A Time Series Database for Global Trade, Production and Consumption Linkage

Zhi Wang, Marinos Tsigas and Jesse Mora
United States International Trade Commission

Li Xin and Daniel Xu
China Center for Economic Research, Perking University

I. 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), 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), 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 current policy debate 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 impact of global rebalancing (Petri, 2010). As a statistical analysis tool, input-output and the closely related supply-use tables are increasingly becoming the 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.

However, contrast with this surged analytical and statistical demand, the lack of consistent global I-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 inter-country I-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], the few available inter-country 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.

This paper try to fill this void by developing a mathematical programming model integrating individual country input-output (I-O) tables with detailed bilateral trade statistics through a three-stage optimization procedure to produce a consistent annual global inter country I-O database. The global balanced trade and I-O statistics are obtained as a solution to a system of simultaneous equations that minimize a quadratic penalty function. The model deals with the data reconciliation problem at the global level first by reconciling UN and IMF reported country total merchandise and service trade statistics with each country reported total exports to and imports from the world at sector level. It results in a set of country and industry level total exports and imports, which satisfy the condition that world total export supply plus a shipping (c.i.f.) margin equals world total import demand. At the second stage, the model reconcile each country's IO table with the global consistent exports supply and import demand data from the first stage and fill the missing I-O data between benchmark years for countries that do not have annual I-O statistics, using industry level annual production, national income account and the global consistent trade data as controls. At the third and final stage, the model integrate individual country's I-O statistics with international bilateral trade statistics by distributing each country's total exports in every industry into its trading partners based on bilateral trade share computed from UN COMTRADE (Commodity Trade Statistics) database and OECD bilateral service trade statistics (as initial value), taking each country's total exports to and imports from the world derived from the first stage as controls and adjusting their distribution among the partner countries to produce a consistent annual global inter-country input-output table (ICIO). The initial allocation of bilateral trade flows in each industry to intermediate and final uses is based on UN BEC (Broad Economic Categories) method. The model can be used annually to

[1] It provides a bulk of the required data and is regularly compiled for about 40 countries across the globe, but integrating them with bilateral trade statistics into global consistent database still remains as a substantial challenge.

update the world ICIO table when industry level production and trade statistics become available.

This time-series database integrates global trade, production and consumption statistics in a consistent accounting framework. It provides a benchmark for dynamic Applied General Equilibrium (AGE model) calibration and baseline validation. It is also particularly useful for the analysis of vertical specialization in global production and the interactions between different industries at different geographic locations that contribute to the same industrial value chain.

A world production or supply chain can be seen as the distribution of value-added share among countries in a particular global industry. Within the supply chain or production network, each producer purchases inputs and then adds value, which then becomes part of the cost of the next stage of production. The sum of the value added by every stage in the chain equals the value of final goods produced by the chain. To precisely define such chains across many countries one need able to quantify the contribution of each country to the total value-added generated in the process of production (supply) of final products. An ICIO table provided the best available and consistent information that allow us to model the value-added generation process among related countries at industry average level.[2] 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 Multi-Country I-O (MCIO) table, which is the core of most current global CGE modeling database such as GTAP database. In short, the ICIO table developed in this paper not only provides 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 the table 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.

Our research in this direction is on two parallel tracks. To provide a workable dataset for our global value chain analysis and AGE modeling of processing trade, we constructed a single

[2] There are product-level approach to estimate the financial value embedded in an product and quantify how it is distributed across the many participants in the supply chain from design and branding to component manufacturing to assembly to distribution and sales. (Jason Dedrick, Kenneth L. Kraemer, Greg Linden, 2008)

year global ICIO table based on version 7 GTAP database and processing trade information from China and Mexico. The initial allocation of bilateral trade flows in the GTAP database into intermediate and final uses is also based on UN BEC (Broad Economic Categories) method and detailed trade statistics at 6-digit HS code. We use China's expanded I/O table with a separate accounts for processing exports from Koopman, Wang and Wei (2008) and obtained 2003 Mexico IO table with separate domestic and Maquiladora accounts from Mexico statistical agency, Instituto Nacional de Estadística, Geografía e Informática (INEGI). We integrate China and Mexico's IO table with version 7 GTAP database by a quadratic mathematical programming model to minimize the deviation between the resulted new data set from original GTAP data. The new database covers 26 countries and 41 sectors and was used to support our initial global AGE modeling and value-chain analysis efforts. The major results are reported in other two papers in this organized session.

Parallel with this track, we also in a process to develop a time series global ICIO database, which integrating individual country's IO tables from OECD with detailed bilateral trade statistics. It currently covers all OECD countries and important non-OECD economies such as Brazil, China and India from 1995 to 2006 and is classified at 2-digit ISIC (48 industries). Its data structure is similar to the database we build up from version 7 GTAP database and we intend to use it as the major data source of our global dynamic AGE model in the future.

Rest of the paper is organized as follows: Section II specifies the three stage optimization procedure for I-O and trade statistics reconciliation. Section III describes major data sources we used to implement the procedure. Session IV presents major description statistics computed from the estimated time series ICIO table and section V concludes the paper with a discussion on limitations of the procedure and directions of future work in this area.

II. The Three Stage Optimization Procedure

The optimization procedure deals with the data reconciliation problem at the global level first and then fills each individual country's missing annual I-O data in order to make sure the adjustment at the individual country level not violate the global consistent condition. In the first stage, the model adjusts total exports to and imports from the world for all countries and all commodity groups according to the condition that global export supply equals global import demand. At the second stage, the model fills in the missing I-O data between benchmark years for countries that do not have annual I-O statistics, taking total exports to and imports from the world derived from the first stage and other statistics reported in country's national income account as controls. Bilateral trade pattern information is integrated at the last stage with each countries total exports to and imports from the world at industry level fixed as controls. In each of the three stages, weights are assigned to the variables in the objective functions reflecting their reliabilities and thus helping determine the adjustment magnitudes. Basically, the more reliable the initial statistics are, the less they need to be adjusted. The goal of the adjustment at each stage is to make the adjusted data satisfy a set of global or bilateral consistency conditions.

2.1 First-stage

The notations used to specify the first stage programming model are as follows:

WX_{it}^{s} = Exports to the world of commodity group *i* by country *s* at year t

WM_{it}^{r} = Imports from the world of commodity group *i* by country *r* at year t

WTX_{t}^{s} = Total exports to the world by country s at year t

WTM_{t}^{r} = Total imports from the world by country r at year t

CIF_{t}^{r} = Cost, Insurance and Freight for country r total imports from the world at time t

CIF_{it} = Cost, Insurance and Freight for commodity group i total imports from the world at time t

Exports are valued at fob price, imports are valued at cif price;

RIX_{i}^{s} = reporter reliability index of commodity group i by exporter s

RIM_{i}^{r} = reporter reliability index of commodity i by importer r[3]

Sectoral index *i* is defined over commodity (industry) set I $\in \{1, 2, \ldots, n\}$; country indices *s* and *r* are defined over country set W $\in \{1, 2, \ldots, g\}$.

Using above notation, the first stage programming model is specified as:

Objective function at each year t:

[3] Definition of these reporter reliability indexes and their estimation will be discussed in detail later at section 3.

$$\text{Min S} = \sum_{s=1}^{g}\sum_{i=1}^{n}\frac{(WX_{it}^{s} - WX0_{it}^{s})^2}{(1-RIX_{i}^{s})WX0_{it}^{s}} + \sum_{r=1}^{g}\sum_{i=1}^{n}\frac{(WM_{it}^{r} - WM0_{it}^{r})^2}{(1-RIM_{i}^{r})WM0_{it}^{r}} + \sum_{r=1}^{g}\frac{(CIF_{t}^{r} - CIF0_{t}^{r})^2}{(1-RIM^{r})CIF_{t}^{r}}$$
$$+1000\sum_{s=1}^{g}\frac{(WTX_{t}^{s} - WTX0_{t}^{s})^2}{(1-RIX^{s})WTX0_{t}^{s}} + \sum_{r=1}^{g}\frac{(WTM_{t}^{r} - WTM0_{t}^{r})^2}{(1-RIM^{r})WTM0_{t}^{r}}) + \sum_{i=1}^{n}\frac{(CIF_{it} - CIF0_{it})^2}{(1-RIM_{i})CIF0_{it}} \quad (1)$$

Constraints at each year t:

Country total exports

$$\sum_{i=1}^{n} WX_{it}^{s} = WTX_{t}^{s} \qquad \text{for all s} \qquad (2)$$

Country total imports

$$\sum_{i=1}^{n} WM_{it}^{r} = WTM_{t}^{r} \qquad \text{for all r} \qquad (3)$$

World market equilibrium at commodity group level

$$\sum_{s=1}^{g} WX_{it}^{s} + CIF_{it} = \sum_{r=1}^{g} WM_{it}^{r} \qquad \text{for all i} \qquad (4)$$

Total world exports equals total world imports minus transport cost

$$\sum_{s=1}^{g} WTX_{t}^{s} = \sum_{r=1}^{g}(WTM_{t}^{r} - CIF_{t}^{r}) \qquad (5)$$

The model is used to reconcile UN national income account reported country total *merchandise* and *services* trade statistics ($WTX0^{s}$ and $WTM0^{r}$) and each country reported total exports to and imports from the world at commodity group level ($WX0_{i}^{s}$ and $WM0_{i}^{r}$) based on UN COMTRADE data.[4] It results in a set of country and sector level total exports and imports, along with value of transport cost by country and commodity group which satisfy the condition that world total export supply plus a shipping cost equal world total import demand. Detailed data source and justification to use UN national income account reported trade statistics as controls will given in next section and appendix.

