How a Global Inter-Country Input-Output Table with Processing Trade Account Can be constructed from GTAP Database

Marinos Tsigas and Zhi Wang
U.S. International Trade Commission*
Mark Gehlhar
U.S. Department of Interior*

(Preliminary draft, not for quotation)

Abstract

We developed a method to construct a global ICIO table from version 8 GTAP database as well as detailed trade data from UN COMTRADE, and two additional IO tables for major emerging economies where processing exports are a large portion of their external trade. We integrate the GTAP database and the additional information with a quadratic mathematical programming model that (a) minimizes the deviation of the resulting new data set from the original GTAP data, (b) ensures that supply and use balance for each sector and every country, and (c) keeps all sectoral bilateral trade flows in the GTAP database constant. Bilateral and aggregate reliability indexes are computed for each GTAP sectors and end use categories which are used to control the relative amount of adjustment for each end-use categories within each original bilateral trade flows from the GTAP database. The new database covers 63 countries/regions and 41 sectors for 2004 and 2007 two years.

Paper for Presentation at the 15th GTAP Conference on Global Economic Analysis

International Trade Center, Geneva June 28, 2012

*The views expressed in this paper are those of the authors alone. They do not necessarily reflect the views of the US International Trade Commission, or U.S. Department of Interior.

1. **Introduction**

There is resurgence in the applications of input-output (I-O) tables in the economic literature during recent years for both analytical and statistical purpose (Norihiko Yamano and Nadim Ahmad, 2006). As an analytical data source and accounting framework, input-output tables provide consistent analysis and measurement of vertical specialization of international trade (Hummels, Ishii, and Yi, 2001), domestic and foreign contents in a country's gross exports (Koopman, Wang and Wei, 2008, 2012), the development of value-chain in global production network (Wang, Power and Wei, 2009), the pattern of gross versus value-added trade around the world (Johnson and Noguera, 2009), the decomposition of gross trade to their value-added contents (Koopma, Powers, Wang and Wei, 2010), and trade flows in intermediate goods and services among OECD countries (Sébastien Miroudot, Rainer Lanz, and Alexandros Ragoussis, 2010). It is also increasingly being used in environmental analysis such as measuring direct and indirect pollutants produced by industrial sectors within an economy and estimate consumption-based emissions, thus accounting 'leakages' between economies (Davis and Caldeira, 2010), as well as policy debates on the role of vertical specialization in the dramatic decline of world trade during recent global financial crisis(Bems, Johnson, and Yi, 2010) and the economic and trade impact of Japan's recent earthquake and tsunami (Escaish, Keck, Nee and Teh, 2011). As a statistical analysis tool, input-output and the closely related supply-use tables are increasingly becoming the most important vehicles used to balance the income, expenditure and production estimates of GDP to satisfy the United Nation standards of System of National Account (SNA) 1993 and 2008.

However, contrast with this surged analytical and statistical demand, the lack of consistent global II-O data sets, especially such data with a time dimension remain as a major obstacle for many economists to address the various issues mentioned above at their hands. This is because global II-O tables are very rare due to the tremendous amount of data required and the differences in statistical classifications across countries. Most existing global I-O databases are a collection of individual country tables such as various version of OECD STAIN I-O database[1],

[1] It provides a bulk of the required data and is regularly compiled for about 50 countries across the globe, but integrating them with bilateral trade statistics into global consistent database still remains as a substantial challenge.

the few available II-O tables, such as the Asian international I-O table compiled by the Institute of Development Economies (IDE) in Japan, cover only a select set of Asian economies and treat other countries (including EU) in the rest of the world as exogenous blocks. In addition, its publication has a significant time delay (the available most recent table is 2000) and its industry classification is ad-hoc, not very easy concord to common used international classifications, so make it very difficult for update using statistics published by UN and other international agencies. Progress has been made in recent years. Most developed countries, such as the 27 European Union member states and the United States, now compile and publish annual supply and use tables. Major initiatives are under way to help developing countries to comply with the 1993 System of National Accounts (SNA), including publishing supply and use tables.[2] The European Commission, has funded a consortium of eleven European research institutions to develop a worldwide time series of national input-output tables, called the World Input Output Database (or WIOD), that are fully linked through bilateral trade data (27 EU member and 13 other major economies), generating a time series, multi-country IO table (for 1995-2009). WIOD contains tables in both current and previous year's prices. The data set just become public accessible since April this year.[3] The OECD is also constructing an inter-country IO table for three benchmark years (1995, 2000 and 2005) by combining their individual country IO databases and STAN bilateral industry trade statistics, covering about 50 countries.

GTAP database is a public accessible global data set to facilitate contemporary applied general equilibrium analysis of global economic issues. It has a broader country and sector coverage than WIOD, with 57 sectors, 109 individual countries and 20 composite regions in its most recent version (version 8); It has also full global coverage and benchmarked on reconciled official trade statistics based on data reliability. For example, re-exports through Hong Kong are systemically adjusted to their origin and destination countries. It uses entropy theoretic methods

---

[2] ADB organized a project with participation of 17 developing countries (RETA 6483) in Asia Pacific to construct supply and use tables for each participating country.

[3] Despite many of advantages, such as improved allocation of imports by end use category; closely linked with EU KLEMS and World KLEMS and with better and detailed capital types and labor skill levels breakdown, there are also obvious shortcomings in the WIOD data set need to be further improved, such as its trade flows are based on import statistics only, and exports to the rest of the world is calculated as residuals and could become negative for some products; each country's data just simply put together and no reconciliation procedure based on data reliability has been used. In addition, while the coverage of the 27 EU member countries is detailed, less than 10 developing countries are included. Processing trade is also not considered.

to reconcile data from different sources and create a consistent database. This consistency is the core advantage of the GTAP data base offered to the CGE modeling community. However, benchmarked only on trade statistics, sector level supply and demand data for individual countries may have large discrepancies with corresponding statistics in national accounts[4]; There is no consistency imposed for different versions of the data, making it difficult to make over time comparisons. In addition, the II-O table underlying GTAP database is based on the so called Multi-Region Input-Output (MRIO) table in the literature, there is no distinction between intermediate and final goods and services trade flows in the data. Therefore, significant transformation has to be made in order to construct an Inter-Country IO (ICIO) table from the GTAP database.

This paper documents how an ICIO table with separate processing trade account can be constructed from GTAP database step by step. It starts with a specification of the mathematical relationship between MCIO and ICIO model, and discuss how exports subsidies, imports tariffs, commodity taxes and international transportation margin in the GTAP database should be treated in the corresponding MCIO and ICIO accounting framework; followed by presentation of a quadratic programming model with various reliability index in its objective function to separate gross bilateral trade flow in the GTAP database into intermediate, consumption and investment goods trade flows, thus transfer the MRIO table embodied in the GTAP database into a ICIO table. The initial allocation of bilateral trade flows in the GTAP database into the three end use categories is based on improved concordance between HS and UN BEC (Broad Economic Categories) and detailed trade statistics at 6-digit HS level from UN COMTRADE(Commodity Trade Statistics). Finally, a mathematical programming model that integrates processing trade information from major developing countries is introduced. China's expanded IO table with a separate accounts for processing exports from Koopman, Wang and Wei (2012) and 2003 Mexico IO table with separate domestic and Maquiladora accounts from Mexico statistical agency, Instituto Nacional de Estadística, Geografía e Informática (INEGI) are merged with the ICIO table constructed from GTAP database using the model by minimizing the deviation between the resulted new data set from original GTAP data. The new database covers 63

[4]For instance, the imports use by sector in the reference year does not correspond to the benchmark year of import matrix information published by the National Statistics Agencies.

countries and 41 sectors and was used to support our initial global AGE modeling of processing trade and global value-chain analysis efforts. The paper concludes with a discussion on remaining issue to be solved and directions of future work to further improve the data.

2. From Multi-country Input-Output (MCIO) Account to Inter-Country Input-Output (ICIO) Account

2.1 MCIO and ICIO accounts and their mathematical descriptions

Assume there are G countries, with N industries in each country. The production in each sector in any country can potentially use intermediate inputs from any sector (including its own) in any country. Assuming a predetermined location of production based on individual country's I-O table that defines the structure of the global production, the deliveries of goods and services between countries are determined by imbalances between supply and demand inside the different countries. A world MCIO table is a comprehensive account of annual transaction and payment flows within and between countries.

