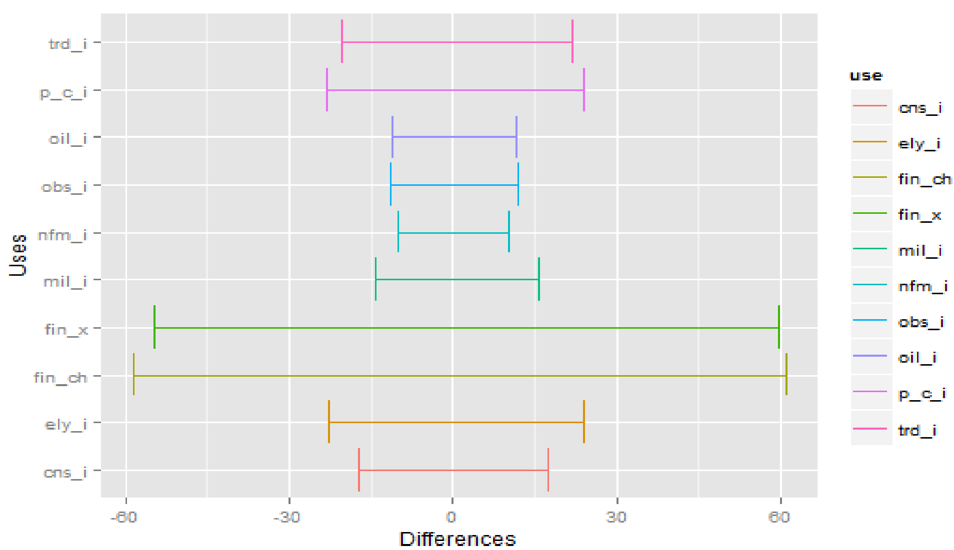


Source: Authors' calculations

The changes in the sales shares depicted in Figure 6 and Table A 4 are of greater concern. As expected from the analysis of the country data the differences in the sales shares were on average much larger than the cost shares, and hence the differences here are also much higher. The share of private consumption (fin_ch) in total sales rises by 1.23 percent and exports by 2.62 percent relative to the initial shares in the contributed IO tables, with standard deviations of more than 20 percent. Investment (fin_if) and government (fin_cg) also have large standard deviations (9.4 and 7.4 respectively). The entropy results (Table 3), which adjust for the size of the initial share, also place final demand in exports, private consumption, government and investment in the top four positions as the shares that have moved the most as a result of the GTAP construction process.

Figure 6: Top 10 largest (percentage point) differences in Cost Shares between initial IO tables and Final GTAP Data Base (across all countries)

Confidence intervals around (percentage point) differences in Sales Shares by Input ordered by mean



Source: Authors' calculations

A lot of these changes in sales shares within private and government consumption and investment occur in their demand for imports and are therefore most likely due to the matching of imports to the trade data. Likewise exports also adjust due to matching to trade data. With trade and final private consumption (and to a lesser extent government) rising as a result of the construction process, there are large movements in both domestic and imported sales across the uses as both imports by commodity and total final demand (total C, I, and G, and X by commodity) are targeted.[7] The large differences between exports and imports reported by individual countries and the imports and exports resulting from the global reconciliation process is well known, however according to these comparisons a significant portion of the changes to the structure of the IO tables stem from the trade data.

One reason for the significant differences may be the growing importance of re-exports. Re-exports are removed from IO tables as part of contribution process, however adjustments to the trade data are only done for a select group of countries – Hong Kong and the Netherlands. A case in point is Cyprus, which is in the top 10 worst countries for changes to the underlying IO data. Cyprus is often the first port of call into the European Union and hence processes a lot of goods for entry into the European Union that it simply re-exports to other parts of the European Union.

[7] Note that changes also occur due to the removal of change in stocks. These have been removed from the rankings.

The most heavily impacted sectors again include a couple of energy sectors, although the sectors most affected are those that rely on imported inputs in a lot of countries, e.g., the other food sector (ofd_i) and trade (trd_i). Paddy rice (pdr_i) and raw milk (rmk_i) are also high on the list. This is most likely due to aggregation issues, since these raw agricultural materials are rarely traded, while their counterparts processed rice (prc) and milk (mlk) are.

Table 3: Comparison of Entropy v Difference Rankings

Cost Shares			Sales Shares		
Top ten Countries (ordered by standard deviation)	Ranking of countries in column I if entropy used	Top 10 Countries ordered by Entropy)	Top ten Countries (ordered by standard deviation)	Ranking of countries in column I if entropy used	Top 10 Countries ordered by Entropy)
Capital (prim_cap)	3	Land (prim_land)	Final Household consumption (fin_ch)	3	Exports (fin_x)
Labor (prim_lab)	5	Imported government services (ogs_m)	Exports (fin_x)	1	Final Government consumption (fin_cg)
Land (prim_land)	1	Capital (prim_cap)	Other food products (ofd_i)	10	Final Household consumption (fin_ch)
Domestic Trade (trd_d)	15	Imported gas (gas_m)	Petroleum (p_c_i)	5	Investment (fin_if)
Domestic electricity (ely_d)	14	Labor (prim_lab)	Investment (fin_if)	4	Petroleum (p_c_i)
Imported oil (oil_m)	6	Imported oil (oil_m)	Electricity (ely_i)	8	Oil (oil_i)
Imported petroleum (p_c_m)	9	Imported wheat (wht_m)	Trade (trd_i)	7	Trade (trd_i)
Domestic petroleum (p_m_d)	28	Imported paddy rice (pdr_m)	Final Government consumption (fin_cg)	2	Electricity (ely_i)
Domestic other business services (obs_d)	38	Imported Petroleum(p_c_m)	Chemicals, rubbers and plastics (crp_i)	9	Chemicals, rubbers and plastics (crp_i)
Imported Gas (gas_m)	4	Imported vegetables and fruit (v_f_m)	Textiles (tex_i)	14	Other food products (ofd_i)