[4] For most high income countries, the UN national income account reported merchandise trade data and COMTRADE data totals are identical. However, the UN provides more accurate totals based on balance of payment information for countries prone to missing or unclassified trade where COMTRADE (Commodity Trade) is lacking.

2.2 Second Stage

To fill the missing I-O statistics between benchmark years and adjust each countries' exports and imports in its I-O table to the global consistent level solved from the first stage, we also use a constrained quadratic programming model which minimizes the weighted sum of squares of deviations from the benchmark I-O table in all components of value-added, intermediate inputs, and gross outputs, and in all final expenditure categories, over all industries. The constraints are that 1) for each industry, the total intermediate inputs purchased from all commodity groups and all sources (domestic and imported) as well as value-added generated by the industry sum up to the industry's total gross output; and 2) for each commodity group, the amount sold to all industries as domestic intermediate inputs, the amount sold to the final users as domestic final goods and services, and the amount of domestic exports sum up to the total commodity output produced by the industries; 3) for each commodity group, the imported intermediates used by all industries, the amount of imported final goods used by all users, and the amount of goods re-exported minus a re-exports make-up, sum to the amount of the total imports of that commodity group, which is fixed at the global consistent level of gross imports at sector level solved from first stage; 4) the domestic exports plus re-exports equals each industry gross exports, which is fixed at the global consistent sector level gross exports solved from the first stage; 5) The sum of each type final demand by sectors plus a commodity tax equals the aggregate final demand categories in each country's GDP by expenditure account.

Formally, the second stage optimization procedure is formulated as follows:

Objective function at each year t:

$$\text{Min } S_t = \sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{u=d,m}\frac{(z_{ikt}^{u}-z0_{ikt}^{u})^2}{wz_{ikt}^{d}}+\sum_{i=1}^{n}\sum_{k=1}^{n}\frac{(x_{ikt}-x0_{ikt})^2}{wx_{ikt}}+\sum_{i=1}^{n}\sum_{f=1}^{l}\frac{(v_{ift}-v0_{ift})}{wv_{it}}$$

$$+\Big(\sum_{k=1}^{n}\sum_{c=1}^{h}\sum_{u=d,m}\frac{(y_{kct}^{u}-y0_{kct}^{u})^2}{wy_{kct}^{u}}+\sum_{k=1}^{n}\sum_{u=d,m}\frac{(e_{kt}^{u}-e0_{kt}^{u})^2}{we_{kt}} \qquad (6)$$

Column and row balance condition for IO accounts:

$$\sum_{i=1}^{n}\sum_{u=d,m} z_{kit}^{u}+\sum_{f=1}^{l} v_{kft}+tc_{it}^{d}=x_{kt} \quad \text{for all k} \qquad (7)$$

$$\sum_{k=1}^{n} z_{kit}^{d}+\sum_{c=1}^{h} y_{ict}^{d}+e_{it}^{d}=x_{it} \quad \text{for all i} \qquad (8)$$

$$\sum_{k=1}^{n} z_{kit}^{m} + \sum_{c=1}^{h} y_{ict}^{m} + e_{it}^{m} = m_{it} == WM_{it}^{r} \qquad \text{for all i} \qquad (9)$$

Gross exports and aggregate expenditure components constraint:

$$e_{it}^{d} + e_{it}^{m} == WX_{it}^{s} \qquad \text{for all i} \qquad (10)$$

$$\sum_{i=1}^{n} (y_{ict}^{d} + y_{ict}^{m}) + tc_{ct}^{d} == GDPE_{t}^{c} \qquad \text{for all c} \qquad (11)$$

GDP by economic activity constraint:

$$\sum_{i=1}^{n} \sum_{f=1}^{l} v_{ift} = GDP_{t} + sd_{t} \qquad (12)$$

GDP by expenditure constraint

$$\sum_{k=1}^{n} [\sum_{u=d,m} (\sum_{c=1}^{h} y_{kct}^{u} + e_{kt}^{u}) - m_{kt}] = GDP_{t} \qquad (13)$$

where x, z, v, y, e, and m are, respectively, gross output, intermediate inputs, value-added, final domestic demands, exports, and imports, and wx, wz, wv, wy, we, wm are their corresponding reliability weights. Subscripts (i, k), f, c and d are, respectively, indexes for industry, value-added category, final demand category and supply sources (d for domestic and m for imported). The variables with “0” stand for the benchmark values of the variables. The total number of industries, commodity groups, value-added categories, and final demand categories are, respectively, 48, 48, 3 (compensation to employees, indirect tax, and operating surplus) and 4 (household consumption, government spending, gross fixed capital formation, and changes in inventory). Detailed sector definition and their concordance with ISIC reversion 3 is follow the standard of 2010 version of OECD IO tables and listed in the appendix. All variables are evaluated at basic price. tc is tax on products and sd is the statistical discrepancy, which could be eliminate during the optimization process.

To gain information on the structural distribution of the statistical discrepancy, this quadratic programming model of the national I-O system can be solve twice, once controlling the estimations by setting right hand of equation (13) equals GDP estimated from the expenditure side; and once controlling the estimates by setting right hand of equation (12) equal to GDP estimate from production side. The differences between the two sets of estimates provide the structural distribution of statistical discrepancy by industry and by final expenditure categories.

2.3 Final Stage

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) from 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 ICIO table is a comprehensive account of annual transaction and payment flows within and between countries. We use the following notation to describe the elements of the world ICIO table (expressed in annual values): x_i^r = Gross output of industry *i* in country *r*; v_i^r = Direct value added by production of industry *i* in country *r*; z_{ij}^{sr} = Delivery of good *i* produced by country *s* and used as an intermediate by sector *j* in country *r*; and y_{ic}^{sr} = Delivery of good *i* produced in country *s* for final use in final demand type 'c' in country *r*. The total number of final demand types, such as private consumption or gross capital formation, is h). Then the following two accounting identities describe the relationship among elements of each row (*i, r*) and column (*j, s*) of the global ICIO table:

$$\sum_{s=1}^{G}\sum_{k=1}^{N} z_{ik}^{sr} + \sum_{s=1}^{G}\sum_{c=1}^{H} y_{ic}^{sr} = x_i^r \qquad (14)$$

$$\sum_{r=1}^{G}\sum_{i=1}^{N} z_{ik}^{rs} + v_k^s = x_k^s \qquad (15)$$

The economic meanings of the two equations are straightforward. A typical row in Equation (14) states that total gross output of commodity group *i* in country *r* is equal to the sum of all deliveries to intermediate and final users in all countries (including itself) in the world. Equation (15) defines the value of gross output for commodity group *k* in production country *s* as the sum of the values from all of its (domestic plus imported) intermediate and primary factor inputs. Equations (14) and (15) must hold for all $i, k \in$ N, $c \in$ H and $s, r \in$ G in each year. This ICIO account has to be consistent with each individual country's national IO account and international trade statistics, which requires the following accounting identities also to be satisfied each year:

$$\sum_{c=1}^{h}\sum_{s=1}^{g} y_{ict}^{sr} + \sum_{i=1}^{n}\sum_{s=1}^{g} z_{ikt}^{sr} = e_{it}^{sd} + e_{it}^{sm} == WX_{it}^{s} \quad \text{for all } s \neq r \qquad (16)$$

$$\sum_{c=1}^{h}\sum_{s=1}^{g} y_{kct}^{sr} + \sum_{i=1}^{n}\sum_{s=1}^{g} z_{kit}^{sr} + e_{kt}^{sm} = m_{kt}^{r} == WM_{kt}^{r} \quad \text{for all } r \neq s \tag{17}$$

Collectively, equations (14) to (17) define a commodity based ICIO account. Equation (16) indicates that all country' total imported intermediate and final demand for commodity group "i" made in country "s" must be met by goods and services exported from country "s", which include both domestic exports and re-exports, while Equation (17) states each country's demand for imported of intermediate and final goods and services plus its re-exports have to equal the country's total gross imports from international markets.

Because we solved the global consistent gross exports to and imports from the world (WX and WM) at the first stage of the optimization procedure and the domestic production, value-added and final demand (x, v, z^d and y^d) at the second stage of the optimization procedure, which will fill in the cells when s=r in equation (14) and (15), we are able to solve the full ICIO table containing number of G^2 N by N different intermediate transaction matrix ($Z^{rs} = [z_{ij}^{sr}]$, r,s$\in$G), and number of G^2 N by H different final demand transaction matrix ($Y^{rs} = [y_{ic}^{sr}]$ r,s$\in$G) at the final stage of the optimization procedure in significantly reduced dimensions.