Following notation will be used to describe the elements of the world MCIO account:

x_i^r = Gross output of commodity ‘i’ in region ‘r’

v_i^r = Value added by production of commodity ‘i’ in region ‘r’

t_i^{sr} = Bilateral trade flows of commodity ‘i’ from source country ‘s’ to destination country ‘r’

z_{ij}^{rr} = Domestic intermediate demand of commodity ‘i’ by sector ‘j’ in country ‘r’

y_{ik}^{rr} = Domestic final demand of commodity ‘i’ by final demand type ‘k’ in country ‘r’

z_{ij}^{mr} = Demand of imported intermediates of commodity ‘i’ by sector ‘j’ in country ‘r’

y_{ik}^{mr} = Demand of final goods of commodity ‘i’ by final demand type ‘k’ in country ‘r’

All variables are measured in annual values. The total number of final demand types, such as private consumption or gross capital formation is K. Then the following three accounting identities describe the relationship among elements of each row (i, r) and column (j, s) of the global MCIO table can be specified as:

$$\sum_{i=1}^{N} z_{ij}^{rr} + \sum_{i=1}^{N} z_{ij}^{mr} + v_j^r = x_j^r \tag{1}$$

$$\sum_{j=1}^{N} z_{ij}^{ss} + \sum_{k=1}^{K} y_{ik}^{ss} + \sum_{r \neq s}^{G} t_i^{sr} = x_i^s \qquad (2)$$

$$\sum_{j=1}^{N} z_{ij}^{rm} + \sum_{k=1}^{K} y_{ik}^{rm} = \sum_{s=1}^{g} t_i^{sr} \qquad (3)$$

Equation (1) specifies the value of gross output of commodity “j” in country “r” is attributed to the value of all sector ‘j’ domestic and imported intermediate input purchases and to the value of services from sector ‘j’ primary factor inputs. Equation (2) indicates that total gross output of commodity “i” in the source country ‘s’ equals the sum of its product deliveries to domestic and international users, but there are no distinction about the type of end users in the international markets. Equation (3) indicates total intermediate and final import demand for commodity ‘i’ in destination country ‘r’ must be met by imports from all source countries. Thus, equations (1) – (3) together consistently defines an accounting framework for the global economy, conventionally called a MRIO table in the literature (Miller and Blair, 1985, Isard, et al. 1998). Such an account guarantees that international production and trade flows exactly meet all countries’ supply and demands, but stops short of assigning specific intermediate or final uses for international trade flows.

The above accounting framework can be extended to an ICIO account by further disaggregating gross goods and services trade flows t_i^{sr} by end use categories to sector and final users. Define:

z_{ij}^{sr} = Intermediate trade flows of commodity ‘i’ produced in source country ‘s’ for use by sector ‘j’ in destination country ‘r’;
y_{ik}^{sr} = Final goods and services trade flows of commodity ‘i’ produced in source country ‘s’ for type ‘k’ final use in destination country ‘r’;
mg_i^{sr} = Margin differences for a special transaction between the source country ‘s’ and destination country ‘r’. Then flowing three identifies will hold:

$$\sum_{s \neq r}^{G} (1 + mg_i^{sr}) z_{ij}^{sr} = z_{ij}^{mr} \qquad (4)$$

$$\sum_{s \neq r}^{G} (1 + mg_i^{sr}) y_{ik}^{sr} = y_{ik}^{mr} \qquad (5)$$

$$\sum_{j=1}^{N} z_{ij}^{sr} + \sum_{k=1}^{K} y_{ik}^{sr} = t_i^{sr} \qquad (6)$$

Insert equation (4) into equation (1),

$$\sum_{r=1}^{G}\sum_{i=1}^{N}(1+mg_i^{rs})z_{ij}^{rs}+v_k^s=x_j^s \tag{7}$$

Insert equations (5) and (6) into equation (2),

$$\sum_{s=1}^{G}\ \sum_{j=1}^{N}(1+mg_i^{sr})z_{ij}^{sr}+\sum_{s=1}^{G}\ \sum_{k=1}^{K}(1+mg_i^{sr})y_{ik}^{sr}=x_i^r \tag{8}$$

The economic meanings of these two equations are straightforward. Equation (7) defines the value of gross output for commodity group *j* in production country *s* as the sum of the values from all of its (domestic plus imported) intermediate and primary factor inputs. Equation (8) states that total gross output of commodity group *i* in destination country *r* is equal to the sum of all deliveries to intermediate and final users from all countries (including itself) in the world. The delivery of intermediate and final goods in this ICIO account should be consistent with international trade statistics each year, which is the bilateral trade flow definition equation (6).

Because this extended accounting framework (Equations (6) to (8)) is mathematically equivalent to equations (1) to (3), this ICIO account is fundamentally consistent with the MCIO account defined earlier, this is the theoretical foundation that a MCIO table can be used as an important intermediate step towards estimating a full-fledged ICIO account. However, because the MCIO account has a much smaller dimension thus significant additional information will be required to empirically separate inter-country trade flows into end use categories that delivery to sector and final users.[5]

An ICIO account provides the best available and consistent information that allow us to model the value-added generation process among related countries at industry average level. It traces inter-country transaction in intermediate inputs and final use separately, matches bilateral trade flow in major end use categories to input-output relations therefore includes more detailed source/destination, supply/use information than a MCIO table, which is the core of the GTAP database. In short, an ICIO table extended from the GTAP database will not only provide the origin and destination of international trade flows in its covered industries, but also specifies every intermediate and/or final use for all such flows. For example, from such an extended table

[5] The aggregate model only has $N(NG+G^2+5G)$ variables and $N(3G+N+5)$ constraints, while the full detailed model has $(N^2G + NHG)(G+1)$ variables and $N(G^2+NG+N+5)$ constraints. It is a much smaller model, having $NG^2(N-1) + NG(HG-5)$ less variables and $NG(G+N-3)$ less constraints.

we will not only know how many electronics produced in China was shipped into the United States, but also can distinguish how many of them used as intermediate inputs in which particular U.S. industry and how many of them used for U.S. private household consumption or capital formation.

2.2 MCIO account in the GTAP Database

About 40 arrays in each version of GTAP database are used to store related data set for each release. Flowing 13 arrays are needed to construct an MCIO account valued in market price:

TVOM(TRAD_COMM,REG) sales of domestic product, at market prices;
VFM(ENDW_COMM,PROD_COMM,REG) primary factor purchases, by firms, at market prices;
EVFA(ENDW_COMM,PROD_COMM,REG) primary factor purchases, at agents' prices;
VDFM(TRAD_COMM,PROD_COMM,REG) domestic purchases by firms at market prices;
VDFA(TRAD_COMM,PROD_COMM,REG domestic purchases, by firms, at agents' prices ;
VDGM(TRAD_COMM,REG) domestic purchases by government at market prices;
VDPM(TRAD_COMM,REG) domestic purchases by households at market prices;
VIFM(TRAD_COMM,PROD_COMM,REG) import purchases, by firms, at market prices;
VIFA(TRAD_COMM,PROD_COMM,REG) import purchases, by firms, at agents' prices;
VIGM(TRAD_COMM,REG) import purchases, by government, at market prices;
VIPM(TRAD_COMM,REG) import purchases, by households, at market prices ;
VST (MARG_COMM,REG) margin exports;
VXMD (TRAD_COMM,REG,REG) non-margin exports, at market prices.

Equations (1) - (3) that define the MCIO account can be written in GTAP notation as follows:

TVOM(i,r) = sum(j, VDFM(j,i,r)) + sum(j, VIFM(j,i,r)) + sum(f,EVFM(f,i,r))
+ sum(j, VDFA(k,i,r)-VDFM(j,i,r)) + sum(k, VIFA(j,i,r)-VIFM(j,i,r))
+ sum(f, EVFA(f,i,r) - VFM (f,i,r)) + (TVOM(i,r) - TVOA(i,r)); (G1)

TVOM(i,r) = sum(k, VDFM(i,j,r)) + VDPM(i,r) + VDGM(i,r) + VDFM(i,"cgd",r)
+ SUM(s, VXMD(i,r,s)) + VST(i,r); (G2)

Sum(s, VXMD(i,s,r)) = sum(j, VIFM(i,j,r)) + VIPM(i,r) + VIGM(i,r)

+ VIFM(i,"cgd",r) (G3)

Equation (G1) specifies the column (cost of production) balance of the MCIO account. Where sum(j, VDFA(j,i,r)-VDFM(j,i,r)), sum(k, VIFA(j,i,r)-VIFM(j,i,r)), and sum(f, EVFA(f,i,r) - VFM (f,i,r)) are taxes of domestic intermediate inputs, imported intermediate inputs and production factor inputs, respectively; (TVOM(i,r) - TVOA(i,r)) is tax on production and TVOA(i,r) = SUM(j,VDFA(j,i,r)+VIFA(j,i,r)) + SUM(f, EVFA(f,i,r)), all these taxes plus the payment to production factors, sum(f,EVFM(f,i,r)), constitute total value-added in country r.

Equation (G2) specifies the row (supply and demand) balance of the MCIO account. Where sum(j, VDFM(i,j,r)), VDPM(i,r), VDGM(i,r) and VDFM(i,"cgd",r) give demand of domestic products for intermediate inputs, private and public consumption as well as investment respectively; The remaining two terms are margin and non-margin commodity exports, the external demand for goods and services produced in country r.

Equation (G3) specifies the import supply and demand balance condition in the MCIO account. It is the same as equation (3), and splits import demand for final goods into three end use categories, i.e. K=3 as private, government and investment demand respectively.

