# A Model to Construct Time Series Databases for Global Trade, Production, and Consumption Linkages

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

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

*The views expressed in this presentation are solely those of the presenter. It is not meant to represent in anyway the views of the U.S. International Trade Commission or any of its individual Commissioners.

# Presentation Outline

- Objective and Motivations
- Basic Ideas of the adjustment framework
- Model Specification
- Data Sources and Model Implementation
- Preliminary Results
- Problems and Future Work

# Objectives

- Develop a formal model integrating individual countries' input-output tables with bilateral trade statistics to estimate an consistent annual inter-country input-output (ICIO) account for the world.

- Implement and test the model with real world data from 1995 to 2006. IO data from OECD, trade in goods from UN,  trade in services from IMF and OECD

# Motivations (1)

**Fill the void between resurgence in application of input-output tables and lack of consistent data**

- ***Analytical data source and accounting framework***
  - measurement of vertical specialization (**Hummels, Ishii, and Yi, 2001**);
  - development of value-chain in global production network (**Wang, Powers and Wei, 2009)**
  - the pattern of gross versus value-added trade (**Johnson and Noguera, 2009);**
  - Accounting intermediate trade flows **(Sébastien Miroudot, Rainer Lanz, and Alexandros Ragoussis, 2010);**
  - environmental analysis (**Davis and Caldeira, 2010);**
  - Collapse of world trade during recent global financial crisis (**Bems, Johnson, and Yi, 2010);**
  - the economic impact of global rebalancing (**Petri, 2010).**
- ***Statistical analysis tool, estimate GDP in consistence with SNA***
- **In contrast with this surge in analytical and statistical demand, there is a lack of consistent, time-series global I-O data sets**

# Motivation (2)

## Facilitate the research and analysis of global value chain and Dynamic AGE modeling

- We first constructed a single year global ICIO table based on version 7 GTAP database and processing trade information from China and Mexico. It covers 26 countries and 41 sectors and was used to support our initial global AGE modeling of processing trade and value-chain analysis efforts. The gross, intermediate and value-added bilateral trade flows and initial decomposition of value-added trade now is available for internal review and test use at ITC.

- We are also developing a time series global ICIO database, which integrates 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).

# Basic Ideas of the Procedure

- Initial estimates of the same economic variables from different sources
- A set of well defined consistency conditions and other constraints
- Reliability information on the initial estimates

# Problems of Proportional Adjustment in National Income Account

## Components of U.S. GDP, 2002, Billion U.S. Dollars

| GDI Components | Value | Value Share | Adjustment by proportion | CV |
|---|---|---|---|---|
| Compensation of employees | 5977.4 | 0.575 | 32.97 | 0.95 |
| Indirect businesss tax and nontax liability | 800.4 | 0.077 | 4.41 | 2.31 |
| Private consumption of fixed capital | 1163.9 | 0.112 | 6.42 | 3.56 |
| Nonfarm proprietors' income | 743.7 | 0.072 | 4.10 | 3.74 |
| Net interest | 684.2 | 0.066 | 3.77 | 5.86 |
| Corporate profits | 787.4 | 0.076 | 4.34 | 8.35 |
| Businesss transfer payments | 44.1 | 0.004 | 0.24 | 19.86 |
| Rental income of persons | 142.4 | 0.014 | 0.79 | 23.45 |
| Subsidies | 32.5 | 0.003 | 0.18 | 26.51 |
| Farm proprietors' income | 12.9 | 0.001 | 0.07 | 48.89 |
| Sum | 10388.9 | 1 | 57.30 | |
| Gross domestic product | 10446.2 | | | |
| Difference | 57.3 | | | |

# Problems of Proportional Adjustment in International Trade Statistics

## China & Hong Kong reported exports and partner reported imports, 2004, Million Dollars

| Country | China reported Exports to Partners | Hong Kong domestic exports to partner | China re-exports to partner via Hong Kong | Partners imports from China and Hong Kong | Statistical discrepancy % |
|---|---|---|---|---|---|
| Malta | 273 | 5 | 20 | 92 | -200.4 |
| Russia | 9,102 | 119 | 361 | 4,744 | -110.4 |
| Korea | 27,810 | 2,111 | 2,832 | 32,853 | -1.8 |
| Japan | 73,222 | 4,268 | 11,977 | 94,911 | 3.4 |

# A Three-Stage Optimization Procedure (1)

Decompose the adjustment process into three Stages

- **Model structure determined by available statistics**
- **Make full use of all available information**
- **Reduce model dimension**
- First Stage:
  - Reconcile total merchandise and service trade statistics with each country's reported total export to and import from the world
  - Constraint: global export supply plus international transportation margin equals global import demand for each goods and service group.