The optimization problem in the last stage of our data reconciliation procedure is formulated to minimize a quadratic penalty function as follows subject to equations (16) and (17) as constraints.

$$\text{Min } S = \frac{1}{2}\left\{\sum_{s=1}^{g}\sum_{r=1}^{g}\sum_{i=1}^{n}\sum_{k=1}^{n}\frac{(z_{ikt}^{sr} - z0_{ikt}^{sr})^2}{wz_{ij}^{sr}} + \sum_{s=1}^{g}\sum_{r=1}^{g}\sum_{i=1}^{n}\sum_{c=1}^{h}\frac{(y_{ict}^{sr} - y0_{ict}^{sr})^2}{wy_{ic}^{sr}}\right. \quad \text{for all } s \neq r \tag{18}$$

There are several desirable theoretical properties of such a mathematical programming approach for data reconciliation. As discussed by Harrigan (1990), Canning and Wang (2004) and Wang, Gehlhar and Yao (2010), by imposing valid binding constraints, the optimization procedure will definitely improve, or at least not worsen, the initial statistics estimates. The weights (wz_{ij}^{sr}, wy_{ic}^{sr}) in the objective functions play very important role in the data reconciliation process. They have a significant impact on the model solution. The model uses these weights to determine by how much an initial estimate may be changed. For instance, using the initial trade statistics as weights has the advantage that each entry of the trade flow data is adjusted in proportion to its magnitude, in order to satisfy those consistency constraints. The variables cannot change signs and the larger the trade flows, the more adjustment takes place. However, while these features are intuitively appealing, the drawback is that the adjustment relates directly to the size of the initial trade statistics, and does not force the unreliable trade data to absorb the bulk of the required adjustment. Indeed, it is only under very special assumptions that this commonly used weighting scheme (and the one underlying RAS) will yield best unbiased estimates. Specifically this requires the following two assumptions: (1) the initial estimates for different trade flows are statistically independent, and (2) each error variance is proportional to the corresponding initial estimates. In practice, they do not hold for trade data. Therefore, the efficiency of the model will be improved if the error structure of the initial trade statistics is available. So, in a more sophisticated weighting scheme, the larger the variance, the smaller its contribution to the objective function, and hence the lesser the penalty for each adjusted trade statistics 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 trade statistics 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. In sum, we would like to adjust the trade data on an unreliably reported route more than the reliably reported one.

Advantages of such an optimization framework in data reconciliation are also significant from an empirical perspective. First, it provides considerable flexibility in achieving the global consistent conditions. It allows a wide range of initial information to be used efficiently in the data adjustment process. Additional constraints can be easily imposed to allow, for example, upper and lower bonds to be placed on unknown elements, or inequality conditions to be added. It is also very flexible regarding to the required known information and allows missing data in certain block of the I-O matrix, as long as the sum of the elements within the block is known. Such flexibility is important in terms of improving the information content of the final balanced estimates as shown by Robinson, Cattaneo, and El-Said. (2001).

Second, the optimization approach permits alternative measures of the reliability of the initial data can be easily included in the reconciliation process. The idea of including data reliability in data reconciliation can be traced back to Stone (1942) when he explored procedures for compiling national income accounts. As noted before, these weights should reflect the relative reliability of the initial statistics. The interpretation is straightforward. Initial statistics with higher reliability should be changed less than initial statistics with a lower reliability, thus the best available information can always be used to insure that initial statistics reported from reliable sources are not perturbed by the reconciliation procedure as much as initial statistics reported from unreliable sources. Using properly selected reliability weights, the optimal solution should yield estimates that deviate less from the initial estimates with higher degrees of reliability than for those with lower degrees of reliability. In Implementation, the reliability weights can be put into an array that has the same dimension and structure as the initial estimates. Therefore, considerable amount information regarding to the quality of the initial statistics could be incorporated into the data reconciliation process.

The three-stage ICIO account reconciliation procedure described above is solved with an optimization software package GAMS/CPLEX.[5] Optimal solutions from this procedure are equivalent to the estimates produced by generalized least square estimation (GLS).[6]

III Data Sources and model initialization

The key in implementing this three-stage optimization procedure to produce a useful ICIO database is to carefully link each variable in the model with best available statistics from the real world which is reported regularly by national or international statistical agencies. This section documents all the data sources we used to initialize the model and discusses their justifications. We also briefly introduce the reliability weights used in the objective function at the first and final stages of the optimization procedure.

3.1 Sector classification and country coverage

Our objective is to construct a preliminary version world ICIO table by integrating the OECD individual country I-O tables and international trade statistics using the three stage optimization procedure specified in this paper. Therefore, the sector classification and country coverage of our database follows closely with the 2010 version OECD STAN I-O database. It is classified into 48 industries, including 3 primary sectors, 22 manufacturing sectors, 4 utility producing sector plus construction and 18 additional services sectors. This is very close to 2-digit ISIC (rev.3) with some exceptions. It covers 40 individual economies, including all OECD

[5] GAMS/CPLEX is a well established, versatile, high-performance optimization system powerful for solving large linear and quadratic programming models.

[6] Since the optimal solutions are equivalent to the GLS estimates, the term “optimal solution” and “estimates” are sometimes used interchangeably here.

countries and important non-OECD economies such as Brazil, China and India plus an aggregated rest of the world block. First year of the data is 1995 and currently cover 12 years until 2006. The detailed sector classification and their concordance with ISIC rev.3 are listed in Appendix table A1 and A2.

3.2 Aggregate international trade statistics

We looked at various sources for goods and services trade data: UN National Accounts Data, UNCTAD, IMF's IFS database, IMF's BOP database, WITS-COMTRADE, and OECD. There were significant differences in the data, mostly due to valuation (trade valued on a f.o.b. (free on board) or c.i.f. (cost, insurance, and freight) basis) and data coverage (data missing for some countries and for some years). Because of this difference, we had to decide which source to use. Ultimately, we decided to use the totals provided by the UN *(National Account Estimates of Main Aggregates, GDP by type of expenditure at current prices*) for a number of reasons. This data matches other sources, once valuation and other differences are taken into account. The additional benefit from using UN data is it being consistent with each country's GDP by expenditure statistics, which is a nice consistence feature in our data reconciliation task and was also available through the UN data website (same source). Furthermore, imports and exports are both in f.o.b. basis, so that we did not have to remove the c.i.f. margin out of imports to create a balanced volume of total world trade. Finally, the data cover both goods and services, so the estimates for goods and services trade are uniformed for the UN data, something that would not have been the case if we use aggregate trade data from different sources.

The only major drawback to using the UN data was that the data (in US dollars), did not separate goods and services trade. To split the total, we used the UN's National Accounts table in local currency, downloaded from the same website under "*National Accounts Official Country Data, Gross Domestic Product by expenditure at current prices*". We used the share of goods and services in total exports and imports in local currency to create each county's share of goods and services by year and flow (i.e. imports and exports). The shares allowed us to split the total goods and services data in the *National Account Estimates of Main Aggregates* into each country's goods and services. To ensure that the split of the UN data was done correctly, we compared the split total numbers with trade in both goods and services from the other sources.

3.2.1 Country total merchandise exports to and imports from the world

The UN data is consistent across various sources for world totals (see Table 3.2.1). For the world's total exports, the UN data are a bit larger than that those found in the UNCTAD, WITS-COMTRADE, IFS, and BOP databases. The difference is due to estimation differences and missing data from some sources (in some cases, such as the WITS-COMTRADE, the "world" totals were created from "all countries" available, which do not include estimates of missing data). Imports of merchandise, on the other hand, tend to be larger for UNCTAD, IFS, and BOP than that for the UN data. This is a result of valuation differences (UNCTAD and IFS are both in c.i.f.; WITS-COMTRADE data is also c.i.f. and, if it weren't for missing data, it would also be larger than the UN data) and to definitional differences (IMF's BOP data is only for merchandise goods, while BOP2 includes merchandise goods plus goods for process, repair of goods, goods procured in ports by carriers, and non-monetary gold). These differences are highlighted when we add imports and exports for each of the sources and compare the totals with that of the UN data (see Figure 3.2.1). WITS-COMTRADE data is significantly smaller in the earlier years, this is because many more countries did not report data in 1995 compared to the later years. The data drops off in 2007 for the same reason. UNCTAD data is quite similar, probably due to similar techniques used to estimate missing data. The fact that IFS data is a bit larger is expected, imports from IFS are values on c.i.f. basis. BOP2 is significantly larger than the UN data even though all trade is valued in f.o.b. basis, we believe this is due to the definition of BOP2, which includes non-monetary gold trade.