3. A Mathematical Programming Model to Separate Gross Bilateral Trade into Trade Flows by End Use Categories

3.1 Estimating ICIO table from existing MCIO table - the optimization model

Assume an MCIO table exists. This implies that all variables on the right side of equations (4) to (8) specified in section 2.2 ($x_i^r, y_i^{mr}, z_{ij}^{mr}, t_i^{sr}$) and value-added by sector in each country (v_i^r) are known and can be treated as parameters. Suppose international transportation margins and tariff information are also available. Then to estimate an ICIO table containing $G\times G$ different intermediate trade flow matrix (Z^{rs}, $r,s\in G$), and $KG\times G$ different final goods flow matrix matrix (Y^{rs}, $r,s\in G$) from the existing MCIO table can be formulated as an optimization model and specify a cross-entropy (Harrigan & Buchanan, 1984, Golan et al., 1994) or a quadratic objective penalty function subject to equations (4) to (8) as constraints[6].

[6] The quadratic function has a numerical advantage in implementing the model. It is easier to solve than the entropy function in very large models because they can use software specifically designed for quadratic programming. As showed by Canning and Wang (2005), the quadratic function is equivalent to the entropy function in the neighborhood

For example, the quadratic objective penalty function for such an optimization model can be specified as follows:

$$\operatorname{Min} S = \frac{1}{2}\left\{\sum_{s=1}^{G}\sum_{r=1}^{G}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{(z_{ij}^{sr}-\bar{z}_{ij}^{sr})^2}{wz_{ij}^{sr}}+\sum_{s=1}^{G}\sum_{r=1}^{G}\sum_{i=1}^{N}\sum_{k=1}^{K}\frac{(y_{ik}^{sr}-\bar{y}_{ik}^{sr})^2}{wy_{ik}^{sr}}\right\} \tag{9}$$

A solution to this quadratic programming model provides a complete set of estimates for a full-fledged ICIO table. It is similar in many aspects with the interregional accounting framework proposed by Batten (1982) two decades ago, who used an entropy formulation based on an uninformed data pooling approach for initial estimates where all weights are equal to one.

In theory, one can construct either informed (e.g., survey based) or uninformed (e.g., data pooling) initial estimates for each endogenous element of the ICIO table—$\bar{z}_{ij}^{sr}$ and $\bar{y}_{ik}^{sr}$, along with reliability measures to weight each initial estimate —wz_{ij}^{sr} and wy_{ik}^{sr}. The “uninformed” initial estimates are derived in the absence of information about variations in row or column structures in the targeted ICIO account. In such cases, one typically adopts proportional allocation methods and assigns weights in these same proportions. The “informed” initial estimates requires using the greatest amount of primary information from multiple sources that collectively provide consistent descriptions of all row or column structures in the targeted ICIO account. Ideally, the primary information sources include statistical measures of reliability that can be used to weight these initial estimates. Therefore, the key steps in implementing this optimization model with real data properly are construct these initial estimates with available information from different sources and select a full set of reliability weights in the objective function in order to obtain a meaningful solution from the model. We will discuss these implementation issues in following sub-sections.

3.2 Construct initial estimates based UN BEC classification and detailed bilateral trade statistics

To estimate detailed inter-industry and inter-country intermediate and final transaction flows in an ICIO table, we need additional information beyond a MCIO table in the GTAP database to (i) distinguish intermediate and final use of imports from different sources in each sector, and (ii) allocate intermediate goods from a particular country source to each sector it is

of initial estimates, under a properly selected weighing scheme.

used within all destination countries. We address the first information issue based on UN Broad Economic Categories (BEC) and detailed trade statistics. However, no additional information is available to properly allocate intermediates of a particular sector from a specific source country to its use industries at the destination economy. Thus, sector *j*'s imported intermediate inputs of a particular product are initially allocated to each source country by assuming they are consistent with the aggregate source structure of that particular product.[7]

Although the GTAP database provides bilateral trade flows, it does not distinguish whether goods are used as intermediates or final goods. Our initial allocation of bilateral trade flows into intermediate and final uses is based on the UN BEC applied to detailed trade statistics at the 6-digit HS level from COMTRADE based on concordance used in WIOD project [8]. This differs from the approaches in Johnson and Noguera (2010) and Daudin, Rifflart, and Schweisguth (2010), which also transform the MCIO table in the GTAP database into an ICIO table. However, they do not use detailed trade data to identify intermediate goods and final goods trade in each bilateral flow. Instead, they apply a proportionality method directly to the trade data in GTAP database; i.e., they assume that the proportion of intermediate to final goods is the same for domestic supply and imported products.

Suppose we could obtain estimates for share of intermediate, consumption and capital goods transactions in each bilateral trade flows based on UN BEC classification and detailed trade statistics as ish (i,s,r) and fsh(i,s,r) respectively, then we can initialize these endogenous variables z_{ij}^{sr} and y_{ik}^{sr} in the model as follows:

1. Compute the share of intermediate goods distributed to its use industries based on data available in the MCIO table and distribute imported intermediate goods by proportion for $s \neq r$

[7] For example, if 20% of U.S. imported intermediate steel comes from China, then we assume that each U.S. industry obtains 20% of its imported steel from China. Such an assumption ignores the heterogeneity of imported steel in different sectors. It is possible that 50% of the imported steel used by the U.S. construction industry may come from China, while only 5% of the imported steel used by auto makers may be Chinese.

[8] We thank Dr. Robert Stehrer at WIIW kindly provides the concordance. Both the zero/one and a weighting scheme are be used in WIOD concordance to allocate bilateral trade flows at 6-digit HS level into the three major UN BEC end use categories, there are 703 (among 5718 in total) 6 digit HS code are identified as dual used products that were split into two or more end use categories in the WIOD concordance. This is better than the zero/one classification from UNSD we used in an earlier version of the paper. Shares based on country-specific information could be applied as weights to further improve the allocation. These are areas for future research.

$$\bar{z}_{ij}^{sr} = \frac{z_{ij}^{mr}}{\sum_{j=1}^{N} z_{ij}^{mr}} ish_i^{sr} f_i^{sr} \quad (10)^{9}$$

2. Compute the share of final goods distributed to its final users based on data available in the MCIO table and distribute imported final goods by proportion for s ≠r

$$\bar{y}_{ik}^{sr} = \frac{y_{ik}^{mr}}{\sum_{K=1}^{K} y_{ik}^{mr}} fsh_i^{sr} f_i^{sr} \quad (11)$$

3. Keep domestic intermediate inputs and final goods use as what in the MCIO table

$$\bar{z}_{ij}^{rr} = z_{ij}^{rr} \qquad \bar{y}_{ik}^{rr} = y_{ik}^{rr}$$

4. Compute margins between the source country's exports and destination country's imports, this could include exports tax or subsidies in the source countries and import duties in the destination countries as well as international transportation cost for each bilateral route.

The use of end-use categories to distinguish imports by their final users is becoming more widespread in the literature and avoids some noted deficiencies of the proportionality method.[10] Feenstra and Jensen (2009) use a similar approach to separate final goods from intermediate inputs in U.S. imports in their recent re-estimation of the Feenstra-Hanson measure of material off shoring. Dean, Fung, and Wang (2011) show that the proportionality assumption underestimates the share of imported goods used as intermediate inputs in China's processing trade. Nordas (2005) states that the large industrial countries have a higher share of intermediates in their exports than in their imports, while the opposite is true for large developing countries. These results imply that the intermediate content of imports differs systematically from the intermediate content in domestic supply.

The less distorted intermediate trade share estimates from end use classification provides a better initial row sum for each block matrix of Z_{sr} in the ICIO flow matrix *Z*, thus giving a better row total control of the most important parameters (the IO coefficients) in an ICIO model.

---

[9] $\bar{z}m_{ij}^{sr}$, $\bar{z}x_{ij}^{sr}$ and $\bar{y}m_{ic}^{sr}$, $\bar{y}x_{ic}^{sr}$, the intermediate and final goods trade flows computed based on the share reported by importers and exporters are used as up and low bound to constraint for model solutions.

[10] The literature notes that the UN BEC classification has shortcomings of its own however, particularly its inability to properly identify dual-use products such as fuels, automobiles, and some food and agricultural products.

However, it still does not properly allocate particular intermediate goods imported from a specific source country to each using industry (the ICIO flows in each cell of a particular row in each block matrix Z_{sr} still have to be estimated by proportionality assumption). This allocation is especially important to precisely estimate value-added by sources for a particular industry, although it is less critical for the country aggregates because total imports of intermediates from a particular source country are fixed by observed data, so misallocations will likely cancel out.

3.3 Additional issues of model initialization in the GTAP database

The international transportation cost often vary for intermediate, capital and consumption goods in each bilateral route and different country may impose different tariff rate for intermediate and final goods. However, the international transportation margin and tariff data in current GTAP database cannot make such distinctions, we have to assume international transportation margins are the same for intermediate and final goods "i" in the same bilateral trade route and split VTWR (trs,i,s,r), the margins commodity array in GTAP database, according to the proportion of each end use category in the bilateral trade flows, and treat them as intermediate inputs from the international transportation margin supply industries (air, water and other transportation sectors) at the source country to the use industries in the destination countries. We also have to assume exports subsidies/import tariffs have the same rate between intermediate and final goods in the same source/destination countries and treat them as part the value-added created by the source/destination countries.