Source: Authors' calculations

5 Conclusions

In this paper, we compared the GTAP IO tables before and after being processed for the GTAP Data Base in order to examine where the largest changes occur in the IO tables as a result of the GTAP construction process. We find that while there is some evidence that data from developing countries with weaker IO tables and less sectors do undergo more changes than those with more robust IO tables, the largest differences occur in the sales shares due to differences in the trade data between the contributed IO tables and the balanced trade dataset used in the GTAP database. We find that the energy sectors, as well as agriculture, other business services and trade also appear to have higher absolute changes in their sales shares than other commodities. Other large changes result from the re-allocation of value added across land, labor and capital as land, natural resources and self-employment are incorporated, however this is to be expected given that land is not included in most IO tables and hence needs to be incorporated.

The issues associated with the trade data need to be examined more carefully, and decisions made about whether more adjustments need to be made for re-exports and/or whether total trade should be adjusted to match known country-specific totals, with the current methodology used to obtain the bilateral detail.

The data used in this paper were obtained after agricultural production targeting. Ideally it would be useful to see how agricultural production targeting also affects these results. Furthermore, econometric analysis might be useful in ascertaining the extent to which certain factors, like age of the table and number of sectors do affect the extent to which the IO tables are changed as a result of the construction process.

Appendix I Cost and Sales Shares by Country

Table A 1: Summary Statistics for Percentage Point Differences between Initial and Final Cost Shares by Region (sorted by Standard Deviation)

region	Standard Deviation	min	max	90% Confidence Interval		Entropy
				lower	upper	
uga	5.50	-95.41	92.68	-13.75	13.75	2.06
moz	5.19	-99.83	97.39	-12.97	12.97	2.05
zmb	4.99	-74.98	94.17	-12.46	12.46	2.02
lux	4.87	-67.12	99.26	-12.17	12.17	1.74
lka	4.67	-81.27	97.33	-11.68	11.68	2.17
per	4.65	-70.89	71.54	-11.62	11.62	1.40
mwi	4.59	-56.17	86.69	-11.48	11.48	3.00
mlt	4.46	-52.36	93.00	-11.14	11.14	2.19
mys	4.37	-94.90	83.59	-10.89	10.97	1.40
cyp	4.26	-71.37	83.93	-10.66	10.66	2.08
tza	4.26	-76.36	87.70	-10.65	10.65	2.15
ury	4.20	-63.97	99.38	-10.50	10.50	1.58
ven	4.19	-94.69	99.53	-10.48	10.48	1.96
bgr	4.18	-49.23	89.48	-10.45	10.45	2.16
xea	3.95	-47.52	84.94	-9.88	9.89	1.52
mdg	3.95	-62.80	89.71	-9.88	9.88	1.67
est	3.91	-79.30	81.65	-9.79	9.79	1.31
xsu	3.87	-45.41	71.15	-9.67	9.67	2.20
sgp	3.84	-69.04	82.63	-9.61	9.61	2.14
phl	3.84	-57.74	81.64	-9.56	9.65	1.10
idn	3.80	-96.26	83.17	-9.47	9.53	0.87
ltu	3.74	-28.83	99.47	-9.35	9.35	1.69
zwe	3.70	-73.31	77.37	-9.26	9.26	1.27
grc	3.62	-67.19	64.68	-9.04	9.04	1.25
xsa	3.57	-46.63	75.62	-8.92	8.92	1.39
xse	3.56	-60.29	57.17	-8.87	8.92	0.81
col	3.56	-61.01	54.72	-8.90	8.90	1.19
fra	3.52	-74.15	85.19	-8.79	8.79	0.90
nld	3.48	-69.92	95.33	-8.70	8.70	1.17
vnm	3.45	-52.38	66.88	-8.63	8.63	1.19
ita	3.44	-55.45	78.08	-8.60	8.60	0.88
bel	3.42	-54.55	72.49	-8.54	8.54	1.32
xsm	3.37	-62.55	65.24	-8.42	8.42	0.97
irl	3.33	-70.38	64.88	-8.32	8.32	1.33