# A Three-Stage Optimization Procedure (2)

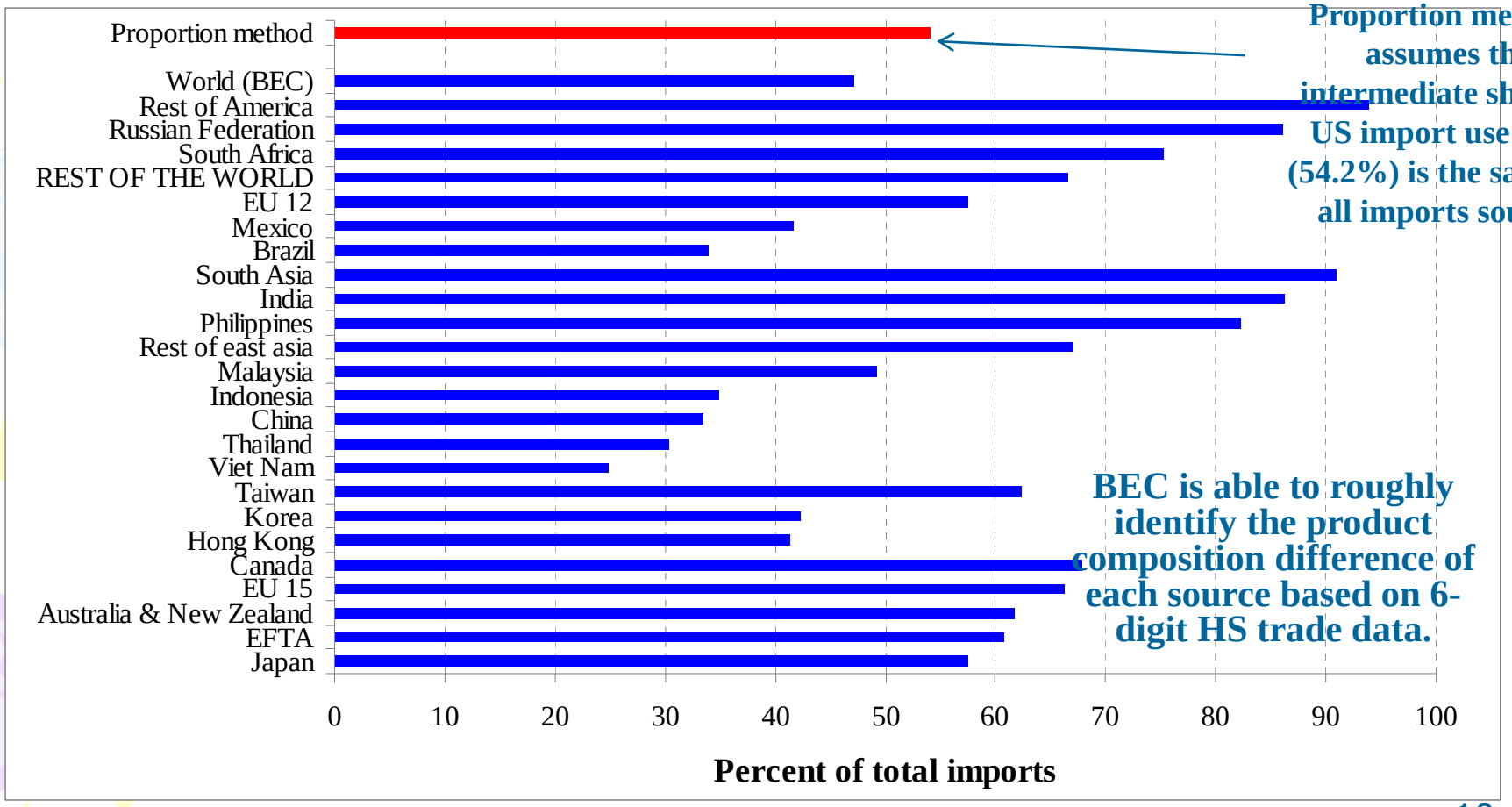
- Second Stages:
  - Reconcile each country's I-O table with the global consistent trade data from the first stage and fill missing I-O data between benchmark years for countries that do not have annual I-O statistics.
  - Constraints:
  - For each industry, total intermediate inputs purchased from all commodity groups and all sources (domestic and imported) plus value-added sum up to the industry's gross output;
  - 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
  - For each commodity group , imported intermediates used by all industries, imported final goods used by all users, and the amount of goods re-exported minus a re-exports mark-up, sum to total imports, which is fixed at the global consistent gross imports solved from first stage;
  - For each commodity group, domestic exports plus re-exports equals gross exports, which is fixed at the global consistent level solved from the first stage;
  - The sum of each type final demand by commodity groups plus a commodity tax equals the aggregate final demand categories in each country's GDP by expenditure account.