Among the 129 country/region in version 8 GTAP database, 20 of them are composite regions. The new ICIO database we constructed from the GTAP database covers 63 countries/regions, 17 of them constituted by more than one country. China and Mexico have normal and processing trade regions. (see Appendix A for country aggregation of the new database from V8 GTAP classifications)[11]. The bilateral trade flows within these composite regions are removed and treated as the composite regions' domestic supply and demand.

The details of these special treatments can be found in the GAMS code in Appendix.

3.4 Selection of reliability indexes in the objective function

[11] The new database has similar sector classification, except most primary sectors. It aggregates the 12 primary agricultural sector into two sectors, oil and gas into one sector, and the 8 food processing sector into 3 sectors.

As pointed by Wang et al (2010), one of the most desirable analytical and empirical properties of this class of data reconciliation models such as the one we specified by equations (4) – (9) is it uses reliability weights in the objective function to control how much an initial estimate may be adjusted. If the selected weights properly reflect the relative reliability of the associated initial estimates, the model will adjust those relatively unreliably reported data more than those relatively reliably reported data in the reconciliation process. In other words, initial estimates with a higher reliability will be adjusted less than initial estimates with a lower reliability, thus the best available information can always be used to insure that statistics reported by reliable trade routes or reporters are not perturbed by the reconciliation process as much as statistics reported by unreliable trade routes or reporters. From statistical point of view, the best way to systematically assign reliability weights in the objective function is to obtain estimates of the variance-covariance matrix of the initial estimates. Then the inverted variance-covariance matrix can be justified as the best index of the reliability of initial estimates. The larger the variance, the smaller the associated term $\frac{(z_{ij}^{sr} - \bar{z}_{ij}^{sr})^2}{wz_{ij}^{sr}}$ or $\frac{(y_{ik}^{sr} - \bar{y}_{ik}^{sr})^2}{wy_{ik}^{sr}}$ contributes to the objective function, and hence the lesser the penalty for the associated variables to move away from their initial value (only the relative, not the absolute size of the variance affects the solution). A small variance of the initial estimates indicates, other things being equal, that it is more reliably reported data and thus should not be required to change by as much. In contrast, a large variance of the initiate estimates indicates unreliably reported data that may be adjusted considerably. However, the lack of consistent historical data often makes the estimation of the variance-covariance matrix associated with the initial estimates very difficult to implement. For example, the common practice in SAM balancing exercises is assign differing degrees of subjective reliabilities to the initial entries of the matrix follow the method proposed by Stone (1984),[12] almost no attempt to date has been made to statistically estimate data reliability such as error variance of the initial estimates from historical data, except Weale (1989), who developed a statistical method that uses time series information on accounting discrepancies to infer data reliability in a system of national accounts. Theoretically speaking, a similar statistical method can be applied to the historically reported discrepancies of bilateral trade data to derive those variances associated with international trade statistics. In practice,

[12] Stone proposed to estimate the variance of x^0_{ij} as $var(x^0_{ij}) = (\theta_{ij}x^0_{ij})^2$, where θ_{ij} is a subjective determined reliability rating, expressing the percentage ratio of the standard error to the initial estimates of x^0_{ij}.

however, the historical data and knowledge of the changes in related country's trade reporting system are too demanding and make such a statistic method less attractable in large empirical applications. Therefore, here we suggest a practical alternative approach to estimate the reliability weights, which is constructed by reporter relative reliability indexes for both exporters and importers.

3.4.1 Reporter reliability indexes

Trade data reported by each country and its partners are often used in the international economic literature to check the quality of trade statistics. An approximate match of mirror statistics suggests that trade data reported via that route are reliable. However, such weights treat the reported trade statistics from both reporters equally and do not distinguish which reporter is more reliable. In the case there is a very unreliable reporter in the pair, it may adjust the reliable data reported by the partner too much thus loss original accurate information from the reliable partner. This is undesirable. To correct this problem, a reporter's relative reliability index needs to be developed. Such an index should be able to deal with three critical issues.

The first issue is related to the difference of reporting countries in their ability to report bilateral commodity trade by end use categories. Variability in reporting quality across countries is highly relevant information for the problem we try to solve in our proposed data reconcilation approach. As discussed earlier, the adjustment process hinges heavily on the relative reliability of the each reporting countries. An indicator of reporter reliability is a measure of how consistency a country reports its trade in each end use categories relative to all its trading partners. However, judging a country's trade data based on a single bilateral flow alone is a poor reference, because a partner can misrepresent its trade thereby potentially discrediting a reliable reporter. Therefore, a good reporter reliability measure should take all reporting countries in the world into account in assessing a country's reporting reliability.

The second issue is what exactly should be captured by the reliability measure. The size of discrepancies could be incorporated into a measure of reliability. However, placing emphasis on the magnitude of discrepancies only may over-penalize the reliability of a legitimate reporter. A poor reporter that makes an error for a given trade flow usually makes a similar error with other partners. For example a reporter that has mistaken the identity of one of its partners has implicitly made a mistake for others. It brings a systemic bias for that reporter. This type of problem should be detected and reflected in the reporter reliability measure without penalizing the reliable reporter.

The third issue is the capability of the measure to reflect both end-use-categories by sector- and country-specific reliability information for each country as an exporter and as an importer. Countries typically have commodity by end use category specific strength and weaknesses. For example one exporting country may have an excellent reporting record on steel used as intermediate goods but at the same time is highly inconsistent in its reporting practice for organic chemical in final goods trade.

All three issues discussed above are effectively dealt with in the reliability index developed by Gehlhar (1996) where reporter reliability indices were used to make a discreet choice whether to disregard or accept reported trade flows. The index is calculated as the share of accurately reported transactions of a reporter's total trade for a particular end use category in a sector using a threshold level. It assesses reporter reliability from a complete set of global reporting partners, captures the reporter's ability to accurately report without interferences from gross discrepancies in reporting, and contains exporter and importer-sector and end use category specific reliability information. Specifically, the importer-sector and end use category specific and exporter-sector and end use category specific reliability indexes in the objective function (equation (9)) are defined as:

$$RIM_{ic}^{r} = \frac{MA_{ic}^{r}}{\sum_{s} M_{ic}^{sr}} \qquad where \qquad MA_{ic}^{r} = \sum_{s \in AL_{ic}^{sr} \le 0.20} M_{ic}^{sr} \qquad AL_{ic}^{sr} = \frac{\left| M_{ic}^{rs} - E_{ic}^{sr} \right|}{M_{ic}^{rs}} \tag{12}$$

$$RIX_{ic}^{s} = \frac{XA_{ic}^{s}}{\sum_{r} E_{ic}^{sr}} \qquad where \qquad XA_{ic}^{s} = \sum_{s \in AL_{ic}^{sr} \le 0.20} E_{ic}^{sr} \qquad AL_{ic}^{sr} = \frac{\left| M_{ic}^{rs} - E_{ic}^{sr} \right|}{M_{ic}^{rs}} \tag{13}$$

Under such defined reporter reliability indexes, the size of the discrepancies becomes immaterial because inaccurate transactions are treated the same regardless of the magnitude of the inaccuracy. The indexes have the flexibility of being implemented at the detailed 6-digit HS level and can be aggregated to any sector level. We computed such reporter reliability measures for each GTAP country/region for the 3 end use categories at the GTAP sector level. Major data are from UN COMTRADE with supplements from country sources.

3.4.2 Reliability weights used in objective function

After obtaining RIM and RIX, there is an additional issue need to be solved before we can empirically compute the reliability weights in the objective function (equation (9)) of the data reconciliation model. There is only one number for trade flow in each route at the sector level in the GTAP database, which is a combination of both reporter and partner reported trade statistics based on reporter's reliability. Therefore, the proportion of such composition for each trade routine at GTAP sector level are used as weights to computer a weighted average of RIM and RIX as the final reporter reliability index and the weights in the objective function are assigned by multiplying one minus these weighted average reporter indexes with their corresponding initial values for each endogenous variable in the model. The complete set of weights in equation (9) is defined as follows:

$$wz_{ij}^{sr} = (1 - RIM_{ii}^{r})\bar{z}m_{ij}^{sr} + (1 - RIX_{ii}^{r})\bar{z}x_{ij}^{sr} \tag{14}$$

$$wy_{ic}^{sr} = (1 - RIM_{ii}^{r})\bar{y}m_{ic}^{sr} + (1 - RIX_{ii}^{r})\bar{y}x_{ic}^{sr} \tag{15}$$

Where $\bar{z}m_{ij}^{sr}$, $\bar{z}x_{ij}^{sr}$ and $\bar{y}m_{ic}^{sr}$, $\bar{y}x_{ic}^{sr}$ are the intermediate and final goods trade flows computed based on the share reported by importers and exporters respectively (shares multiple t_{ij}^{sr}, the bilateral trade flows in GTAP database). With such a weighting scheme, we achieve our goal to encourage the model to change those unreliable initial data more than those reliable ones in the reconciliation process. It means the reconciled solution from the model not only adjust less to the reliable routes than the unreliable ones, but also adjust more to the relative unreliable reporter than the relative reliable reporter in each trade route, although in a rough manner.