region	Standard Deviation	min	max	90% Confidence Interval		Entropy
				lower	upper	
esp	3.31	-62.38	98.92	-8.27	8.27	0.93
pol	3.29	-60.55	73.78	-8.24	8.24	0.96
mex	3.29	-74.01	42.70	-8.23	8.23	0.85
prt	3.28	-51.15	71.80	-8.21	8.21	1.17
deu	3.24	-50.53	81.24	-8.10	8.10	1.07
xsc	3.22	-55.58	86.82	-8.06	8.06	1.04
xca	3.22	-43.21	81.99	-8.04	8.04	1.04
tha	3.21	-62.60	48.91	-7.99	8.05	0.71
zaf	3.18	-60.13	91.10	-7.96	7.96	0.83
xna	3.13	-51.99	81.20	-7.83	7.83	0.76
chl	3.13	-58.87	71.41	-7.82	7.82	0.94
lva	3.11	-39.59	76.51	-7.78	7.78	0.98
swe	3.10	-57.88	72.36	-7.74	7.74	1.02
mar	3.09	-89.32	81.91	-7.72	7.72	0.84
rus	3.08	-50.84	94.32	-7.70	7.70	0.85
hun	3.02	-35.83	74.44	-7.55	7.55	1.13
kor	3.00	-72.25	45.54	-7.50	7.50	0.72
svn	3.00	-42.01	82.77	-7.50	7.50	1.14
cze	2.99	-61.82	70.68	-7.47	7.47	1.22
xnf	2.98	-44.09	64.32	-7.44	7.44	1.12
xcb	2.91	-53.52	55.99	-7.27	7.27	0.74
xef	2.90	-48.20	66.34	-7.24	7.24	0.95
che	2.88	-42.03	82.36	-7.21	7.21	1.06
hkg	2.87	-83.49	98.05	-7.16	7.16	0.60
fin	2.86	-73.91	48.86	-7.14	7.14	0.73
svk	2.81	-33.81	82.63	-7.03	7.03	0.97
bgd	2.67	-41.87	85.75	-6.68	6.68	1.10
dnk	2.66	-49.53	50.66	-6.64	6.64	0.92
gbr	2.65	-43.68	97.48	-6.64	6.64	0.84
jpn	2.53	-47.70	83.87	-6.33	6.33	0.54
xer	2.51	-59.04	45.91	-6.28	6.28	0.50
bra	2.46	-45.16	45.31	-6.16	6.16	0.72
twn	2.41	-42.30	97.33	-6.02	6.04	0.39
hrv	2.39	-55.75	41.25	-5.98	5.98	0.59
tun	2.34	-48.57	48.20	-5.85	5.85	0.58
tur	2.24	-52.47	40.55	-5.61	5.61	0.57
usa	2.23	-43.90	46.36	-5.57	5.57	0.49
can	2.18	-52.22	42.84	-5.45	5.45	0.62
ind	2.17	-36.16	41.99	-5.43	5.43	0.76
xoc	2.15	-37.63	86.10	-5.38	5.38	0.54
chn	1.98	-36.62	28.89	-4.94	4.94	0.72
alb	1.84	-42.32	59.44	-4.59	4.59	0.22

region	Standard Deviation	min	max	90% Confidence Interval		Entropy
				lower	upper	
aus	1.78	-37.57	32.38	-4.45	4.45	0.48
arg	1.64	-30.36	26.11	-4.11	4.11	0.34
bwa	1.60	-31.07	47.08	-4.00	4.00	0.45
nzl	1.41	-34.70	26.22	-3.53	3.53	0.30

Source: Authors' calculations

Table A 2: Summary Statistics for Percentage Point Differences between Initial and Final Sales Shares by Region (sorted by standard deviation)

region	Standard Deviation	min	max	90% Confidence Interval		Entropy
				lower	upper	
mwi	9.26	-99.07	98.91	-23.16	23.15	6.44
uga	9.18	-96.35	99.90	-22.96	22.96	5.93
lka	8.58	-96.07	99.98	-21.45	21.46	4.83
ven	7.78	-88.28	99.88	-19.45	19.45	5.38
ury	7.47	-99.65	99.63	-18.67	18.67	3.65
mlt	7.46	-97.19	98.23	-18.66	18.66	3.62
xsa	7.34	-99.34	93.84	-18.35	18.35	3.57
cyp	7.29	-66.50	98.13	-18.22	18.22	3.70
xea	7.25	-99.45	97.72	-18.11	18.14	3.20
bgr	7.20	-74.64	97.11	-17.99	17.99	4.21
chl	7.17	-95.97	99.32	-17.93	17.93	2.93
xsu	7.13	-75.56	89.62	-17.83	17.83	4.84
zmb	7.08	-97.92	99.69	-17.69	17.69	3.53
mys	7.05	-97.47	99.47	-17.60	17.63	3.42
xca	6.92	-99.14	98.49	-17.29	17.29	2.56
xsm	6.74	-98.56	98.28	-16.86	16.86	2.36
lux	6.72	-95.44	91.32	-16.80	16.80	2.56
xsc	6.63	-96.47	94.66	-16.58	16.58	2.88
phl	6.55	-99.46	96.62	-16.36	16.39	2.63
moz	6.48	-86.62	85.02	-16.20	16.20	2.72
grc	6.46	-91.86	86.66	-16.15	16.15	2.54
tza	6.44	-61.06	94.20	-16.10	16.10	3.20
vnm	6.36	-98.62	97.84	-15.89	15.89	2.41
per	6.33	-99.98	89.05	-15.83	15.83	2.64
nld	6.33	-99.67	95.87	-15.83	15.83	2.41
che	6.29	-94.13	95.22	-15.72	15.72	1.93
zwe	6.27	-95.05	81.31	-15.69	15.69	2.25
tha	6.27	-89.95	96.38	-15.66	15.69	2.33
deu	6.25	-94.28	93.22	-15.62	15.62	2.05
xef	6.24	-83.61	93.33	-15.59	15.59	2.22
swe	6.21	-96.18	92.63	-15.52	15.52	2.19
irl	6.19	-81.83	89.71	-15.49	15.49	2.25
sgp	6.19	-77.80	98.82	-15.48	15.48	2.51
bra	6.16	-98.77	99.28	-15.40	15.40	2.13
est	6.11	-84.21	96.96	-15.27	15.27	2.82
xse	6.04	-77.49	87.57	-15.09	15.12	1.75
idn	6.02	-94.23	92.63	-15.03	15.05	2.32
ltu	6.01	-92.38	99.42	-15.02	15.02	2.33
col	5.97	-87.80	99.49	-14.92	14.92	2.46
ind	5.91	-97.83	92.26	-14.79	14.79	2.03