Table 3.2.1: Comparing World Goods or Services Trade Data (Various Sources as a Percent of UN National Accounts data)

year		1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Source	Type	Exports												
UNCTAD	Goods	97	97	97	98	97	98	98	98	99	99	98	99	98
WITS	Goods	89	92	94	95	95	96	97	97	98	98	97	98	96
IFS	Goods	94	95	98	98	98	98	98	99	99	99	100	100	98
BOP	Goods	84	85	83	85	85	84	84	82	83	83	83	82	82
BOP2	Goods	90	91	91	110	109	107	108	109	110	110	109	108	108
UNCTAD	Serv.	101	101	101	101	101	101	101	100	100	100	100	102	104
BOP	Serv.	85	84	85	88	88	88	88	87	87	87	87	88	91
		Imports												
UNCTAD	Goods	100	101	101	101	100	101	101	102	102	102	102	102	101
WITS	Goods	92	95	98	98	99	100	100	101	101	101	101	101	100
IFS	Goods	98	99	103	103	102	102	102	102	103	103	103	102	102
BOP	Goods	84	85	84	87	87	87	86	86	86	86	86	86	86
BOP2	Goods	90	91	91	109	109	109	108	109	110	109	110	110	110
UNCTAD	Serv.	101	99	99	99	99	99	100	98	98	98	98	99	101
BOP	Serv.	85	83	82	85	86	85	86	84	83	83	82	82	84

Source: UN, UNCTAD, WITS-COMTRADE, IMF BOP, and IMF IFS databases

Figure 3.2.1: Comparing Data Sources for Goods (g) and Services (s): World Imports + Exports (Various Sources as a Percent of UN National Accounts Data)

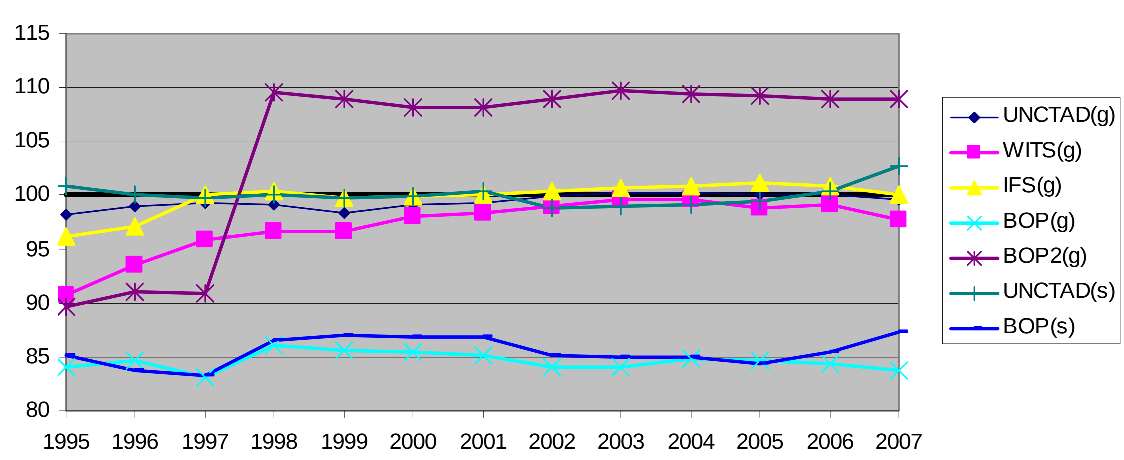


Table 3.2.2 provides the same comparison for merchandise trade, but looking only at four countries: China, Japan, Germany, and the US. By focusing on these four major exporters and importers, we can provide a more accurate comparison between the UN data and other sources as it eliminates one key difference: estimation of missing data. Looking at these four countries, we

can clearly see that the UN data is quite close to that of other sources, especially in the case of merchandise exports. For merchandise exports, UN data totals are about 100 percent for all years for China, Germany, and the US. UN export data for Japan seems to be smaller than that found in the OECD, WITS, and IFS, but 100 percent for that found in BOP2. Data for BOP is smaller, but that is expected due to its definition (see above). Merchandise imports for most sources are clearly larger than the UN data, with the exception of the BOP2 database. The data from UNCTAD, WITS, and IFS are on average about 5 percent larger for China, 1 percent for Germany, 10 percent for Japan, and 2 percent for the US; these difference are a result of the c.i.f. margin. The data from BOP2 is about 100 percent, which is has the same valuation as that of the UN data (f.o.b.). These results convinced us that there we more benefits to using the UN data than there were benefits from using any other source.

Table 3.2.2: Comparing Merchandise Trade Data for Selected Countries
(Various Sources as a Percent of UN National Accounts data)

reporter	Source	Exports							Imports						
		1995	1997	1999	2001	2003	2005	2007	1995	1997	1999	2001	2003	2005	2007
China	UNCTAD	102	100	100	100	100	100	100	107	104	104	105	105	105	106
	WITS	102	100	100	100	100	100	100	107	104	104	105	105	105	106
	IFS	102	100	100	100	100	100	100	107	104	104	105	105	105	106
	BOP	88	45	43	44	45	45	49	89	51	55	61	60	58	61
	BOP2	88	100	100	100	100	100	100	89	100	100	100	100	100	100
Japan	UNCTAD	104	103	104	105	105	105	106	113	110	110	111	111	109	109
	WITS	104	103	104	105	105	105	106	113	110	110	111	111	109	109
	IFS	104	103	104	105	105	105	104	113	110	111	111	111	109	108
	BOP	100	98	99	98	98	98	98	100	96	95	96	95	95	95
	BOP2	100	100	100	100	100	100	100	100	99	99	100	99	100	100
Germany	UNCTAD	100	100	100	100	100	99	99	101	102	100	101	100	99	99
	WITS	100	100	100	100	100	99	99	101	102	100	101	100	99	99
	IFS	100	100	100	100	100	99	98	101	102	100	101	100	99	99
	BOP	94	94	93	94	93	93	95	92	93	93	93	93	93	95
	BOP2	99	99	99	99	99	99	100	99	100	99	100	100	100	101
United States	UNCTAD	100	100	99	100	100	100	101	102	101	101	101	102	102	102
	WITS	100	100	99	100	100	100	101	102	101	101	101	102	102	102
	IFS	100	100	100	100	100	100	101	102	102	101	101	102	102	102
	BOP	97	97	97	97	98	97	97	98	98	98	98	98	98	98
	BOP2	98	98	98	98	99	98	98	98	98	98	98	98	98	99

Source: UN, UNCTAD, WITS-COMTRADE, OECD, IMF BOP, and IMF IFS databases

3.2.2 Country total services exports to and imports from the world

As mentioned above, total trade in services was created using the same source (UN data) and technique used for splitting total goods trade from total goods and services. The split of services trade data is also fairly consistent with other sources. For example, world totals found in UNCTAD data on services trade are almost 100 percent of those of the split UN data (see Table 3.2.1). However, UN totals are between 9 and 18 percent larger than those found in the IMF's BOP database; the difference in highlighted in Figure 3.2.1. This variation is a result of having to create the "world" from all the countries that had data; the IMF's BOP database does not estimate a "world" total and many countries had missing data. This difference in totals, however, does not exist in the country totals. For example, Table 3.2.3 shows that services trade data for most years, from most sources, including the BOP database, are 100 percent of the UN data for both services exports from, and import to, China and Germany. They are about 30 and 17 percent larger for Japan's exports and imports, respectably. For the US, the services trade data are about 5 and 3 percent larger for US exports and imports, respectively.

Table 3.2.3: Comparing Services Trade Data for Selected Countries (Various Sources as a Percent of UN National Accounts data)

Reporter	Source	Exports							Imports						
		1995	1997	1999	2001	2003	2005	2007	1995	1997	1999	2001	2003	2005	2007
China	UNCTAD	88	100	100	100	100	100	100	89	100	100	100	100	100	100
	BOP	88	100	100	100	100	100	100	89	100	100	100	100	100	100
	OECD	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Japan	UNCTAD	124	128	130	131	130	131	137	111	116	117	117	118	117	120
	BOP	124	128	130	131	130	131	137	111	116	117	117	118	117	120
	OECD	0	0	130	131	129	131	0	0	0	114	116	115	117	0
Germany	UNCTAD	100	99	95	99	104	105	103	100	100	98	98	101	101	101
	BOP	100	99	95	100	104	105	106	100	100	101	101	101	101	101
	OECD	0	0	96	100	103	99	0	0	0	102	101	101	99	0
United States	UNCTAD	95	95	95	94	95	96	96	97	97	97	96	98	98	98
	BOP	95	95	95	94	95	96	98	97	97	97	96	98	98	97
	OECD	0	0	96	95	96	96	0	0	0	97	96	98	99	0

To separate the total data into IO sectors we used the IMF's BOP database which has trade in services for all of our covered countries (except Taiwan) by sectors. We mapped these sectors to our IO services sectors (see IMF to IO concordance in Appendix A)**.** However, most of the data were not very disaggregated. We will have to split this data using a different technique. We also used data from Hong Kong Census and statistical department as the data from the BOP database were not as disaggregated as the data we received from Hong Kong official sources.

3.2.3 Total exports to and imports from the world and total world trade in value

Another benefit of the UN data is that it is fairly balanced. Looking at the share of imports over exports of world totals (see Table 3.2.4), allows us to compare the global trade balance of the different sources; in a perfectly balanced world trade this share would equal to 100 percent. The data show that on average imports account for 99 percent of exports for the UN data (goods, services, and total). Imports from UNCTAD, IFS, and WITS are predictably larger by about 2 percent. This difference results when exports are valued on a f.o.b. basis and imports are valued on a c.i.f. basis.