4. Include Processing Trade Information from Major Developing Countries

The World Trade Organization has identified more than 130 countries that use some form processing exports (WTO and IDE JETRO, 2011) and reports that about 20% of developing country exports come from Export Processing Zones (EPZs). Such processing regimes provide incentives to use imported intermediate inputs, provided that the resulting final goods are entirely exported. Processing trade can thus dramatically increase the imported content of exports relative to domestic use. Failure to account for processing trade can dramatically overstate the domestic content of exports (Koopman, Wang, and Wei, 2008).

To reflect the reality and importance of processing trade and Export Processing Zones (EPZs) in emerging economies and their role in global value-added trade and production network, we extend Koopman, Wang, and Wei (2008, 2012) to a multi-country global setting that separates standard input-output tables of a subset countries in our database into normal and processing trade accounts. In what follows we first specify a mathematical programming model that is able to split a standard ICIO tables into normal and processing trade accounts for a subset countries, then briefly discuss the data sources and major implementation issues.

4.1 Mathematical programming model to separate processing trade account for a subset of developing countries

The objective of this second stage optimization model is to split the economies with processing trade information in the ICIO table estimated from the first stage optimization model into separate normal and processing accounts, each with their own input-output structure. i.e further split z_{ij}^{sr}, y_{ik}^{sr} and v_j^s in the ICIO account specified in equations (6) to (8) into zn_{ij}^{sr} and zp_{ij}^{sr}, yn_{ik}^{sr} and yp_{ik}^{sr}, vn_j^s and vp_j^s for a subset countries respectively. The additional letter "n" and "p" in the related variables represent normal and processing economy respectively. The basic idea is to use information from the ICIO table to determine sector-level bilateral imports/exports, and additional information of processing exports/imports from trade statistics in a subset countries to determine the relative proportion of processing and normal trade flows within each sector, thus use up all available data to split the subset economies into processing and non-processing blocks, each with its own IO structure. The first step (using trade data from the ICIO table to determine sector-level total imports/exports) helps to ensure that the balance conditions in the ICIO account are always satisfied, and that the separate processing and non-processing accounts in the subset economies are consistent with the ICIO table. The second step (using data from trade statistics to determine the relative proportion of processing and normal flows within each sector level bilateral trade route) helps to ensure that the estimated new ICIO table with processing trade account for subset countries is consistent with the trade structures implied by official trade statistics obtained from these economies, i.e $t_i^{sr} = tn_i^{sr} + tp_i^{sr}$ always hold.

Assume there are P countries in the G country world engage in processing trade with G>P. Assume all output from the P economies with processing trade is exported to the international market, then output of the normal economies in each of the P country can be obtained by subtracting processing exports to all destination from the source country's sector level total output. The ICIO table with processing trade account can be specified as follows:

Column balance of these economies with processing trade account

$$\sum_{r=1}^{G}\sum_{i=1}^{N}(1+mg_i^{rs})zn_{ij}^{rs}+vn_k^s = x_j^s - \sum_{s=1}^{G}\sum_{i=1}^{N}zp_{ij}^{sr} - \sum_{s=1}^{G}\sum_{k=1}^{K}yp_{ik}^{sr} \tag{16}$$

$$\sum_{r=1}^{G}\sum_{i=1}^{N}(1+mg_i^{rs})zp_{ij}^{rs}+vp_k^s = \sum_{s=1}^{G}\sum_{i=1}^{N}zp_{ij}^{sr} + \sum_{s=1}^{G}\sum_{k=1}^{K}yp_{ik}^{sr} \tag{17}$$

Column balance of these economies with processing trade account

$$\sum_{s=1}^{G}\sum_{j=1}^{N}(1+mg_i^{sr})zn_{ij}^{sr} + \sum_{s=1}^{G}\sum_{k=1}^{K}(1+mg_i^{sr})yn_{ik}^{sr} = x_i^r - \sum_{s=1}^{G}\sum_{i=1}^{N}zp_{ij}^{sr} - \sum_{s=1}^{G}\sum_{k=1}^{K}yp_{ik}^{sr} \tag{18}$$

Trade flow balance for imports from and exports to all other G-P countries without processing trade account:

$$\sum_{j=1}^{N}zn_{ij}^{sr} + \sum_{k=1}^{K}yn_{ik}^{sr} = tn_i^{sr} \tag{19}$$

$$\sum_{j=1}^{N}zp_{ij}^{sr} + \sum_{k=1}^{K}yp_{ik}^{sr} = tp_i^{sr} \tag{20}$$

Trade flow among all the P countries with processing trade account:

$$\sum_{j=1}^{N}znn_{ij}^{sr} + \sum_{k=1}^{K}ynn_{ik}^{sr} + \sum_{j=1}^{N}zpn_{ij}^{sr} + \sum_{k=1}^{K}ypn_{ik}^{sr} = tn_i^{sr} \tag{21}$$

$$\sum_{j=1}^{N}znp_{ij}^{sr} + \sum_{k=1}^{K}ynp_{ik}^{sr} + \sum_{j=1}^{N}zpp_{ij}^{sr} + \sum_{k=1}^{K}ypp_{ik}^{sr} = tp_i^{sr} \tag{22}$$

Adding up conditions

$$vn_j^s + vp_j^s = v_j^s \tag{23}$$

$$zn_{ij}^{sr} + zp_{ij}^{sr} = z_{ij}^{sr} \tag{24}$$

$$yn_{ik}^{sr} + yp_{ik}^{sr} = y_{ik}^{sr} \tag{25}$$

$$tn_{i}^{sr} + tp_{i}^{sr} = t_{i}^{sr} \tag{26}$$

The basic balance condition of ICIO table, Equations (6)-(8) continue to hold for the G-P economies without processing trade account.

The second stage optimization model can be constructed with following quadratic penalty function as objective function and equations (6) to (8) as well as equations (16) to (26) as constraints.

$$\mathrm{Min}\, S = \frac{1}{2}\ \{\sum_{s=1}^{G}\sum_{r=1}^{G}\ \sum_{i=1}^{N}\sum_{j=1}^{N}\left[\frac{(\ zn_{ij}^{sr} - \overline{zn}_{ij}^{sr}\ )^2}{wzn_{ij}^{sr}} + \frac{(\ zp_{ij}^{sr} - \overline{zp}_{ij}^{sr}\ )^2}{wzp_{ij}^{sr}}\right]$$
$$+\sum_{s=1}^{G}\sum_{r=1}^{G}\sum_{i=1}^{N}\sum_{k=1}^{K}\left[\frac{(yn_{ik}^{sr} - \overline{yn}_{ik}^{sr}\ )^2}{wyn_{ik}^{sr}} + \frac{(yn_{ik}^{sr} - \overline{yn}_{ik}^{sr}\ )^2}{wyn_{ik}^{sr}}\right] + \sum_{s=1}^{G}\sum_{j=1}^{N}\left[\frac{(vn_{j}^{s} - \overline{vn}_{j}^{s})^2}{wvn_{j}^{s}} + \frac{(vp_{j}^{s} - \overline{vp}_{j}^{s})^2}{wvp_{j}^{s}}\right]\} \tag{27}$$

4.2 Data source and major implementation issues

Due to data limitation, only two countries, China and Mexico, are selected into the subset economies to empirically implement the model described in last subsection. We use an expanded Chinese IO table with separate accounts for processing exports and a 2003 Mexican IO table with separate domestic and Maquiladora accounts,[13]to initialize related variables in the model.

China and Mexico are the two largest users of export processing regimes in the developing world, and together account for about 85% of worldwide processing exports. During 2000-2008, China alone accounted for about 67% of all reported processing exports in the world while Mexico represents another 18% (Maurer and Degain, 2010).[14] Therefore, using processing trade information from these two countries that involve major processing trade activities in the world, the constructed database should get the large picture right for the world production and trade patterns. When similar information from other developing country becomes accessible, the model can be extended to cover more developing countries easily.

[13]The Mexican table is from the Mexican statistical agency Instituto Nacional de Estadística, Geografía e Informática (INEGI).

[14] Similarly, based on IMF BOP statistics provided by Andreas Maurer, we estimate that China and Mexico together accounted for about 80% of goods for processing in the world in 2005 and 2007.

5 Mean absolute percentage adjustment for major variables in the GTAP database

Among the 13 data array used to construct MCIO account from GTAP database, bilateral trade flows (VXMD), Total gross output (TVOM), primary factor demand (VFM) and supply of international transportation margin (VST) are fixed as constant in the optimization model[15], but allow domestic and imported purchase goods and services to adjust to fit the balance condition in the ICIO table in construction. The data reconciliation procedure produces a different set of estimates for those domestic and imported purchases than what gave in the GTAP database, it is desirable to know how much each set of estimates differs from the original GTAP data. However, it is difficult to use a single measure to compare the original and adjusted data, since there are so many dimensions in the data. It is meaningful to use several measures to gain more insight on the model performance. Generally speaking, it is the proportionate deviation and not the absolute deviation that matters; therefore, we compute the "Mean Absolute Percentage Adjustment" with respect to the original GTAP data for different country and sector aggregations. Consider the following aggregate index measure for country and commodity group total adjustment for both intermediate and final demand.