region	Standard Deviation	min	max	90% Confidence Interval		Entropy
				lower	upper	
bel	5.88	-75.28	89.44	-14.71	14.71	2.30
hun	5.87	-94.15	95.46	-14.67	14.67	2.27
pol	5.77	-89.84	97.83	-14.43	14.43	1.93
hkg	5.74	-69.71	95.28	-14.35	14.35	1.33
mdg	5.74	-80.36	98.84	-14.35	14.35	2.11
dnk	5.70	-87.12	91.13	-14.25	14.25	2.00
xnf	5.67	-98.13	75.85	-14.17	14.17	2.02
fin	5.63	-85.36	83.20	-14.08	14.08	1.79
svn	5.62	-92.31	92.62	-14.04	14.04	2.35
mex	5.59	-81.10	88.52	-13.99	13.99	1.81
rus	5.55	-91.87	78.75	-13.89	13.89	1.99
ita	5.54	-72.40	93.21	-13.84	13.84	2.04
esp	5.49	-97.43	98.31	-13.73	13.73	1.74
xcb	5.48	-97.78	86.23	-13.69	13.69	1.64
zaf	5.42	-93.60	88.10	-13.55	13.55	1.90
prt	5.41	-75.01	96.19	-13.52	13.52	2.02
svk	5.39	-63.13	88.45	-13.47	13.47	1.93
tur	5.32	-82.99	97.21	-13.30	13.30	1.65
mar	5.18	-99.29	97.74	-12.95	12.95	1.47
can	5.08	-94.27	63.26	-12.69	12.69	1.66
xna	4.93	-62.36	62.29	-12.33	12.33	1.36
fra	4.91	-75.46	97.97	-12.27	12.27	1.75
cze	4.91	-44.38	82.90	-12.26	12.26	2.10
gbr	4.80	-87.40	80.69	-12.00	12.00	1.57
lva	4.65	-67.32	92.99	-11.62	11.62	1.51
xer	4.61	-67.34	71.46	-11.53	11.53	1.13
jpn	4.55	-94.66	94.76	-11.38	11.38	1.06
chn	4.44	-95.73	93.42	-11.10	11.10	1.36
bgd	4.29	-77.29	99.19	-10.73	10.73	1.01
kor	4.18	-98.15	84.55	-10.46	10.46	0.94
xoc	4.18	-75.73	61.63	-10.46	10.46	0.95
usa	3.80	-60.62	86.97	-9.51	9.51	0.95
aus	3.80	-90.48	69.25	-9.50	9.50	1.02
arg	3.64	-92.59	73.45	-9.11	9.11	0.79
hrv	3.40	-65.71	96.81	-8.51	8.51	0.55
bwa	3.35	-87.65	93.19	-8.38	8.38	0.46
alb	3.30	-77.33	63.69	-8.26	8.26	0.56
tun	3.29	-65.20	70.52	-8.23	8.23	0.53
nzl	3.26	-59.92	69.62	-8.16	8.16	0.60
twn	2.47	-38.08	79.98	-6.17	6.20	0.38

Source: Authors' calculations

Appendix II Cost and Sales Shares by Inputs and Uses

Table A 3: Summary Statistics for Percentage Point Differences between Initial and Final Cost Shares by Input (sorted by standard deviation)