Table 3.2.4: World Data (Share of Imports over Exports by Source)

	Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
UN (Total)		98	98	98	99	99	100	100	99	99	99	99	98	98
UN	Goods	98	98	97	98	99	99	100	99	99	99	99	99	98
UNCTAD	Goods	101	101	101	101	102	103	103	102	102	102	103	101	101
IFS	Goods	102	103	102	102	103	103	104	103	103	103	102	101	102
BOP	Goods	97	98	99	100	101	103	103	103	103	103	103	103	103
BOP2	Goods	97	98	98	97	99	101	100	99	99	99	100	100	100
WITS	Goods	101	102	102	101	103	103	103	102	102	103	103	102	103
UN	Serv.	100	100	100	100	100	100	100	100	100	98	98	96	96
UNCTAD	Serv.	100	99	98	98	98	98	99	98	98	96	96	94	93
BOP	Serv.	100	98	97	97	97	96	97	96	96	94	92	90	88

Source: UN, UNCTAD, WITS-COMTRADE, IMF BOP, and IMF IFS databases

3.3 I-O Statistics from OECD and National Sources

Our major benchmark I-O statistics are taken from 2010 version OECD STAN individual country I-O database as is with the exception of China and Taiwan as well as those tables valued at producer prices. We also use annual I-O data from EU and BEA of the United States as supplementary I-O data source. We also constructed 3 individual I-O tables for rest of the world from version 5, 6, 7 GTAP database and compiled annual I-O tables for Hong Kong from 1995 to 2007 under the assistance of Hong Kong Census and Statistics Department.[7] The country and year coverage, price valuation, table types, currency unit and their sources of the I-O statistics used to initialize our model in the second stage are summarized in table 3.3.1.

OECD I-O database only include three years' US commodity by commodity I-O tables (1995, 2000, 2005). We take the annual use, make and import use matrix published by the Bureau of Economic Analysis (BEA) to construct C by C annual IO tables for 1996-1999, 2001-2004 and 2006 to ensure time continuity and data source consistence, The Commodity-by-Commodity symmetric annual US I-O tables at 65 sectors are aggregated or disaggregated into the OECD 48 sectors according to a bridge map between NAICS1997 to ISIC Rev.3.

OECD I-O database includes 1995, 2000, and 2005 China I-O tables. However, there are 14 industries among the 48 OECD IO sectors with all zero entries in both rows and columns in these tables. The missing sector information may due to the original China I-O tables are too aggregate since China only construct survey based detailed benchmark tables at each 2s and 7s, while only publish an extension table in 42 sector in each 5s and 0s. In order to better reflect the real economic structure in China, we process the 1997, 2002 and 2007 benchmark I-O data provide by the National Bureau of Statistics and developed detailed concordance between China's benchmark I-O tables and OECD I-O sectors. There are 124 sectors in 1997, 122 sectors in 2002 and 135 sectors in 2007 of China's benchmark I-O tables. We first concord 2002 version of CSIC (classification of economic activities, GB/T4754) to ISIC Rev.3 and use it as bridge to aggregate China's benchmark I-O data to OECD I-O sectors. During the process, some sectors in the benchmark IO were further split into 2 or more sectors by gross output data from NBS industrial firm survey to make better match with the OECD sector classification. Further work still need to better split service sectors and estimate commodity taxes (tax on product).

[7] The initial data collection and model set up work was done at Hong Kong during January 13 to February 17 jointly with professor Sung Yin-win during Dr. Zhi Wang visit Department of Economics at Chinese University of Hong Kong as a short term Fulbright Scholar.

Table 3.3.1 Summary of I-O Data used to initialize the model

ISO3	Country	Year	Valuation	Type	Currency[a]	Source:
ARG	Argentina	1997	Basic	ii	Mill. Pesos	OECD IO (2010ed)
AUS	Australia	98/99、04/05	Basic	ii	Mill. AUD	OECD IO (2010ed)
AUT	Austria	1995、2000、2004、2005	Basic	ii	Mill. Euros[b]	OECD IO (2010ed)
BEL	Belgium	1995、2000、2004	Basic	ii	Mill. Euros	OECD IO (2010ed)
BRA	Brazil	1995、2000、2005	Basic	ii	Mil. Real	OECD IO (2010ed)
CAN	Canada	1995、2000、2005	Basic	ii	Mill. CAN D	OECD IO (2010ed)
CHE	Switzerland	2001	Basic	ii	Mill. CHF	OECD IO (2010ed)
CHN	China	1997、2002、2007	Producer	Cc	Mill. RMB	NBS of China
CZE	Czech	2000、2005	Basic	ii	Mill. CZK	OECD IO (2010ed)
DEU	Germany	1995、2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
DNK	Denmark	1995、2000、2004、2005	Basic	ii	Mill. Kroner	OECD IO (2010ed)
EST	Estonia	1997、2000、2005	Basic	ii	Mill. Kroon	OECD IO (2010ed)
ESP	Spain	1995、2000、2004、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
FIN	Finland	1995、2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
FRA	France	1995、2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
GBR	UK	1995、2000、2003、2005	Basic	ii	Mill. Pound	OECD IO (2010ed)
GRC	Greece	1995、1999、2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
HUN	Hungary	1998、2000、2005	Basic	ii	Mill. Forint	OECD IO (2010ed)
IND	India	93/94、98/99、03/04	Basic	ii	Mill. NAC	OECD IO (2010ed)
IRL	Ireland	1998、2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
ITA	Italy	1995、2000、2004、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
LUX	Luxembourg	1995、2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
NLD	Netherlands	1995、2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
NZL	New Zealand	95/96、02/03	Basic	ii	Mill. NZD	OECD IO (2010ed)
NOR	Norway	1995、2000、2001、2005	Basic	ii	Mill. Krone	OECD IO (2010ed)
MEX	Mexico	2003	Basic	ii	Mill. Pesos	OECD IO (2010ed)
POL	Poland	1995、2000、2004、2005	Basic	ii	Mill. Zloty	OECD IO (2010ed)
PRT	Portugal	1995、1999、2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
SVK	Slovak Republic	1995、2000、2005	Basic	ii	Mill. Koruna	OECD IO (2010ed)
SVN	Slovenia	2000、2005	Basic	ii	Mill. Euros	OECD IO (2010ed)
SWE	Sweden	1995、2000、2005	Basic	ii	Mill. Krona	OECD IO (2010ed)
TWN	Taiwan	1996、1999、2001、2004□2006	Producer	Cc	Mill. NT$	Directorate General of Budget, Accounting and Statistics.
ZAF	South Africa	1993、2000	Basic	ii	Mill. Rand	OECD IO (2010ed)
USA	United State	1995, 1997 — 2007	Producer	Cc	Mill. USD	U.S. BEA & OECD
IDN	Indonesia	1995	Basic	Cc	Mill. Rps	OECD IO (2010ed)
		2000	Basic	Cc	Mill. Rps	
		2005	Producer	Cc	Mill. Rps	

JPN	Japan	1995	Producer	Cc	Mill. JPY	OECD IO (2010ed)
		2000	Basic	ii	Mill. JPY	
		2005	Producer	Cc	Mill. JPY	
KOR	Korea	2000	Basic	Cc	Mill. Won	OECD IO (2010ed)
		2005	Producer	Cc	Mill. Won	
RUS	Russia	1995	Basic	Cc	Mill. RUB	OECD IO (2010ed)
		2000	Basic	ii	Mill. RUB	
TUR	Turkey	1996	Producer	ii	Bill. TRL	OECD IO (2010ed)
		1998	Basic	ii	Bill.TRL	
		2002	Basic	ii	Bill. TRL	See footnote 7
HKG	Hong Kong	1995-2007	Basic	Cc	Million HKD	
ROW	The rest of World	1997, 2001, 2004	Basic	Cc	Mill. USD	GTAP 7 and 5 version

ii: Industry-by-industry symmetric matrix
cc: Commodity-by-Commodity symmetric matrix
All the countries include separate domestic and imported tables

Notes:

In original OECD 2010ed I-O database, ARG is at 1000 pesos unit, IDN is at Rs. In Lakhs, TUR of 1998 is at Mill. TRL. Here, in order to consistent to the production data, we adjusted IDN and ARG to Mill. Unit and TUR to Bill.

In Euros from 1999 onwards, before 1999, AUT, BEL, FIN, FRA, DEU, IRL, ITA, LUX, NLD, PRT, ESP are converted to Euro with the 1999 official fixed Euro conversion, which the office exchange rate are 13.7603 ATS/EUR, 40.3399 BEF/EUR, 5.94573 FIM/EUR, 1.95583 DEM/RUE, 0.787564 IEP/EUR, 1936.27 ITL/EUR, 40.3399 LUF/EUR, 2.20371 NLG/EUR, 200.482 PTE/EUR & 166.386 ESP/EUR. GRC is converted to Euros with the 2001, the rate is 340.750 GRD/EUR.

Taiwan statistical agency (Directorate General of Budget, Accounting and Statistics, Chinese Taipei) publish very detailed I-O statistics: benchmark table at 596 sectors details in 1996, 610 sector details in 2001, and 554 sectors details in 2006 as well as extension tables at 162 sectors in 1999 and 165 sectors in 2004. Therefore, we re-aggregate the 1996, 1999, 2001, 2004 and 2006 Taiwan I-O data to OECD I-O sector classifications using ISIC rev. 3 code as the bridge (we constructed a concordance file between Taiwan I-O industry classification at 1996, 2001 and 2006 to OECD I-O sectors).