Domestic intermediate demand:

$$MAPADI^{r} = \frac{100 \bullet \sum_{j=1}^{N} \sum_{i=1}^{N} | z_{ij}^{rr} - \mathrm{VDFM}_{ijr} |}{\sum_{i=1}^{N} \sum_{j=1}^{N} VDFM_{ijr}} \tag{28}$$

$$MAPADI^{i} = \frac{100 \bullet \sum_{r=1}^{G} \sum_{j=1}^{N} | z_{ij}^{rr} - \mathrm{VDFM}_{ijr} |}{\sum_{r=1}^{G} \sum_{j=1}^{N} VDFM_{ijr}} \tag{29}$$

Imported intermediate demand

$$MAPAII^{r} = \frac{100 \bullet \sum_{j=1}^{N} \sum_{i=1}^{N} | \sum_{s \neq r}^{G} z_{ij}^{sr} - \mathrm{VIFM}_{ijr} |}{\sum_{i=1}^{N} \sum_{j=1}^{N} VIFM_{ijr}} \tag{30}$$

[15] Another three arrays EVFA, VDFA and VIFA are used to compute taxes.

$$MAPAII^{i} = \frac{100 \bullet \sum_{r=1}^{G} \sum_{j=1}^{N} | \sum_{s \neq r}^{G} z_{ij}^{sr} - \mathrm{VIFM}_{ijr} |}{\sum_{r=1}^{G} \sum_{j=1}^{N} VIFM_{ijr}} \qquad (31)$$

Domestic final demand

$$MAPADF^{r} = \frac{100 \bullet \sum_{i=1}^{N} | \sum_{k} y_{ik}^{rr} - VDPM_{ir} - VDGM_{ir} - \mathrm{VDFM}_{\mathrm{i,'cgd',r}} |}{\sum_{i=1}^{N} (VDPM_{ir} + VDGM_{ir} + \mathrm{VDFM}_{\mathrm{i,'cgd',r}})} \qquad (32)$$

$$MAPADF^{i} = \frac{100 \bullet \sum_{r=1}^{G} | \sum_{k} y_{ik}^{rr} - VDPM_{ir} - VDGM_{ir} - \mathrm{VDFM}_{\mathrm{i,'cgd',r}} |}{\sum_{r=1}^{G} (VDPM_{ir} + VDGM_{ir} + \mathrm{VDFM}_{\mathrm{i,'cgd',r}})} \qquad (33)$$

Imported final demand

$$MAPAIF^{r} = \frac{100 \bullet \sum_{i=1}^{N} | \sum_{r \neq s}^{G} \sum_{k} y_{ik}^{sr} - VIPM_{ir} - VIGM_{ir} - \mathrm{VIFM}_{\mathrm{i,'cgd',r}} |}{\sum_{i=1}^{N} (VIPM_{ir} + VIGM_{ir} + \mathrm{VIFM}_{\mathrm{i,'cgd',r}})} \qquad (34)$$

$$MAPAIF^{i} = \frac{100 \bullet \sum_{r=1}^{G} | \sum_{r \neq s}^{G} \sum_{k} y_{ik}^{sr} - VIPM_{ir} - VIGM_{ir} - \mathrm{VIFM}_{\mathrm{i,'cgd',r}} |}{\sum_{r=1}^{G} (VIPM_{ir} + VIGM_{ir} + \mathrm{VIFM}_{\mathrm{i,'cgd',r}})} \qquad (35)$$

The numerical results for the 8 aggregate indexes defined above are reported in tables 1 and 2 for the year 2007, and tables 3 and 4 for the year of 2004 respectively.

We focus on results for country total adjustments to illustrate some key characteristics of the adjustment process. Each country's reliability as an exporter and importer is a key factor that governs the magnitude of adjustment of its exports and imports. Generally speaking, there are three noticeable features of the adjustment made in the data reconciliation process. First, the adjustment made for developed countries is smaller than the adjustment made for developing countries in average, reflecting the facts that the data quality is better in developed countries than that in most developed countries in the GTAP database. Second, the adjustment made for domestic demand is smaller than the adjustment made for imported demand , reflecting that the information on how and where imported commodity were sourced and used are generally poor than information on how and where domestic products were used in the GTAP database. Finally the adjustment for domestic intermediate inputs purchase is generally larger than the adjustment made for domestic final demand, but it is in the opposite for the adjustment in imported demand, indicating the share of imported final good usage in the extended database is quite different from the original GTAP database, whether this caused by the inaccuracy final demand information in the GTAP database or due to our BEC to HS concordance needs further investigation. Looking into the adjustment at sector level, seems these sectors have large portion of their products could be used as both intermediate and final goods often associated with large adjustments.

6. Concluding Remarks

This paper describes how a Global Inter-Country Input-Output Table with processing trade account can be constructed from GTAP database. It first provides a theoretical foundation that explains how the MRIO table embodied in GTAP database could be consistent with an ICIO table and what additional information is needed for the transformation. Using a quadratic programing model with reliability weights in its objective function, we constructed two preliminary ICIO tables for the year 2004 and 2007 from version 8 GTAP database, covering 63 countries and 41 sectors. Additional work is needed to further improve the HS to UN BEC concordance and extend it to services trade. Better methods also need to be developed to properly distribute imports to domestic users either based on sector specific information, or cross country statistical surveys of the domestic distribution of imports or linked firm level and Customs transaction-level trade data. This will need joint efforts by statistical agencies and academic communities across the world.

References

Brooke, Kendrick, Meeraus, and Raman, 2005, "GAMS -- User's Guide" GAMS Development Cooperation, Washington, DC.

Byron, Ray P. 1978. "The Estimation of Large Social Account Matrix," *Journal of Royal Statistical Society*, A, 141 (Part 3), 359-367.

Patrick Canning and Zhi Wang "A Flexible Mathematical Programming Model to Estimate Interregional Input-Output Accounts." *Journal of Regional Sciences* 45(3):539-563, August 2005.

Ferrantino Michael and Zhi Wang, "Accounting for Discrepancies in Bilateral Trade: The Case of China, Hong Kong, and the United States" *China Economic Review*, 19(4): 502-520, October 2008.

Gehlhar, Mark, 1996, "Reconciling Bilateral Trade Data for Use in GTAP," *GTAP Technical Paper no 10*, Purdue University.

Harrigan, J. Frank 1990, "The Reconciliation of Inconsistent Economic Data: the Information Gain," *Economic System Research*, Vol.2, No.1, pp. 17-25

Ploeg, van der F, 1984, "General Least Squares Methods for Balancing Large Systems and Tables of National Accounts," *Review of Public Data Use*, 12, 17-33

Hummels, D., J. Ishii, and K. Yi, 2001, "The Nature and Growth of Vertical Specialization in World Trade," *Journal of International Economics* 54:75–96.

Johnson, Robert, and Guillermo Noguera, 2009, " Accounting for Intermediates: Production Sharing and Trade in Value-added," Mimeo, Princeton University, June.

Koopman, Robert, Zhi Wang and Shang-jin Wei, 2008, "How much Chinese exports is really made in China – Assessing foreign and domestic value-added in gross exports," NBER Working Paper 14109, June.

Robert Koopman, Zhi Wang and Shang-jin Wei "Estimating domestic content in exports when processing trade is pervasive." *Journal of Development Economics 99*(2012):178-189.

Robinson, Sherman, Andrea Cattaneo and Moataz El-Said. "Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods" *Economic System Research*, 13(1), March 2001, p. 47-64.

Sébastien Miroudot, Rainer Lanz and Alexandros Ragoussis, 2010 "TRADE IN INTERMEDIATE GOODS AND SERVICES" OECD Trade Policy Working Paper No. 93, January.

Stone, Richard, David G. Champernowne and James E. Meade, 1942, "The Precision of National Income Estimates," *Review of Economic Studies*, 9(2), 110-125.

Stone, Richard, “The Precision of National Income Estimates,” *The Review of Economic Studies*, Vol. 9, no. 2, (summer) 1942, p. 111-125.

Wang, Zhi, Mark Gehlhar and Shunli Yao, “Reconciling Trade Statistics from China, Hong Kong and Their Major Trading Partners -- A Mathematical Programming Approach,” *GTAP Technical Paper no 27*, Purdue University, 2007.

Zhi Wang, Mark Gehlhar and Shuli Yao, “A Globally Consistent Framework for Reliability-based Trade Statistics Reconciliation in the Presence of an Entrepôt”, *China Economic Review*, 21(1):161-189, March 2010.

Wang, Zhi, William Powers and Shang-jin Wei, 2009, “Value Chains in East Asian Production Networks: *An International Input-Output Model Based Analysis” Working paper, U.S. International Trade commission.*

Weale, Martin R, 1985, "Testing Linear Hypotheses on National Account Data," *Review of Economics and Statistics*, 67, 685-689.