input	Mean	Standard Deviation	min	max	90% Confidence Interval		Entropy
					lower	upper	
prim_cap	-4.36	18.26	-96.26	89.48	-50.01	41.29	9.68
prim_lab	-0.96	15.49	-99.83	88.59	-39.68	37.76	7.17
prim_land	10.57	12.20	-48.97	45.40	-19.93	41.07	77.57
trd_d	-0.42	6.67	-34.23	91.49	-17.08	16.25	3.51
ely_d	0.51	6.39	-43.68	93.00	-15.46	16.47	2.80
oil_m	0.23	5.16	-52.47	87.70	-12.66	13.12	1.61
p_c_m	0.57	4.94	-69.04	98.05	-11.79	12.93	2.35
p_c_d	0.45	4.66	-61.01	68.54	-11.21	12.11	2.40
obs_d	1.19	4.61	-64.91	65.31	-10.34	12.71	2.83
gas_m	0.24	4.21	-37.57	99.47	-10.29	10.78	1.58
ofi_d	-0.37	4.16	-51.15	53.03	-10.77	10.03	1.89
oap_d	-0.11	4.06	-63.97	83.87	-10.25	10.04	1.55
oil_d	0.08	4.04	-47.16	87.46	-10.01	10.17	1.23
ofd_d	-0.11	4.00	-48.68	52.15	-10.12	9.90	1.95
otn_m	0.30	3.74	-49.84	81.99	-9.04	9.65	1.17
rmk_d	-0.14	3.70	-67.12	66.64	-9.40	9.12	1.57
pdr_m	0.10	3.64	-59.04	99.26	-9.01	9.21	0.80
ctl_d	-0.05	3.63	-73.31	61.34	-9.12	9.01	1.13
pdr_d	0.07	3.55	-71.37	99.53	-8.80	8.94	0.63
ofd_m	0.34	3.51	-16.30	85.76	-8.42	9.10	1.18
c_b_d	0.08	3.49	-76.36	83.59	-8.64	8.80	0.91
wht_m	0.10	3.32	-36.42	97.33	-8.21	8.41	0.87
otp_d	0.09	3.29	-39.77	45.11	-8.13	8.31	1.49
crp_m	0.52	3.18	-25.66	46.20	-7.44	8.48	1.51
osd_m	0.19	3.15	-48.97	97.33	-7.68	8.05	0.70
i_s_m	0.23	3.13	-57.88	61.62	-7.59	8.05	0.99
coa_d	0.00	3.11	-46.53	86.04	-7.76	7.77	1.19
gas_d	0.00	3.07	-56.95	71.15	-7.67	7.68	1.35
osg_m	0.24	3.00	-4.19	92.58	-7.27	7.74	1.65
cns_d	-0.06	2.98	-40.35	63.58	-7.52	7.40	1.22
osg_d	-0.16	2.95	-77.78	49.02	-7.54	7.21	1.38
ocr_d	0.03	2.87	-35.45	83.17	-7.16	7.21	1.43
ros_d	-0.40	2.67	-34.75	80.96	-7.08	6.27	1.66
crp_d	-0.42	2.64	-25.66	57.71	-7.01	6.18	1.28

input	Mean	Standard Deviation	min	max	90% Confidence Interval		Entropy
					lower	upper	
v_f_d	-0.18	2.55	-70.72	43.06	-6.55	6.19	1.23
mvh_m	-0.02	2.53	-65.89	69.28	-6.34	6.29	0.84
tex_m	0.25	2.50	-47.51	38.46	-5.99	6.49	0.62
tex_d	-0.26	2.49	-61.17	29.60	-6.48	5.96	0.59
gro_d	0.00	2.43	-41.10	86.69	-6.08	6.08	0.78
ome_m	0.27	2.39	-26.46	26.97	-5.69	6.23	1.14
omn_d	-0.07	2.32	-74.98	29.33	-5.86	5.73	0.79
ele_m	0.06	2.30	-43.31	46.39	-5.70	5.81	0.72
osd_d	0.07	2.29	-30.86	40.31	-5.65	5.79	0.62
otp_m	0.07	2.15	-37.86	88.62	-5.30	5.44	0.67
frs_d	0.02	2.13	-26.50	56.21	-5.31	5.35	0.63
v_f_m	0.04	2.12	-51.88	82.36	-5.27	5.34	0.76
gdt_d	-0.12	2.04	-37.19	50.53	-5.21	4.97	1.36
ofi_m	0.18	2.01	-7.28	73.20	-4.84	5.20	0.69
lum_m	0.23	1.91	-9.53	60.37	-4.55	5.01	0.54
i_s_d	-0.17	1.89	-36.23	20.44	-4.88	4.54	0.60
nfm_d	-0.13	1.87	-32.82	60.58	-4.81	4.56	0.62
isr_d	-0.03	1.87	-20.16	43.08	-4.70	4.65	0.65
vol_m	0.06	1.83	-18.52	84.94	-4.51	4.62	0.48
omt_d	-0.14	1.82	-41.91	43.35	-4.69	4.40	0.76
omn_m	0.02	1.77	-24.54	66.74	-4.40	4.44	0.65
ros_m	-0.02	1.77	-49.53	55.81	-4.43	4.40	0.99
mil_d	0.13	1.74	-41.80	31.80	-4.21	4.47	0.53
nfm_m	0.01	1.72	-35.98	23.23	-4.28	4.30	0.51
ocr_m	0.04	1.71	-48.72	50.57	-4.23	4.30	0.63
ome_d	-0.13	1.70	-23.60	16.28	-4.37	4.12	0.83
cmt_d	-0.13	1.67	-31.20	54.14	-4.32	4.05	0.61

Table A 4: Summary Statistics for Percentage Point Differences between Initial and Final Sales Shares by Use (sorted by standard deviation)