# A Three-Stage Optimization Procedure (3)

- Third Stage: Integrate individual country's I-O statistics with international bilateral trade statistics to produce a consistent annual global inter-country input-output table (ICIO)
  - Allocate each country's total export and imports in every commodity group into its trading partners based on UN COMTRADE (merchandise trade) and OECD (Services) bilateral trade data.
  - Allocate of each commodity group into to intermediate and final uses is based on the UN BEC method.
  - Each country's total exports and imports from the world from the first stage as controls
  - Bilateral flows at each trade route are controlled in the interval between mirrored trade statistics (both partner reported data) or with minimum deviation

# Why BEC is Better than Proportional Assumption

## Intermediates as % of U.S. Imported Electronic Machinery, 2004



# What BEC Can Help and What It Can't

- Intermediate goods identified from gross trade flows are the row sum of each block matrix $A_{rs}$ in the IO coefficient matrix A.
- End use classification such as BEC can help improve the accuracy of IO coefficients in IRIO table by giving better row total control for each block matrix in A.
- To allocate intermediate inputs to each industry after they enter the importing country, we still rely on proportion assumptions, which may distort the value-added distribution at industry level computed from IRIO table. More reliable data collected by national and international agencies are needed to overcome this limitation.

# Reliability of Reported Trade Statistics

- Mirror trade statistics are the major data source to estimate the reliability weights

- Econometric analysis of discrepancies between the two reported trade data of the same trade flows provide estimates of data reliability

# Auto regression with dummy variables

$$e_{it} = a_i\, e_{it-1} + b_i^0 + \sum_{k=1}^{n} b_i^k D_t^k + \mu_{it}$$

**e: mirror trade statistics discrepancies**
**b: symmetric bias**
**D: dummy variables. Represent events have a significant impact on the reporting practice in the two reporting countries**

**the variance:**

$$V(e_{it}) = V(\mu) \Big/ (1 - a_i^2)$$

# Route Specific Reliability Indexes

- **The initial trade flow estimates of the model can be used to construct such indexes:**

$$REL_{it}^{sr} = \frac{|M_{it}^{sr} - E_{it}^{sr}|}{0.5(M_{it}^{sr} + E_{it}^{sr})} = 2\frac{|M_{it}^{sr} - E_{it}^{sr}|}{(M_{it}^{sr} + E_{it}^{sr})}$$

- **It has a value between 0 and 2. A smaller value indicates the initial estimates are relatively more reliable for the associated trade route than other routes.**

# Reporter Specific Reliability Indexes

- **It is the share of accurately reported trade in total trade (less than 20 percent discrepancies in mirrored data)**
- **All bilateral trade data in the world need to be used to construct the reporter specific reliability indexes**
- It has a value between 0 and 1. A large value indicates the initial estimates reported by the country are relatively more reliable for its reported exports or imports than other reporters
- The weights in the objective function of the model can be assigned by multiplying one minus the indexes by their corresponding initial values

# The Adjustment Problems

Adjust a given set of initial trade and IO statistics according to an objective function to satisfy accounting and consistence constraints. Such as

$$\text{Min} \quad S = \left\{ \begin{array}{l} \sum_{s=1}^{g}\sum_{i=1}^{n} \frac{(WX_{it}^{s} - WX0_{it}^{s})^2}{(1-RIX_{i}^{s})WX0_{it}^{s}} + \sum_{r=1}^{g}\sum_{i=1}^{n} \frac{(WM_{it}^{r} - WM0_{it}^{r})^2}{(1-RIM_{i}^{r})WM0_{it}^{r}} \\ +100\left(\sum_{s=1}^{g} \frac{(WTX_{t}^{s} - WTX0_{t}^{s})^2}{(1-RIX^{s})WTX0_{t}^{s}} + \sum_{r=1}^{g} \frac{(WTM_{t}^{r} - WTM0_{t}^{r})^2}{(1-RIM^{r})WTM0_{t}^{r}}\right) \\ + \sum_{i=1}^{n} \frac{(CIF_{it} - CIF0_{it})^2}{(1-RIM_{i})CIF0_{it}} + \sum_{r=1}^{g} \frac{(CIF_{t}^{r} - CIF0_{t}^{r})^2}{(1-RIM^{r})CIF_{t}^{r}} \end{array} \right.$$

# Theoretical Properties

- By imposing valid binding constraints,  the optimization procedure will definitely improve, or at least not worsen, the initial statistics estimates.
- In all but the trivial case, posterior estimates derived from entropy or quadratic loss function will always be closer to the unknown, true values than the associated initial statistics(Harrigan,1990)
- Statistical interpretations underlying the model differ when different reliability weights are used
- The choice of weights in the objective function has a large impact on the estimation results. The model uses these weights to determine by how much an initial estimate may be changed. The larger the weights, the smaller its contribution to the objective function, and hence the lesser the penalty for each adjusted variable to move away from its initial value (only the relative, not the absolute size of the weights affects the solution).