Table 1 Mean Absolute Percentage Adjustment from GTAP Database (V8) by Region, 2007

	Intermediate demand		Final demand			Intermediate demand		Final demand	
Countries	Domestic	Imports	Domestic	Imports	Countries	Domestic	Imports	Domestic	Imports
Australia	10.3	61.8	6.1	96.0	France	8.3	37.8	2.4	92.3
New Zealand	14.1	77.6	3.7	94.6	Germany	9.7	33.6	3.9	89.4
China	30.2	65.6	13.0	90.2	Greece	19.9	43.3	3.8	97.0
Hong Kong	12.3	43.1	5.8	97.3	Hungary	22.4	40.6	7.7	96.4
Japan	5.3	44.7	1.3	95.2	Ireland	19.9	31.1	9.0	91.0
Korea	12.5	44.8	4.2	91.1	Italy	8.9	46.1	2.3	92.1
Taiwan	15.7	33.9	4.1	94.8	Netherlands	13.6	41.6	4.1	97.7
Indonesia	10.4	54.4	3.9	97.9	Poland	13.7	45.4	3.7	93.6
Malaysia	13.3	33.3	6.3	100.6	Portugal	13.9	48.1	4.1	94.5
Philippines	19.1	38.0	4.8	93.3	Rest of EEU	24.2	39.9	13.3	97.3
Singapore	30.0	25.0	8.7	107.9	Spain	11.2	41.9	2.9	91.6
Thailand	13.8	31.0	5.2	95.7	Sweden	15.9	50.4	4.9	91.3
Viet Nam	25.5	36.1	9.2	100.6	United Kingdom	10.1	49.4	2.8	91.3
Rest of East Asia	20.1	58.5	9.4	94.1	EFTA	15.2	37.7	3.7	97.3
India	9.4	48.2	3.1	94.9	Bulgaria	24.3	52.0	7.7	100.3
rest of south Asia	23.3	44.2	15.8	95.4	Romania	18.3	53.8	4.1	95.2
Canada	8.7	34.8	1.7	98.5	Russian Federation	7.4	69.7	3.9	94.7
United States	5.0	39.6	1.3	93.6	Rest of East Europe	19.0	38.6	12.0	97.5
Mexico	64.0	123.4	18.2	96.5	Rest of Former SU	16.9	54.2	11.9	97.7
Argentina	12.8	71.7	2.7	97.8	Turkey	14.2	53.8	3.5	93.2
Brazil	6.4	65.7	1.4	96.0	Saudi Arabia	17.5	47.3	6.4	78.5
Rest of Mercosur	26.3	78.5	8.2	104.0	Rest of Western Asia	19.5	43.4	12.7	95.9
Chile	14.1	56.1	4.2	105.4	Egypt	25.3	64.3	7.1	96.2
Peru	12.4	137.3	8.2	96.1	Morocco	18.0	69.7	9.1	89.0
CAFTA	21.9	46.0	14.5	95.8	Rest of North Africa	19.0	51.4	11.4	98.4
Colombia	11.8	72.3	2.0	99.0	West Africa	32.5	55.2	20.2	94.7
Rest of America	16.1	50.4	5.8	99.0	Central Africa	23.9	49.8	14.3	92.5
Austria	20.3	42.3	4.7	101.8	East Africa	24.5	52.2	18.0	93.0
Belgium and Lux	18.7	20.6	5.4	106.2	South Africa	9.9	56.7	4.0	91.5
Czech and SVK Republic	17.4	37.1	8.1	91.6	Rest of South African Customs Union	38.8	82.7	15.7	108.6
Denmark	23.6	50.0	6.6	96.1	Rest of World	23.3	54.7	9.9	95.3
Finland	19.3	58.1	5.9	96.0	World Total	12.9	45.7	4.0	93.9

Table 2 Mean Absolute Percentage Adjustment from GTAP Database (V8) by Sector, 2007

		Intermediate demand		Final Demand	
GTAP sector		Domestic	Imports	Domestic	Imports
agp	Crop production	11.6	55.6	15.5	88.3
ani	Animal husbandry	20.1	74.7	7.6	97.6
frs	Forestry	39.7	91.4	42.7	98.0
fsh	Fishing	80.4	69.8	13.1	91.8
coa	Coal	74.1	33.8	308.1	118.2
oil	Oil and gas	8.1	8.8	142.9	1050.5
omm	Minerals nec	25.4	43.0	33.3	102.3
met	Meat and Dairy products	23.0	66.2	6.2	89.3
ofd	Food products nec	16.7	61.6	8.0	84.9
b_t	Beverages and tobacco products	24.9	79.1	7.5	86.9
tex	Textiles	18.5	50.5	24.8	68.7
wap	Wearing apparel	48.0	73.3	5.7	59.3
lea	Leather products	88.4	65.7	15.2	103.4
lum	Wood products	35.2	49.7	13.7	78.3
ppp	Paper products publishing	15.2	69.0	12.1	89.0
p_c	Petroleum coal products	19.8	63.5	19.8	99.0
crp	Chemical rubber plastic products	18.0	32.9	21.9	51.1
nmm	Mineral products nec	13.7	73.1	45.0	64.1
i_s	Ferrous metals	14.2	38.9	95.3	101.0
nfm	Metals nec	23.3	29.6	143.9	148.4
fmp	Metal products	14.0	60.7	24.0	59.2
mvh	Motor vehicles and parts	18.8	31.4	9.1	62.4
otn	Transport equipment nec	34.3	38.0	11.3	72.6
ele	Electronic equipment	22.4	26.6	22.9	69.5
ome	Machinery and equipment nec	20.8	35.5	13.8	56.3
omf	Manufactures nec	26.3	77.2	12.4	71.5
ely	Electricity	10.1	94.0	3.4	97.5
gdt	Gas manufacture and distribution	10.9	106.9	17.3	98.0
wtr	Water	8.1	760.0	6.0	99.7
cns	Construction	3.5	134.2	0.4	99.5
trd	Trade	3.9	87.5	1.4	98.2
otp	Other transportation	8.5	99.5	3.8	95.9
wtp	Water transportation	21.4	50.8	37.9	87.7
atp	Air transportation	23.9	66.7	32.6	59.9
cmn	Communication	7.6	115.8	4.7	98.8
ofi	financial services nec	5.7	78.2	3.4	97.7
ins	Insurance	12.7	89.2	5.1	97.4
obs	business services nec	6.5	80.8	4.9	97.1
ros	recreational and other services	10.5	144.6	2.6	98.1
osg	public admin and defence education health	7.0	96.6	0.4	99.3
dwe	Dwellings	5.6	100.0	0.1	100.0
Tot	Total	12.9	45.7	4.0	93.9

Table 3 Mean Absolute Percentage Adjustment from GTAP Database (V8) by Region, 2004

	Intermediate demand		**Final Demand**			**Intermediate demand**		**Final Demand**	
Countries	Domestic	Imports	Domestic	Imports	Countries	Domestic	Imports	Domestic	Imports
Australia	20.3	129.4	8.3	97.8	France	16.1	98.2	4.6	98.3
New Zealand	53.9	218.2	17.0	97.4	Germany	18.4	98.2	5.4	96.9
China	12.7	84.3	5.6	102.0	Greece	59.0	133.0	7.3	101.3
Hong Kong	28.3	107.2	15.0	106.9	Hungary	56.7	101.5	16.6	98.6
Japan	10.4	133.0	2.6	100.1	Ireland	46.4	85.6	21.8	94.2
Korea	30.9	114.2	11.0	103.6	Italy	15.7	118.2	4.8	100.8
Taiwan	45.1	97.4	15.1	109.1	Netherlands	24.8	103.3	6.6	102.8
Indonesia	44.1	137.9	9.5	101.8	Poland	31.8	109.5	7.2	98.2
Malaysia	42.7	82.9	29.4	112.8	Portugal	48.1	134.3	9.1	100.0
Philippines	73.2	101.1	18.4	102.5	Rest of EEU	51.5	123.9	19.0	105.0
Singapore	67.9	72.8	21.1	139.4	Spain	20.8	107.3	3.5	98.8
Thailand	66.0	105.8	28.7	116.4	Sweden	30.8	117.6	7.2	97.5
Viet Nam	79.9	133.9	26.2	106.8	United Kingdom	16.5	104.4	5.6	96.8
Rest of East Asia	59.9	187.8	39.2	97.2	EFTA	28.9	101.6	5.6	101.3
India	25.4	125.7	6.7	103.8	Bulgaria	80.1	193.3	47.0	117.6
rest of south Asia	58.0	157.9	17.6	99.0	Romania	64.6	148.3	14.1	99.9
Canada	19.3	72.5	3.3	102.3	Russian Federation	23.1	157.4	10.0	97.0
United States	11.3	112.3	2.5	98.8	Rest of East Europe	57.5	130.9	12.6	117.8
Mexico	22.3	82.4	3.9	100.0	Rest of Former SU	48.9	157.3	31.1	101.1
Argentina	56.2	187.8	12.7	96.8	Turkey	42.0	148.3	5.7	99.3
Brazil	25.0	143.9	4.4	98.9	Saudi Arabia	38.3	126.4	11.3	85.4
Rest of Mercosur	68.9	240.4	49.5	110.7	Rest of Western Asia	36.1	103.1	12.4	102.1
Chile	61.9	185.0	14.2	106.7	Egypt	70.5	138.1	17.3	96.6
Peru	65.1	441.8	25.8	97.7	Morocco	72.0	202.8	19.9	97.1
CAFTA	67.2	167.6	35.0	96.2	Rest of North Africa	57.7	161.1	18.8	101.8
Colombia	60.2	210.0	14.8	97.7	West Africa	61.5	154.2	27.7	97.8
Rest of America	40.4	139.1	9.0	101.0	Central Africa	86.8	179.0	35.6	99.0
Austria	39.6	102.7	10.1	102.8	East Africa	63.1	160.6	25.7	96.3
Belgium and Lux	33.2	66.6	6.6	112.5	South Africa	34.7	152.0	9.1	99.7
Czech and SVK Republic	37.8	91.3	10.0	100.0	Rest of South African Customs Union	80.8	197.0	53.6	131.5
Denmark	39.0	124.5	13.4	97.7	Rest of World	61.1	175.1	23.9	97.2
Finland	41.2	120.2	9.4	101.8	World Total	20.4	109.4	5.6	99.7