Use	Mean	Standard Deviation	Min	max	90% Confidence Interval lower	90% Confidence Interval upper	Entropy
fin_ch	1.23	23.88	-99.45	99.98	-58.47	60.94	16.17
fin_x	2.62	22.88	-98.62	98.99	-54.58	59.83	15.53
ofd_i	-0.22	10.63	-95.44	95.55	-26.79	26.36	5.70
p_c_i	0.25	9.42	-98.59	99.69	-23.29	23.80	4.29
fin_if	-0.38	9.39	-99.24	90.54	-23.87	23.10	6.14
ely_i	0.50	9.36	-96.66	98.82	-22.91	23.91	4.19
trd_i	0.83	8.42	-94.09	96.62	-20.23	21.88	7.01
fin_cg	0.08	7.42	-99.98	98.23	-18.48	18.64	5.40
crp_i	-0.16	7.28	-86.61	99.90	-18.36	18.05	4.41
tex_i	-0.14	6.99	-98.85	96.96	-17.61	17.33	2.59
cns_i	0.13	6.94	-77.91	99.28	-17.22	17.47	3.46
osg_i	-0.12	6.61	-96.18	99.63	-16.64	16.39	4.19
mil_i	0.65	6.00	-71.53	99.47	-14.36	15.66	2.68
sgr_i	-0.06	5.85	-91.86	99.88	-14.70	14.57	2.09
cmt_i	-0.29	5.71	-99.46	95.49	-14.57	14.00	2.06
omt_i	-0.20	5.26	-66.02	94.76	-13.35	12.95	2.23
b_t_i	-0.17	5.12	-95.05	98.63	-12.97	12.64	2.60
pcr_i	-0.06	5.00	-94.13	88.22	-12.57	12.44	1.50
ros_i	-0.40	4.95	-95.73	86.97	-12.77	11.96	2.96
vol_i	-0.16	4.77	-94.65	88.82	-12.07	11.76	1.74
otp_i	0.04	4.74	-80.43	61.20	-11.80	11.88	2.21
obs_i	0.25	4.64	-75.02	62.88	-11.36	11.86	2.68
ofi_i	-0.33	4.62	-83.99	54.07	-11.87	11.21	1.64
oil_i	0.28	4.54	-62.71	99.49	-11.07	11.64	3.04
oap_i	-0.29	4.24	-84.95	82.19	-10.88	10.30	1.75
nmm_i	-0.21	4.20	-75.68	99.19	-10.72	10.30	1.72
nfm_i	0.16	4.08	-80.51	91.33	-10.05	10.37	1.39
lum_i	-0.05	4.06	-81.26	96.93	-10.20	10.11	1.06
lea_i	-0.15	4.06	-99.67	94.67	-10.30	10.00	1.38
wap_i	0.00	4.04	-97.43	85.73	-10.11	10.11	1.61
cmn_i	0.06	3.67	-97.83	88.64	-9.12	9.25	0.87
i_s_i	-0.10	3.54	-72.91	79.98	-8.95	8.75	1.38
wtp_i	0.08	3.52	-50.09	81.56	-8.71	8.87	1.15
omf_i	-0.15	3.38	-81.99	82.37	-8.59	8.29	1.63
gdt_i	-0.31	3.31	-98.15	44.76	-8.59	7.98	1.62
ctl_i	-0.18	3.18	-71.66	90.96	-8.14	7.78	1.30
omn_i	-0.01	3.16	-62.30	97.74	-7.91	7.89	1.39
ele_i	-0.26	3.11	-58.71	69.80	-8.04	7.51	1.25
rmk_i	-0.11	3.09	-92.57	51.18	-7.84	7.63	1.27
isr_i	-0.03	3.02	-82.48	91.80	-7.57	7.52	0.94

Use	Mean	Standard Deviation	Min	max	90% Confidence Interval		Entropy
					lower	upper	
ppp_i	-0.15	3.00	-64.49	86.17	-7.65	7.34	1.13
atp_i	0.02	2.85	-83.91	73.63	-7.12	7.15	0.98
v_f_i	-0.02	2.70	-62.37	63.01	-6.78	6.75	1.18
ome_i	-0.04	2.62	-50.32	53.70	-6.60	6.51	1.16
ocr_i	-0.11	2.48	-54.39	71.18	-6.31	6.09	0.97
mvh_i	-0.10	2.28	-70.02	56.60	-5.79	5.59	0.93
c_b_i	-0.12	2.16	-85.69	54.66	-5.51	5.27	0.80
fmp_i	-0.18	2.11	-31.75	39.80	-5.47	5.10	0.88
fsh_i	-0.06	2.10	-61.02	60.31	-5.31	5.18	1.18
gas_i	0.03	2.00	-31.89	61.22	-4.96	5.02	1.04
wol_i	-0.04	2.00	-56.79	66.52	-5.03	4.94	0.49
frs_i	-0.07	1.99	-34.12	39.77	-5.05	4.91	0.91
pdr_i	-0.02	1.86	-86.17	75.60	-4.66	4.62	0.38
dwe_i	-0.09	1.71	-51.07	38.87	-4.37	4.18	0.94
gro_i	-0.10	1.71	-84.79	24.82	-4.36	4.17	0.56
wtr_i	-0.04	1.63	-59.00	71.36	-4.12	4.03	0.50
coa_i	-0.03	1.63	-46.53	69.44	-4.09	4.03	0.97
osd_i	-0.03	1.56	-74.34	33.21	-3.93	3.87	0.38
otn_i	-0.02	1.54	-33.05	45.87	-3.87	3.82	0.53
wht_i	-0.05	1.46	-37.98	54.79	-3.70	3.59	0.44
pfb_i	0.00	1.32	-41.45	47.04	-3.30	3.31	0.41