In many aspects, Hong Kong has high quality statistics in world standard. However, there are no officially compiled input-output tables as most of other economies in the world. The few IO tables that existed in history were all complied by academic researchers. The most updated input-output table for Hong Kong was constructed by Sung (1983) for the year 1980. It has 90 sectors, including 69 manufacturing sectors, reflecting the importance of manufacturing at that time. It was compiled by using industry production surveys data alone, no data from the expenditure sources was used so there is no final demand block in the table and is also too obsolete for both analytical and statistical applications. Working with researcher at Chinese university of Hong Kong, we obtained industry survey data and private expenditure by rather detailed categories from Hong Kong Census and Statistical department, so are able apply our optimization procedure to estimates the first complete annual I-O tables (with final demand) for Hong Kong. Detailed data sources and special constraints added based on the features of Hong Kong production, trade and consumption statistics available to us is documented in a separate paper by Sung and Wang(2010).

Except the 40 individual countries' I-O statistics, we aggregate all the other countries as a whole as a rest of World (ROW) block. We built 1997, 2001 and 2004 ROW domestic and imported I-O tables from version 5, 6 and 7 GTAP database. The 2004 ROW I-O table includes all countries in GTAP database except the 40 individual countries covered by our ICIO database. Different from the 2004, the ROW 2001table does not included rest of EFTA (Iceland and Liechtenstein), and the ROW 1997 table in addition does not include Lesotho, Namibia and Swaziland.

Since the concordance between OECD IO sector and GTAP sector are in perfect, some sector need be further split. We use the average proportion of 10 non-OECD countries' information to split the sector need to be disaggregated, because countries in the ROW are all belonging to developing and underdeveloped countries. These countries including IND, IDN, CHN, TWN, ZAF, RUS, MEX, ARG, BRA and NZL.

For which sectors in the original data were disaggregated into which OECD I-O sectors please see the country date notes in appendix A3 for detail.

3.4 Gross output, intermediate input and value-added by country and industry between benchmark years

Our production database include gross output, intermediate inputs, total value added, labor compensation, gross operating surplus and net taxes on production. They are used as column controls when we fill the missing I-O data between benchmark years. All production data for OECD countries come from STAN Industrial Analysis Database 2010ed. In this newest version, most of the countries are updated to year 2006, except CAN and AUS. For these two countries we obtained the data from the *United Nations Statistics Division's website,* and disaggregated the aggregated sectors into more detailed OECD IO sector by last year's proportion.

For the most Non-OECD countries except China, Taiwan and Hong Kong, including IDN, BRA, IND, RUS, TUR, ZAF, ARG, the production information are obtained from UN database and the detailed OECD IO sectors have been split by data draw from their respect benchmark years' I-O tables. The year of coverage, evaluation price, currency unit and their source of all production data are summarized in table

3.5 Macro control data from UN national income account and exchange rate

We use the GDP by major expenditure components statistics as each country's macro control variables. The data are downloaded from *National Accounts Official Country data* of *UN statistics division,* at current price, in thousand USD. The UN database including all the countries we considered except TWN. TWN's GDP by expenditure data is come from Directorate-General of Budget, Accounting & Statistics (DGBAS).

The exchange rate (national currency per USD), which is year average data from 1995 to 2008, is downloaded from *Key Global Indicators* of *UN statistic Division*.

Table 3.5.1 Summary of production data

ISO3	Countries	Type	Available	Source	Currency
AUS	Australia	Basic	1995-2006	OECD STAN 2010ed &UN	AUD Mill.
AUT	Austria	Basic	1995-2008	OECD STAN 2010ed	Mill. Euros
BEL	Belgium	Basic	1995-2008	OECD STAN 2010ed	Mill. Euros.
CAN	Canada	Basic	1995-2006	OECD STAN 2010ed & UN	CAD Mill.
CZE	Czech Republic	Basic	1995-2007	OECD STAN 2010ed	CSK Mill.
DNK	Denmark	Basic	1995-2008	OECD STAN 2010ed	DKK Mill.
FIN	Finland	Basic	1995-2008	OECD STAN 2010ed	Mill. Euros
FRA	France	Basic	1995-2007	OECD STAN 2010ed	Mill. Euros
DEU	Germany	Basic	1995-2007	OECD STAN 2010ed	Mill. Euros
GRC	Greece	Basic	1995-2008	OECD STAN 2010ed	Mill. Euros.
HUN	Hungary	Basic	1995-2007	OECD STAN 2010ed	HUF Mill.
IRL	Ireland	Basic	1995-2007	OECD STAN 2010ed	Mill. Euros.
ITA	Italy	Basic	1995-2007	OECD STAN 2010ed	Mill. Euros.
JPN	Japan	Producer	1995-2006	OECD STAN 2010ed	JPY Mill.
KOR	Korea	Basic	1995-2007	OECD STAN 2010ed	KRW Mill.
LUX	Luxembourg	Basic	1995-2008	OECD STAN 2010ed	Mill. Euros.
NLD	Netherlands	Basic	1995-2008	OECD STAN 2010ed	Mill. Euros.
NZL	New Zealand	Basic	1995-2006	OECD STAN 2010ed	NZD Mill.
NOR	Norway	Basic	1995-2007	OECD STAN 2010ed	NOK Mill.
POL	Poland	Basic	1995-2007	OECD STAN 2010ed	PLZ Mill.
PRT	Portugal	Basic	1995-2006	OECD STAN 2010ed	Mill. Euros.
SVK	Slovak Republic	Basic	1995-2008	OECD STAN 2010ed	Mill. SKK.
ESP	Spain	Basic	1995-2007	OECD STAN 2010ed	Mill. Euros.
SWE	Sweden	Basic	1995-2006	OECD STAN 2010ed	SEK Mill.
CHE	Switzerland	Basic	1995-2007	OECD STAN 2010ed	CHF Mill.
GBR	United Kingdom	Basic	1995-2006	OECD STAN 2010ed	GBP Mill.
USA	United State	Producer	1995-2007	OECD STAN 2008ed	Mill USD
CHN	China	Producer	1995-2007	Estimated by Chinese Statistics Yearbook & Industrials Statistics Yearbook	Mill. RMB
MEX	Mexico	Basic	1995-2006	OECD STAN 2005ed & UN	Mill. Pesos
TWN	Chinese Taipei	Producer	1995-2008	Directorate General of Budget, Accounting and Statistics (DGBAS) of Taiwan	Mill. NT
SVN	Slovenia	Basic	1995-2006	EUKLEMS database, UN	Mill. Euros
EST	Estonia	Basic	1995-2006	EUKLEMS database, UN	Mill. of Kroons
ARG	Argentina	Basic	1995-2006	LAB (95-97), OTX (95-97), GOP (95-97) come from UN Database, 98-06 is estimated according to benchmark year.	Mill. Pesos
BRA	Brazil	Basic	1995-2006	UN Database	Mill. Real
IND	India	Basic	1995-2006	UN Database	Mill. INR
RUS	Russia	Basic	1995-2007	UN Database	Mill. RUB
TUR	Turkey	Basic	1995-2007	UN Database	Bill. TRY
ZAF	South Africa	Basic	1995-2007	UN Database	Mill. ZAR
IDN	Indonesia	Basic	1995-2006	UN Database,I1-3, I26-48 are come from Statistics Indonesia	Mill. Rps
ROW	The rest of World	Basic	1995-2008	UN Database	Mill.USD

3.6 Bilateral trade statistics

We used WITS-COMTRADE bilateral merchandise trade and OECD bilateral services trade data to split the countries trade totals by partner and IO sector. It should be noted that Belgium and Luxembourg data before 1999 is not split in some sources (both for Belgium-Luxembourg as a reporter and as a partner). In such cases, we split the data using the latest year when Belgium and Luxembourg were separated to split the pre-split data from Belgium-Luxembourg. Doing this allowed us to split the data, but we were unable to recreate the intra-country trade, which is not an insignificant amount of each countries trade.

3.6.1 Bilateral merchandise trade statistics from WITS-UNCOMTRADE

For our bilateral merchandise imports and exports data we used the most comprehensive, disaggregated source: WITS-COMTRADE. We downloaded bilateral HS 6- digit level of desegregation by partner. We mapped the HS 6-digit data to the 48 IO sectors (see HS combined to IO sectors concordance in Appendix B1). We used the HS to ISIC Rev. 3 concordance with the ISIC Rev.3 to IO concordance to create an HS combined to IO sector concordance. We manually fixed the concordance for the HS that did not mapped to any, or mapped to multiple, IO sectors.