# Why Adjusted Estimates are Better?

- **D0**: Initial estimates
- **W** variance matrix of initial estimates ,
- **A** coefficient matrix of all linear constraints $AD^* = 0$
- The BLUE :

$$\mathbf{D}^* = [\mathbf{I} - \mathbf{W}\mathbf{A}^T(\mathbf{A}\mathbf{W}\mathbf{A}^T)^{-1}\mathbf{A}]$$

- $\mathbf{D}^*$ will never be worse than D0 with equal or smaller variance

$$Var(D^*) = [\mathbf{I} - \mathbf{W}\mathbf{A}^T(\mathbf{A}\mathbf{W}\mathbf{A}^T)^{-1}\mathbf{A}]\mathbf{W} = \mathbf{W} - \mathbf{W}\mathbf{A}^T(\mathbf{A}\mathbf{W}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{W}$$

# Empirical Advantages

- **Completeness**
  - Complete use of all information from official statistics
- **Flexibility**
  - Additional constraints can be easily imposed to allow, for example, upper and lower bonds to be placed on unknown elements (this is very common in mirrored trade statistics), 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.
- **Incorporation of data reliabilities in a systemic way**
  - Using properly selected reliability weights, the optimal solution yields estimates deviate less from the initial estimates with higher degrees of reliability than for those with lower degrees of reliability. The reliability weights can be put into an array that has the same dimension and structure as the initial estimates. Therefore, considerable amount of information regarding to the quality of the initial statistics could be incorporated into the data reconciliation process.

# Data Source (1)

## Trade Data

- There were significant differences in values among different sources (UN National Accounts, UNCTAD, IMF's IFS and BOP database, WITS- COMTRADE, and the OECD trade statistic)
- We use country total gross exports to and imports from the world from the UN *National Account Estimates* as aggregate controls
- WITS-COMTRADE for country sector total and detailed bilateral merchandise trade (aggregate from HS-6)
- IMF BOP for country sector total and OECD statistics for detailed bilateral services trade.
- All trade data are from 1995-2006

# Total World Trade from UN National Account is Very Close to Data Reported by Other Sources

| 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 | | | | | | | | | | | | | | |

# Total Service Trade from UN National Account is Very Close to Data Reported by Other Sources

**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 |

# Data Source (2)

## IO Data

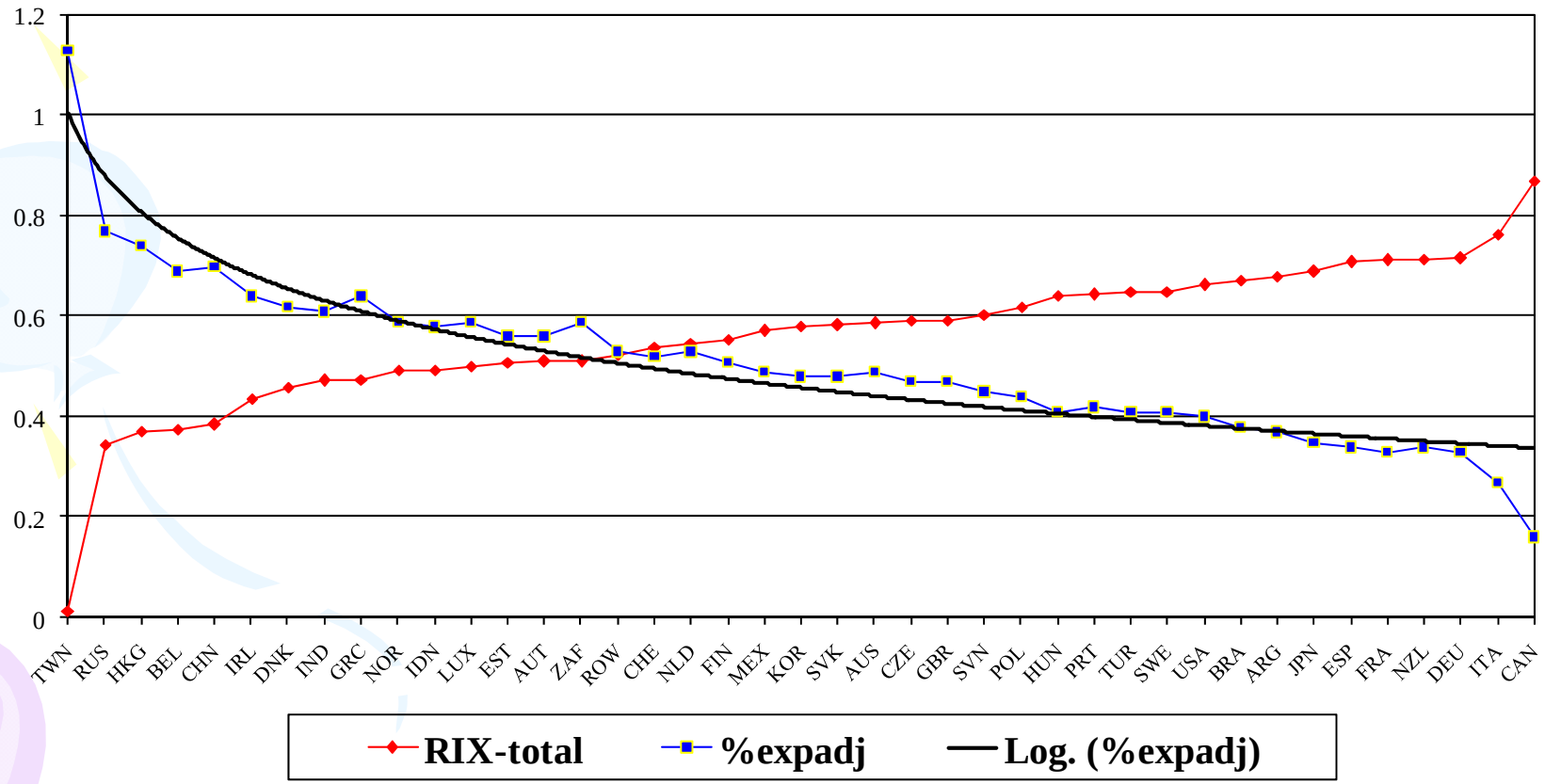
- 2010 version OECD STAN individual country IO tables, which include 39 countries for 3 benchmark years (around 1995,2000 and 2005)
- Additional year IO tables for 6 EU countries: DNK, FIN, FRA, DEU, NLD, SVN
- 1997-2006 U.S. annual IO tables from BEA
- China 1997, 2002 and 2007 Bench mark IO tables from NBS
- Taiwan 1996, 2001 and 2006 Benchmark and 1999, 2004 extension IO table from Taiwan Statistical agency
- Cooperate with Chinese University of Hong Kong, compile Hong Kong annual IO tables from 1995 to 2007 under the assistance of Hong Kong Census and Statistical Department
- Construct 1997, 2001, 2004 IO tables for the rest of the world from version 5,6,7 GTAP database