Source: Authors' calculations

Appendix III Basic Information on Contributed IO tables

Table A 5: Basic Information about the contributed Table

	GTAP Version IO table was contributed	Total No. of sectors in contribution (max 57)	No. of Agricultural sectors (max 12)	No. of Processed food sectors (max 7)	No. of Manufacturing sectors (max 37)
AUS	8	54	11	6	37
NZL	9	41	4	4	33
CHN	9	45	2	7	36
HKG	1	37	6	5	26
JPN	8.1	57	12	8	37
KOR	8.1	56	12	7	37
MNG	8	34	2	3	29
TWN	8.1	57	12	8	37
BRN	9	34	1	1	32
KHM	7	57	12	8	37
IDN	7	53	10	7	36
LAO	7	31	12	1	18
MYS	8	46	6	5	35
PHL	8	50	9	6	35
SGP	8.1	43	2	5	36
THA	8	51	9	7	35
VNM	7.1	47	6	7	34
BGD	5	57	12	8	37
IND	8	50	10	4	36
NPL	8	57	12	8	37
PAK	9	32	8	4	20
LKA	7	29	5	3	21
CAN	7	51	8	8	35
USA	7	57	12	8	37
MEX	8	37	2	2	33
ARG	6	57	12	8	37
BOL	7.1	33	4	5	24
BRA	8.1	52	11	8	33
CHL	7	40	3	5	32
COL	8.1	56	12	7	37
ECU	7	41	5	5	31
PRY	9	36	7	6	23
PER	7	50	12	8	30
URY	7	50	12	8	30

	GTAP Version IO table was contributed	Total No. of sectors in contribution (max 57)	No. of Agricultural sectors (max 12)	No. of Processed food sectors (max 7)	No. of Manufacturing sectors (max 37)
VEN	8	46	4	6	36
CRI	7	37	9	7	21
GTM	7	50	12	7	31
HND	8	42	7	7	28
NIC	7	34	5	5	24
PAN	7	27	10	2	15
SLV	8	38	7	6	25
DOM	9	30	4	4	22
JAM	9	35	6	5	24
PRI	9	38	1	7	30
TTO	9	43	7	7	29
AUT	7.1	54	12	8	34
BEL	7.1	54	12	8	34
CYP	7.1	54	12	8	34
CZE	7.1	54	12	8	34
DNK	7.1	54	12	8	34
EST	7.1	54	12	8	34
FIN	7.1	54	12	8	34
FRA	7.1	54	12	8	34
DEU	7.1	54	12	8	34
GRC	7.1	54	12	8	34
HUN	7.1	54	12	8	34
IRL	7.1	54	12	8	34
ITA	7.1	54	12	8	34
LVA	7.1	54	12	8	34
LTU	7.1	54	12	8	34
LUX	7.1	54	12	8	34
MLT	7.1	54	12	8	34
NLD	7.1	54	12	8	34
POL	7.1	54	12	8	34
PRT	7.1	54	12	8	34
SVK	7.1	54	12	8	34
SVN	7.1	54	12	8	34
ESP	7.1	54	12	8	34
SWE	7.1	54	12	8	34
GBR	7.1	54	12	8	34
CHE	8	30	3	1	26
NOR	8	38	3	1	34
ALB	5.3	57	12	8	37
BGR	7.1	54	12	8	34

	GTAP Version IO table was contributed	Total No. of sectors in contribution (max 57)	No. of Agricultural sectors (max 12)	No. of Processed food sectors (max 7)	No. of Manufacturing sectors (max 37)
BLR	8.1	40	12	8	20
HRV	5.2	57	12	8	37
ROU	7.1	54	12	8	34
RUS	7	41	4	4	33
UKR	7	46	1	8	37
KAZ	7	34	1	1	32
KGZ	7	31	1	1	29
ARM	7	30	6	4	20
AZE	7	32	1	1	30
GEO	7	57	12	8	37
BHR	8	37	5	1	31
IRN	7	43	6	3	34
ISR	8	43	9	5	29
JOR	9	45	6	7	32
KWT	8	37	5	1	31
OMN	8	37	5	1	31
QAT	8	37	5	1	31
SAU	8	37	5	1	31
TUR	9	42	1	6	35
ARE	8	37	5	1	31
EGY	6.2	31	3	4	24
MAR	7	46	12	6	28
TUN	6	37	3	5	29
BEN	8.1	29	5	1	23
BFA	8.1	38	8	6	24
CMR	8	57	12	8	37
CIV	8	34	4	2	28
GHA	8	41	9	4	28
GIN	8.1	30	9	3	18
NGA	8.1	38	12	4	22
SEN	9	31	2	4	25
TGO	8.1	30	3	4	23
ETH	7.1	39	6	5	28
KEN	8	33	10	3	20
MDG	6	57	12	8	37
MWI	8.1	36	9	4	23
MUS	6.1	42	11	5	26
MOZ	8.1	37	9	3	25
RWA	8.1	34	10	4	20
TZA	8.1	39	11	6	22

	GTAP Version IO table was contributed	Total No. of sectors in contribution (max 57)	No. of Agricultural sectors (max 12)	No. of Processed food sectors (max 7)	No. of Manufacturing sectors (max 37)
UGA	8	33	3	2	28
ZMB	8.1	36	9	4	23
ZWE	5.1	57	12	8	37
BWA	5	57	12	8	37
NAM	8	29	3	2	24
ZAF	7	41	3	6	32