One serious problem with the WITS-COMTRADE data is that some data is not disaggregated for all trade, as a result rarely will the sum of 6-digit data equal the totals. Another flaw with the data is that some countries do not provided a split between exports or re-exports. Some countries provide both exports and imports separate from re-exports and re-imports, some only provide net imports and exports (i.e. not including re-exports and re-imports) and others provide only gross imports and exports (i.e. including re-exports and re-imports). To see a list of countries that do not provide separate data for re-exports or re-imports, see Appendix B2. Most of the totals data discusses in Section 3.2 includes re-exports and re-imports. Thus, to be consistent, we use “gross” imports and exports for all countries, which is consistent with the UN classification. Another issue that needs to be resolved is that WITS-COMTRADE data for merchandise imports are in c.i.f. basis, which is not consistent with our total trade, which is on an f.o.b. basis. Finally, we should note that the HS trade data comes in three different HS classifications: HS 1988/92 (H0), HS 1996 (H1), and HS 2002 (H2). See Appendix B3 in data availability list. For most countries, the totals are the same and we used the latest classification available per year. For those countries that had different totals by year per HS classification, we used the classification with the most data (i.e. whose value was the largest).

3.6.2 Bilateral Service trade statistics from IMF and OECD

For bilateral trade in services we used the OECD's trade in services database: *Detailed tables by service category Vol 2008 release 01.* However, the data available is mostly from developed countries and is only available between 1999 and 2006 (see Appendix B4 on country availability for the OECD services trade database for details). We mapped the OECD services sectors to our IO services sectors (for OECD services sectors to IO sectors concordance see Appendix B5)**.**

3.7 **The choice of the reliability weights in the objective function**

As we discussed in the model specification section, the reliability weights in the objective function play a significant role in determine the magnitude of adjustment to the initial estimates. As pointed out by (Wang, Gehlhar and Yao, 2010), the best reliability weights is the inverse of variance-covariance matrix of the initial statistics from a statistical point of view, however, the historical data and knowledge of the changes in related country's statistical system are too demanding and make such a statistic method less attractable in practice. So they suggested using reporter reliability indexes as an empirical alternative for the weights in the quadratic penalty function. We adopted their approach in implementing our model.

The index is calculated as the share of accurately reported transactions of a reporter's total trade 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 specific reliability information. Specifically, the importer-sector specific and exporter-sector specific reliability indexes are defined as:

$$RIM_i^r = \frac{MA_i^r}{\sum_s M_i^{sr}} \qquad where \qquad MA_i^r = \sum_{s \in AL_{it}^{sr} \leq 0.20} M_i^{sr} \qquad AL_i^{sr} = \frac{\left| M_i^{sr} - X_i^{sr} \right|}{M_i^{sr}} \qquad (19)$$

$$RIX_i^s = \frac{XA_i^s}{\sum_r X_i^{sr}} \qquad where \qquad XA_i^s = \sum_{s \in AL_{it}^{sr} \le 0.20} X_i^{sr} \qquad AL_i^{sr} = \frac{\left| M_i^{sr} - X_i^{sr} \right|}{M_i^{rs}} \tag{20}$$

Weighted by related trade flows, the reporter reliability indexes for each country could be obtained as follows:

$$RIM^r = \frac{WM_i^r}{\sum_r WM_i^r} RIM_i^r \tag{21}$$

$$RIX^s = \frac{WX_i^s}{\sum_s WX_i^s} RIX_i^s \tag{22}$$

where M_t^{sr} and X_t^{sr} are sector *i* imports and exports reported by country *r* and *s* respectively, both measured at fob prices. 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. The weights in the quadratic penalty functions (equations (1) and (18)) are assigned by multiplying one minus these indexes with their corresponding initial values for each variable in the model.

With such a weighting scheme, we also achieve the goal to encourage the model to change those initial trade statistics reported by unreliable reporters more than those reported by reliable ones in the data reconciliation process.

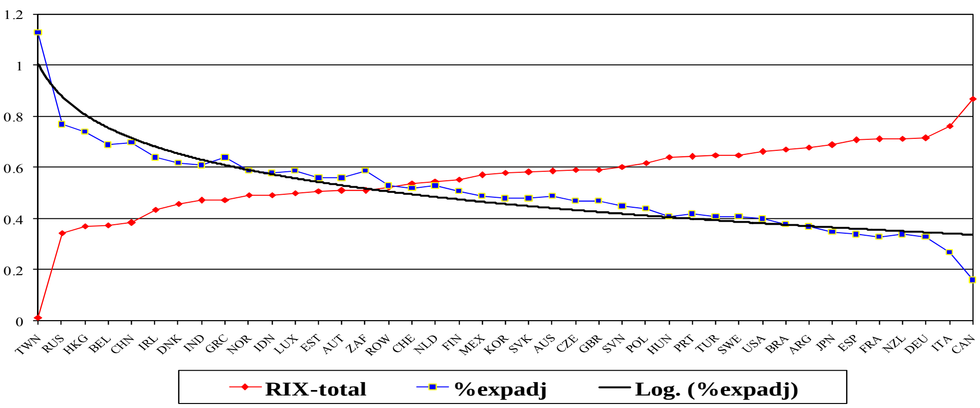
IV. Preliminary Results from the Model

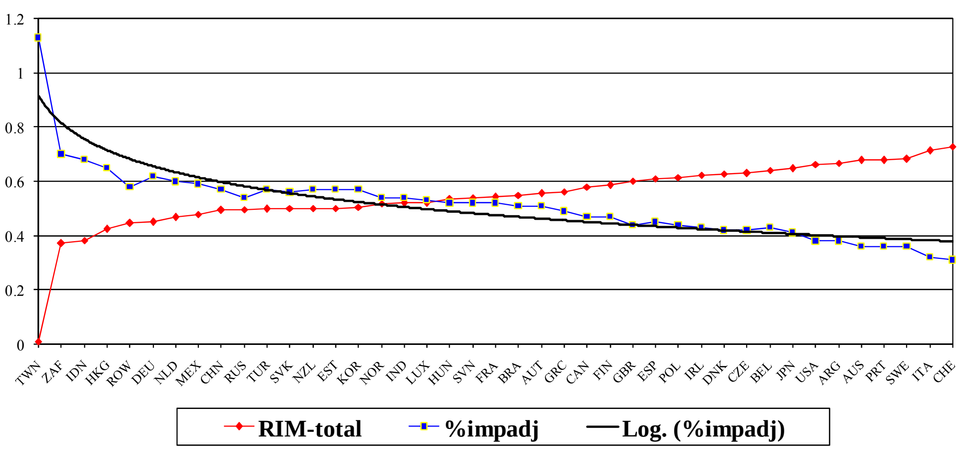
4.1 Major results from the first stage optimization of the model: Global consistent total exports and imports by country and commodity groups

Our model entails enforcing global consistency which takes place in the first stage. We first establish consistency between country-reported commodity trade data and UN's official total merchandise trade statistics. The model solves for the adjusted country total exports to and imports from the world for each covered commodity group and these sector totals are retained for the second and final stages as controls.

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 total exports and imports (Figures 4.1.1a-b). For the UN country totals, the exporter and importer reliability curves (line with dot markers) follow quite close in shape and magnitudes with their respective adjustment curves (line with square markers). The magnitude of adjustment made by the model is relative small, less than 2 percent for most countries in earlier years and less than 1 percent in more recent years. We note also that exporters and importers' reliability is also fairly consistent for adjustments magnitudes made to covered sectors. As expected, both the country and sector patterns of the adjustments reflect their negative relationship with reporter's reliability, with the exception of a few outliers.

Figure 4.1.1 Reporter Reliability and Percentage of Adjustment UN country total exports, and imports, 2006





4.2 Results from the second stage optimization of the model: Difference between adjusted country I-O table with existing tables

4.3 Results from final stage optimization of the model: Major structure index from the fully integrated ICIO database

V. Unsolved issues and direction of future work

Conceptual issue

Most OECD individual country I-O tables are industry by industry symmetric tables (it is the only source with all country tables in the same format available to us), while global ICIO database are commodity group based tables, to solve this inconsistence, supply and use tables from each covered country should be used (which have both an industry and an commodity group dimensions and consistent with the UN standard of System of National Account (SNA)) and our reconciliation procedure in the second stage could be modified to jointly adjust individual countries' supply and use table to satisfy the global consistent condition, but leave the model in the first and final stages no change.

Classification issue

Detailed classification is desirable but not practical. Even at 2 digit ISIC level, Current OECD I-O tables already have number of country tables with many empty row and columns (**Lin Xin, please help count how many of zero row and columns in each of the 40 country tables, especially for non-OECD countries**), may be a more aggregated industry and commodity group classification are more feasible and with less development cost.

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.

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

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

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.

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

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.

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

Zhi Wang, 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.*

Canning, Patrick and Zhi Wang "A Flexible Modeling Framework to Estimate Interregional Trade Patterns and Input-Output Accounts." Policy Research Working Paper 2934, The World Bank, June 2004.

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.