# Mean Absolute Percentage Adjustment

- Measurement of adjustment from official statistics: only the proportionate deviation and not the absolute deviation that matters

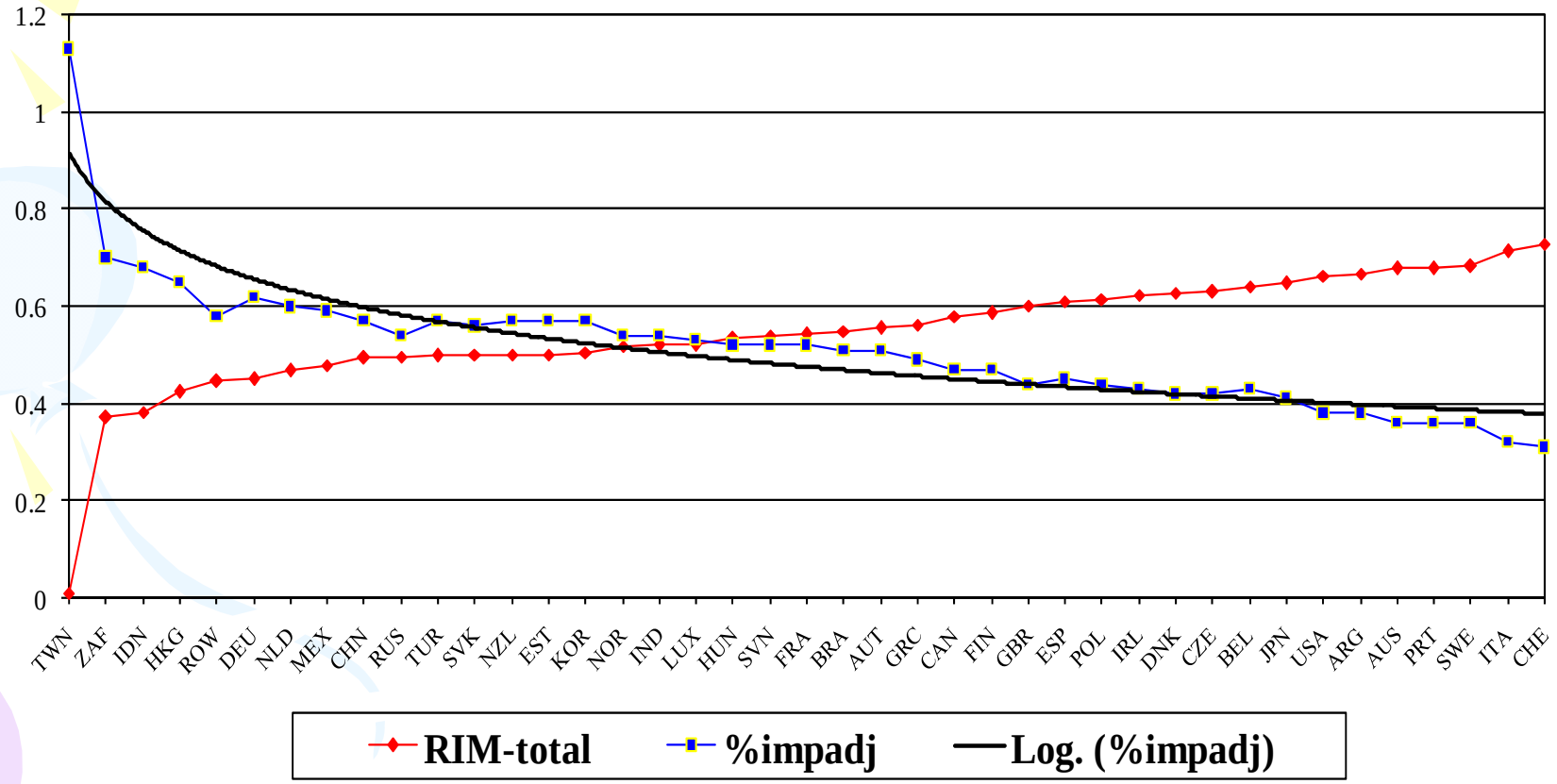
$$MAPA^{C} = \frac{100 \bullet \sum_{t=1995}^{2006} \sum_{r=1}^{g} | \overline{X}_{rt} - X0_{rt} |}{\sum_{t=1995}^{2006} \sum_{r=1}^{g} X0_{rt}}$$