Source: Annual Report on the Regional I-O Tables in the GTAP Data Base

Appendix IV Notation used

Table A 6: Sectoral listing

Number	Code	Description
1	PDR	Paddy rice
2	WHT	Wheat
3	GRO	Cereal grains nec
4	V_F	Vegetables, fruit, nuts
5	OSD	Oil seeds
6	C_B	Sugar cane, sugar beet
7	PFB	Plant-based fibers
8	OCR	Crops nec
9	CTL	Bovine cattle, sheep and goats, horses
10	OAP	Animal products nec
11	RMK	Raw milk
12	WOL	Wool, silk-worm cocoons
13	FRS	Forestry
14	FSH	Fishing
15	COA	Coal
16	OIL	Oil
17	GAS	Gas
18	OMN	Minerals nec
19	CMT	Bovine meat products
20	OMT	Meat products nec
21	VOL	Vegetable oils and fats
22	MIL	Dairy products
23	PCR	Processed rice
24	SGR	Sugar
25	OFD	Food products nec
26	B_T	Beverages and tobacco products
27	TEX	Textiles
28	WAP	Wearing apparel
29	LEA	Leather products
30	LUM	Wood products
31	PPP	Paper products, publishing
32	P_C	Petroleum, coal products
33	CRP	Chemical, rubber, plastic products
34	NMM	Mineral products nec
35	I_S	Ferrous metals
36	NFM	Metals nec
37	FMP	Metal products
38	MVH	Motor vehicles and parts

Number	Code	Description
39	OTN	Transport equipment nec
40	ELE	Electronic equipment
41	OME	Machinery and equipment nec
42	OMF	Manufactures nec
43	ELY	Electricity
44	GDT	Gas manufacture, distribution
45	WTR	Water
46	CNS	Construction
47	TRD	Trade
48	OTP	Transport nec
49	WTP	Water transport
50	ATP	Air transport
51	CMN	Communication
52	OFI	Financial services nec
53	ISR	Insurance
54	OBS	Business services nec
55	ROS	Recreational and other services
56	OSG	Public Administration, Defense, Education, Health
57	DWE	Dwellings

Table A 7: Other codes for assistance

Number	Code	Description
1	_d	Post-script indicating Domestic variety
2	_m	Post-script indicating Imported variety
3	_i	Post-script indicating Industry or activity
4	prim_land	Land
5	prim_cap	Capital
6	prim_lab	Labor
7	fin_x	Exports
8	fin_cg	Government consumptions
9	fin_ch	Private household consumptions
10	fin_if	Investment

Bibliography

Bacharach, M. (1970). Biproportional matrices and input-output change, Cambridge.

Boumellassa, H., D. Laborde and C. Mitaritonna (2009). A picture of tariff protection across the World in 2004 : MAcMap-HS6, Version 2. IFPRI Discussion Paper 903. Washington DC, International Food Policy Research Institute (IFPRI).

European Commission, IMF, OECD, United Nations and W. Bank (2008). "System of National Accounts."

Gehlhar, M., Z. Wang and S. Yao (2008). Reconciling Merchandise Trade Data. Global Trade, Assistance, and Production: The GTAP 7 Data Base. B. Narayanan and T. L. Walmsley. West Lafayette, IN, Center for Global Trade Analysis, Purdue University.

Harslett, P., (2013). "The GTAP Data Base Construction Procedure." GTAP Working paper, 76, Center for Global Trade Analysis, Purdue University, West Lafayette, IN, USA

Huff, K. M., R. McDougall and T. L. Walmsley (1998). Contributing IO tables to the GTAP Data Base GTAP Technical Paper Series. C. f. G. T. Analysis. West Lafayette, IN.

James, M. and R. McDougall (1993). FIT: An input-output data update facility for SALTER. SALTER working paper. A. I. Commission.

Marcel, T., A. A. Erumban, G. J. d. Vries, I. Arto, V. A. A. Genty, J. Frederik Neuwahl, M. Rueda-Cantuche, A. Villanueva, J. Francois, O. Pindyuk, J. Pöschl, R. Stehrer and G. Streicher (2012). The World Input-Output Database (WIOD): Contents, Sources and Methods. Working Paper Number.

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

Organisation for Economic Cooperation and Development and International Energy Agency (2003). IEA Online Database Services, Organisation for Economic Cooperation and Development and International Energy Agency, Washington D.C.

Theil, H. (1967). Economics and information theory. Amsterdam, North-Holland.

Tukker, A., E. Poliakov, R. Heijungs, T. Hawkins, F. Neuwahl, J. Rueda-Cantuche, S. Giljum, S. Moll, J. Oosterhaven and M. Bouwmeester (2009). "Towards a global multi-regional environmentally extended input-output database." Ecological Economics **68**(7): 1929-1937.

Walmsley, T. L. and R. McDougall (2004). Using Entropy to Compare Shares. GTAP Research Memorandum. C. f. G. T. Analysis. West Lafayette, IN, 9.

Walmsley, T. L., A. Aguiar and B. G. Narayanan (2013). "A Global Dataset of Input-Output Tables Linked by International Trade and Policy Data." GTAP Working Paper, 67, Center for Global Trade Analysis, Purdue University, IN, USA