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

Appendix A1. OECD I-O Database Industry Classification and Concordance with ISIC Rev.3

ISIC Rev.3 code	OECD I-O	Description
1+2+5	1	Agriculture, hunting, forestry and fishing
10+11+12	2	Mining and quarrying (energy)
13+14	3	Mining and quarrying (non-energy)
15+16	4	Food products, beverages and tobacco
17+18+19	5	Textiles, textile products, leather and footwear
20	6	Wood and products of wood and cork
21+22	7	Pulp, paper, paper products, printing and publishing
23	8	Coke, refined petroleum products and nuclear fuel
24ex2423	9	Chemicals excluding pharmaceuticals
2423	10	Pharmaceuticals
25	11	Rubber and plastics products
26	12	Other non-metallic mineral products
271+2731	13	Iron & steel
272+2732	14	Non-ferrous metals
28	15	Fabricated metal products, except machinery and equipment
29	16	Machinery and equipment, nec
30	17	Office, accounting and computing machinery
31	18	Electrical machinery and apparatus, nec
32	19	Radio, television and communication equipment
33	20	Medical, precision and optical instruments
34	21	Motor vehicles, trailers and semi-trailers
351	22	Building & repairing of ships and boats
353	23	Aircraft and spacecraft
352+359	24	Railroad equipment and transport equipment n.e.c.
36+37	25	Manufacturing nec; recycling (include Furniture)
401	26	Production, collection and distribution of electricity
402	27	Manufacture of gas distribution of gaseous fuels through mains
403	28	Steam and hot water supply
41	29	Collection, purification and distribution of water
45	30	Construction
50+51+52	31	Wholesale and retail trade; repairs
55	32	Hotels and restaurants
60	33	Land transport; transport via pipelines
61	34	Water transport
62	35	Air transport
63	36	Supporting & auxiliary transport activities; activities of travel agencies
64	37	Post and telecommunications
65+66+67	38	Finance and insurance
70	39	Real estate activities
71	40	Renting of machinery and equipment
72	41	Computer and related activities
73	42	Research and development

74	43	Other Business Activities
75	44	Public administration and defense ; compulsory social security
80	45	Education
85	46	Health and social work
90-93	47	Other community, social and personal services
95+99	48	Private households with employed persons & extra-territorial organizations & bodies

Appendix A2 Countries include in the database

ISO3	**Countries**	**ISO3**	**Countries**
AUS	Australia	LUX	Luxembourg
AUT	Austria	NLD	Netherlands
BEL	Belgium	NZL	New Zealand
CAN	Canada	NOR	Norway
CZE	Czech Republic	POL	Poland
DNK	Denmark	PRT	Portugal
FIN	Finland	SVK	Slovak Republic
FRA	France	ESP	Spain
DEU	Germany	SWE	Sweden
GRC	Greece	CHE	Switzerland
HUN	Hungary	GBR	United Kingdom
IRL	Ireland	USA	United State
ITA	Italy	CHN	China
JPN	Japan	MEX	Mexico
SVN	Slovenia	TWN	Chinese Taipei
EST	Estonia	RUS	Russia
ARG	Argentina	TUR	Turkey
BRA	Brazil	ZAF	South Africa
IND	India	IDN	Indonesia
HKG	Hong Kong	ROW	The rest of World

Appendix A3 Variables code and description of production database

Var. code	Definition
PRO	Production (Gross Output) at current prices
INI	Intermediate inputs at current prices
VAU	value added at current prices
LAB	labor compensation of employees
EXP	exports of goods at current prices
IMP	imports of goods at current prices
OTX	other taxes less subsidies on production
GOP	Gross operating surplus and mixed income

Notes:

1. Production, Gross Output (PRO) = Intermediate inputs (INI) + Value added at Basic Prices (VAU)
2. Value added (VAU) = Value added at factor costs (VAF) + [Other taxed on production – subsidies on production] (OTX)
3. Value added at factors costs (VAF) = Compensation of employees (LAB) + Gross Operating Surplus and mixed Income (GOP)
4. Gross Operating Surplus and mixed income (GOP) = Net Operating surplus and mixed income (NOP) + Consumption of fixed capital (CFC)

Appendix A4 Country data notes – original sectors have been split

1. China, Mainland (CHN):

I-O data
The following manufacture sectors are disaggregated by industrial firm survey information:

In 1997
IO.34 "Cultural goods" split into OECD I11 & I25;
IO.64 "Other general industrial machinery" split into OECD I15_16;
IO.81 "Cultural and office equipment" split into OECD I17 & I20;
IO.110 "Resident services" split into OECD I47_48;
IO.114 "Other social services" split into OECD I40, I41 &I43;
IO.121 "General technical services" split into OECD I43 & I47;
In 2002
IO.34 "Stationary and related products" split into OECD I11 & I25;
IO.63 "Metal casting, forging process" split into OECD I15_16;
IO.70 "Aircraft, spacecraft and other transport equipment" split into OECD I23_24;
IO.81 "Cultural and office equipment" split into OECD I17, I20;
IO.85 "Electricity, steam and hot water production and supply" split into OECD I26 & I28;
IO.115 "Resident and other personal services" split into OECD I47 & I48;

In 2007
IO.68 "Metal casting, forging process" split into OECD I15_16;
IO.72 "Manufacture of other special purpose machinery" split into OECD I16, I20;
IO.76 "aircraft and spacecraft and other transport equipment" split into OECD I23_24;

IO.89 "Manufacture of machinery for cultural activity & office work" split into OECD I17, I20;
IO.92 "Production and supply of electric power and heat power" split into OECD I26, I28;;
IO.124 "Service to households and domestic services" split into OECD I47, I48;

Production data

The production data of CHN is coming from China Statistic Yearbook & China Industry Economy Statistical Yearbook 1996-2008.

Because there are no details for some service sectors, which is corresponding to OECD I-O I40 "Renting of machinery and equipment"; I41 "Computer and related activities"; I42 "Research and development"; I44 "Public administration and defense, compulsory social security"; I47 "Other community, social and personal services" and I48 "Private households with employed persons, extra-territorial organizations & bodies". We disaggregated them by the proportion of I-O tables in the benchmark year (1997, 2002, and 2007).

2. United States (USA):

The following sectors are split by the OECD benchmark USA I-O information.

NAICS_213 "Support activities for mining" split into OECD I2_3;
NAICS_22 "Utilities" split into OECD I26_29;
NAICS_331 "Primary metal manufacturing" split into OECD I13_14;
NAICS_334 "Computer and electronic product manufacturing" split into OECD I17_19;
NAICS_3364/65/66/69 "Other transportation equipment manufacturing" split into OECD I22_24;
NAICS_511 "Publishing industries (split into software)" split into OECD I41, 43, 47;
NAICS_561 "Administrative and support services" split into OECD I43, 47;

3. Chinese Taipei (TWN):

The Variables (PRO, INI, VAU, GOP, LAB, OTX) in production data of TWN is come from "Directorate-General of Budget, Accounting & Statistics (DGBAS)".[8] In DGBAS Database, some sectors needed to be disaggregated into OECD I-O 48 sectors according to TWN classification comparison to ISIC Rev.3 and OECD. These sectors including:

a. "CK. Chemical Products Manufacturing" which is corresponding to OECD I-O I9-10;

b. "CP. Basic Metal Manufacturing" corresponding to OECD I-O I13-14;

c. "CW. Other Transport Equipment Manufacturing" corresponding to OECD I-O I22-24;

d. "EA. Water Supply" corresponding to OECD I-O I28-29;

e. "M. Professional, Scientific and Technical Services" corresponding to OECD I-O I40-42;

The disaggregated proportions for these sectors are coming from the benchmark year of TWN I-O tables. (1995 corresponding to IO 1996; 1997/1998 to IO 1999; 2000/2002 to IO 2001; 2003/2005 to IO 2004; 2005/2007 to IO 2006).

[8] (http://win.dgbas.gov.tw/dgbas03/bs7/sdds/english/calendar.htm)

4. The Rest of World (ROW)

I-O data

The detail of disaggregated sectors are as following:

I9 Chemicals excluding pharmaceuticals split into I10 and I11
I16 Machinery and equipment nec. split into I18
I17 Office accounting and computing machinery split into I19 and I20
I24 Railroad equipment and transport equipment split into I22 I23
I29 Collection purification and distribution of water split into I28
I32 Hotels and restaurants split into I47 I48
I33 Land transport split into I36
I39 Real estate activities (only include privities' household dwelling)
For sector I39, which included in other business service in GTAP, since no data could make the split, here, we just put the data of household dwelling into I39 and keep the Other Business services in GTAP to include OECD Io sector 40-43.
I40 Renting of machinery and equipment split into I41-43 and commercial real estate
I44 Public administration and defence split into I45 I46

Production data

The UN Statistic Division Database just include PRO, VAU, INI, OTX, GOP, LAB data for 87 additional countries except the countries covered in our ICIO database. Therefore, the gross industrial output is less than the real data, so we choose the PRO data from the *World Economic Outlook 2009*, IMF, which including 147 countries and more close to the real world. And then use the proportion between VAU to PRO, OTX to VAU, LAB to VAU, GOP to VAU from the UN 87 countries to estimate the VAU, LAB, GOP, OTX and INI for ROW.

5. Production data split from UN aggregated sectors:

Following country's production data are split by proportion from UN national income account data:

CAN and AUS 2006,

IDN, BRA, IND, RUS, TUR, ZAF, ARG: 1995 to 2006

Split sector	Aggregated sector in UN database
I2_3	"Mining and Quarrying"
I26_29	"Electricity, gas and water supply"
I31_32	"Wholesale retail trade, repair of motor vehicle, motorcycles, etc.; hotels and restaurants"
I33_37	"Transport, storage and communications"
I38_43	"Finance, Insurance, real estate and business services"
I45_47	"Education, health and social work, other community, social and personal services"