# Reporter Reliability and Mean Absolute Percentage Adjustment of UN country total goods exports, 1995 -2006



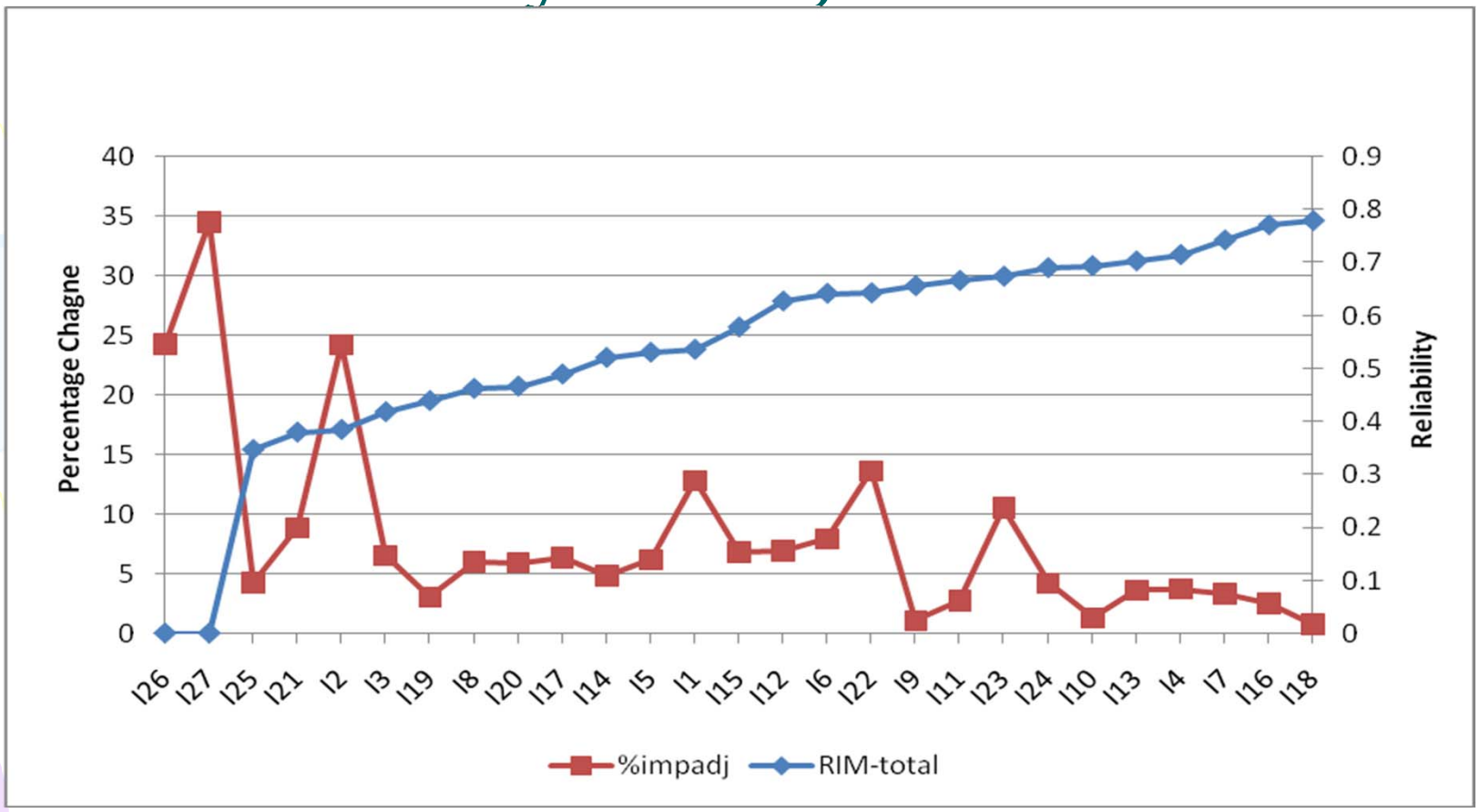

**More reliable data get less adjustment**

# Reporter Reliability and Mean Absolute Percentage Adjustment