Table 4 Mean Absolute Percentage Adjustment from GTAP Database (V8) by Sector, 2004

		Intermediate demand		Final Demand	
GTAP sector		Domestic	Imports	Domestic	Imports
agp	Crop production	24	133	16	92
ani	Animal husbandry	23	236	21	93
frs	Forestry	71	221	65	88
fsh	Fishing	64	412	46	92
coa	Coal	29	100	132	1605
oil	Oil and gas	25	100	203	20659
omm	Minerals nec	43	76	68	435
met	Meat and Dairy products	35	242	9	95
ofd	Food products nec	28	168	10	94
b_t	Beverages and tobacco products	40	385	11	96
tex	Textiles	35	129	24	79
wap	Wearing apparel	80	542	25	66
lea	Leather products	58	349	38	93
lum	Wood products	36	124	17	84
ppp	Paper products publishing	20	126	14	91
p_c	Petroleum coal products	28	111	21	82
crp	Chemical rubber plastic products	27	65	28	77
nmm	Mineral products nec	22	142	45	81
i_s	Ferrous metals	21	100	74	3506
nfm	Metals nec	36	100	115	4083
fmp	Metal products	19	112	27	83
mvh	Motor vehicles and parts	25	83	13	65
otn	Transport equipment nec	43	142	18	84
ele	Electronic equipment	35	63	19	79
ome	Machinery and equipment nec	33	79	20	85
omf	Manufactures nec	43	270	20	85
ely	Electricity	22	99	8	95
gdt	Gas manufacture and distribution	36	171	23	99
wtr	Water	38	210	18	100
cns	Construction	12	458	1	100
trd	Trade	11	332	3	99
otp	Other transportation	15	163	8	98
wtp	Water transportation	48	131	50	89
atp	Air transportation	42	156	40	82
cmn	Communication	15	142	5	99
ofi	financial services nec	11	121	5	99
ins	Insurance	24	138	4	99
obs	business services nec	9	115	6	98
ros	recreational and other services	19	211	2	99
osg	public admin and defence education health	17	238	1	100
dwe	Dwellings	56	100	1	100
Tot	Total	20.4	109.4	5.6	99.7

Appendix A

New database country/region		GTAP #	GTAP Reg	Country Name
AUS	Australia	1	AUS	Australia
NZL	New Zealand	2	NZL	New Zealand
CHN	China	4	CHN	China
HKG	Hong Kong	5	HKG	Hong Kong
JPN	Japan	6	JPN	Japan
KOR	Korea	7	KOR	Korea
TWN	Taiwan	9	TWN	Taiwan
IDN	Indonesia	12	IDN	Indonesia
MYS	Malaysia	14	MYS	Malaysia
PHL	Philippines	15	PHL	Philippines
SGP	Singapore	16	SGP	Singapore
THA	Thailand	17	THA	Thailand
VNM	Viet Nam	18	VNM	Viet Nam
IND	India	21	IND	India
CAN	Canada	26	CAN	Canada
USA	United States of America	27	USA	United States of America
MEX	Mexico	28	MEX	Mexico
ARG	Argentina	30	ARG	Argentina
BRA	Brazil	32	BRA	Brazil
CHL	Chile	33	CHL	Chile
COL	Colombia	34	COL	Colombia
PER	Peru	37	PER	Peru
AUT	Austria	49	AUT	Austria
BEL	Belgium and Lux	50	BEL	Belgium
		64	LUX	Luxembourg
CEZ	Czech and SVK Republic	52	CZE	Czech Republic
DNK	Denmark	53	DNK	Denmark
FIN	Finland	55	FIN	Finland
FRA	France	56	FRA	France
DEU	Germany	57	DEU	Germany
GRC	Greece	58	GRC	Greece
HUN	Hungary	59	HUN	Hungary
IRL	Ireland	60	IRL	Ireland
ITA	Italy	61	ITA	Italy
NLD	Netherlands	66	NLD	Netherlands
POL	Poland	67	POL	Poland
PRT	Portugal	68	PRT	Portugal
ESP	Spain	71	ESP	Spain
SWE	Sweden	72	SWE	Sweden
GBR	United Kingdom	73	GBR	United Kingdom
BGR	Bulgaria	78	BGR	Bulgaria
ROU	Romania	81	ROU	Romania
RUS	Russian Federation	82	RUS	Russian Federation
SAU	Sadi Aribia	98	SAU	Saudi Arabia
TUR	Turkey	99	TUR	Turkey
EGY	Egypt	102	EGY	Egypt
MAR	Morocco	103	MAR	Morocco
ZAF	South Africa	127	ZAF	South Africa
EFTA	EFTA	74	CHE	Switzerland

		75	NOR	Norway
		76	XEF	Rest of EFTA
XEA	Rest of East Asia	3	XOC	Rest of Oceania
		8	MNG	Mongolia
		10	XEA	Rest of East Asia
		11	KHM	Cambodia
		13	LAO	Lao People's Democratic Republic
		19	XSE	Rest of Southeast Asia
XSA	rest of south Asia	20	BGD	Bangladesh
		22	NPL	Nepal
		23	PAK	Pakistan
		24	LKA	Sri Lanka
		25	XSA	Rest of South Asia
XMC	Rest of Mercosur	31	BOL	Bolivia
		36	PRY	Paraguay
		38	URY	Uruguay
ROA	Rest of America	29	XNA	Rest of North America
		35	ECU	Ecuador
		40	XSM	Rest of South America
		45	PAN	Panama
		47	XCA	Rest of Central America
		39	VEN	Venezuela
		48	XCB	Caribbean
CFT	CAFTA	41	CRI	Costa Rica
		42	GTM	Guatemala
		43	HND	Honduras
		44	NIC	Nicaragua
		46	SLV	El Salvador
XE12	Rest of EEU	51	CYP	Cyprus
		54	EST	Estonia
		62	LVA	Latvia
		63	LTU	Lithuania
		65	MLT	Malta
		69	SVK	Slovakia
		70	SVN	Slovenia
XEEU	Rest of East europe	77	ALB	Albania
		79	BLR	Belarus
		80	HRV	Croatia
		83	UKR	Ukraine
		84	XEE	Rest of Eastern Europe
		85	XER	Rest of Europe
XSU	Rest of Former SU	86	KAZ	Kazakhstan
		87	KGZ	Kyrgyztan
		88	XSU	Rest of Former Soviet Union
		89	ARM	Armenia
		90	AZE	Azerbaijan
		91	GEO	Georgia
XWS	Rest of Western Asia	92	BHR	Bahrain
		93	IRN	Iran Islamic Republic of
		94	ISR	Israel
		95	KWT	Kuwait
		96	OMN	Oman

		97	QAT	Qatar
		100	ARE	United Arab Emirates
		101	XWS	Rest of Western Asia
XNF	Rest of North Africa	104	TUN	Tunisia
		105	XNF	Rest of North Africa
XWF	West Africa	106	CMR	Cameroon
		107	CIV	Cote d'Ivoire
		108	GHA	Ghana
		109	NGA	Nigeria
		110	SEN	Senegal
		111	XWF	Rest of Western Africa
XCF	Central Africa	112	XCF	Central Africa
		113	XAC	South Central Africa
XEC	East Africa	114	ETH	Ethiopia
		115	KEN	Kenya
		116	MDG	Madagascar
		117	MWI	Malawi
		118	MUS	Mauritius
		119	MOZ	Mozambique
		120	TZA	Tanzania
		121	UGA	Uganda
		122	ZMB	Zambia
		123	ZWE	Zimbabwe
		124	XEC	Rest of Eastern Africa
XSC	Rest of South African Customs Union	125	BWA	Botswana
		126	NAM	Namibia
		128	XSC	Rest of South African Customs Union
ROW	Rest of World	129	XTW	Rest of the World
		130	NRT	Non-Reporters