of UN country total goods imports, 1995 -2006



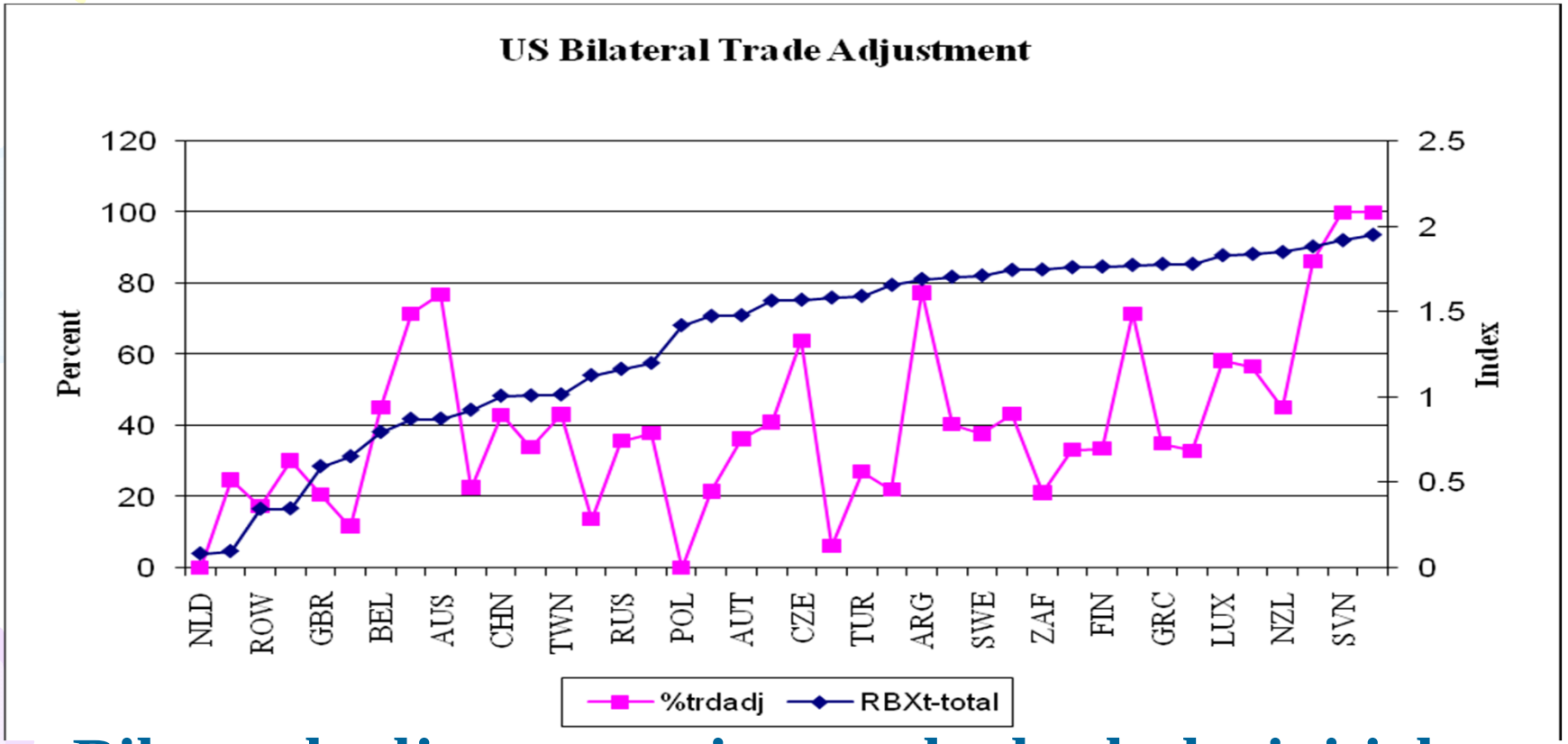

## More reliable data get less adjustment

# Reporter Reliability and Mean Absolute Percentage Adjustment of World Goods Trade by sectors, 1995 -2006



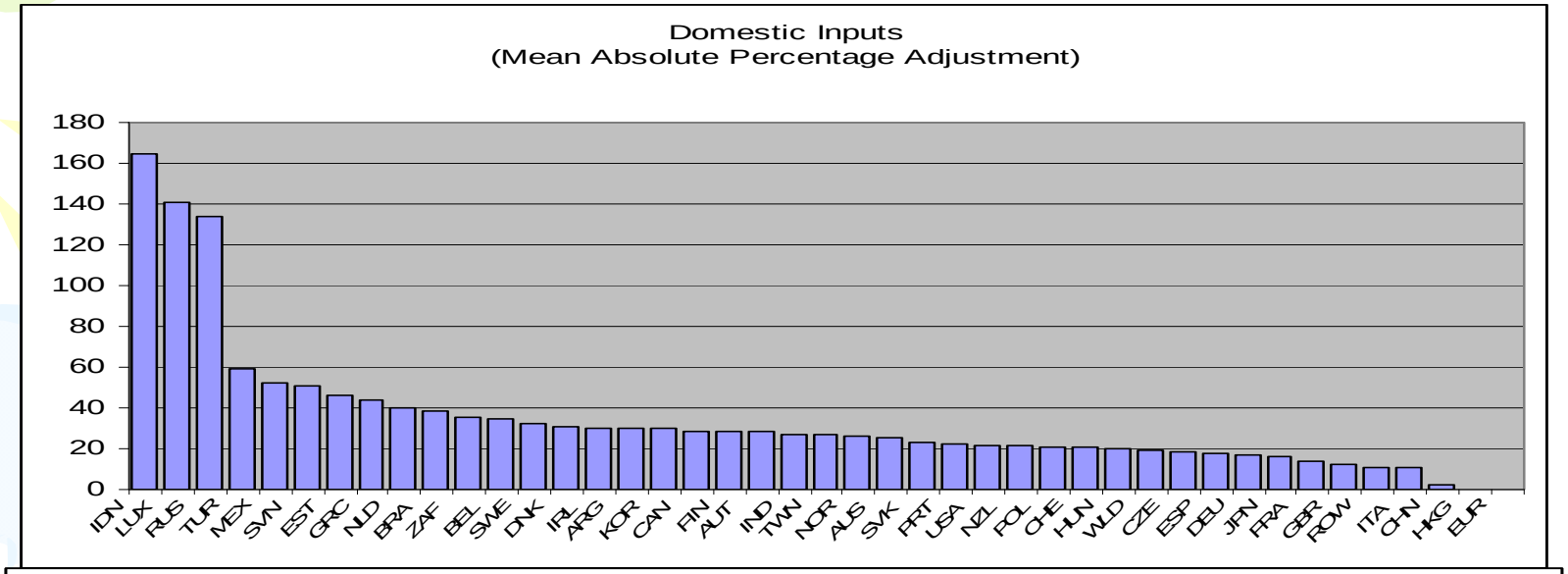

**Sector adjustments impact by both the initial discrepancies and data reliability**

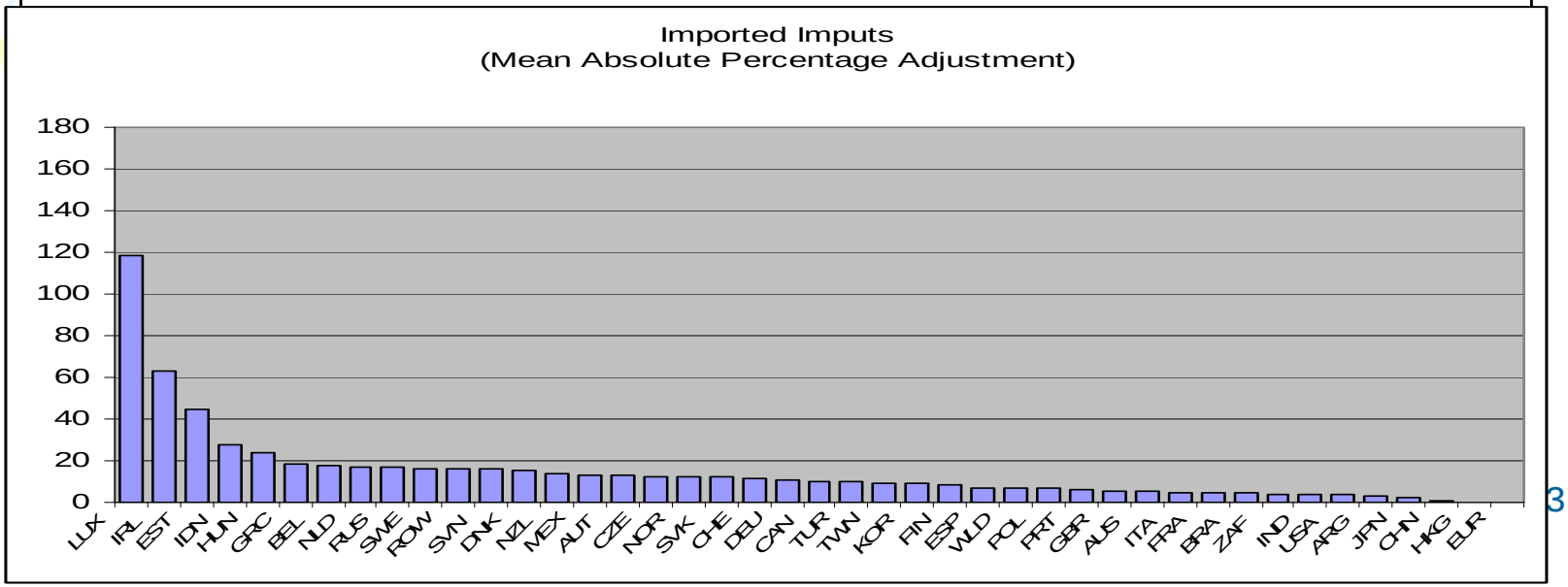
# Route Reliability and Mean Absolute Percentage Adjustment of U.S. Bilateral Goods Trade, 1995 -2006



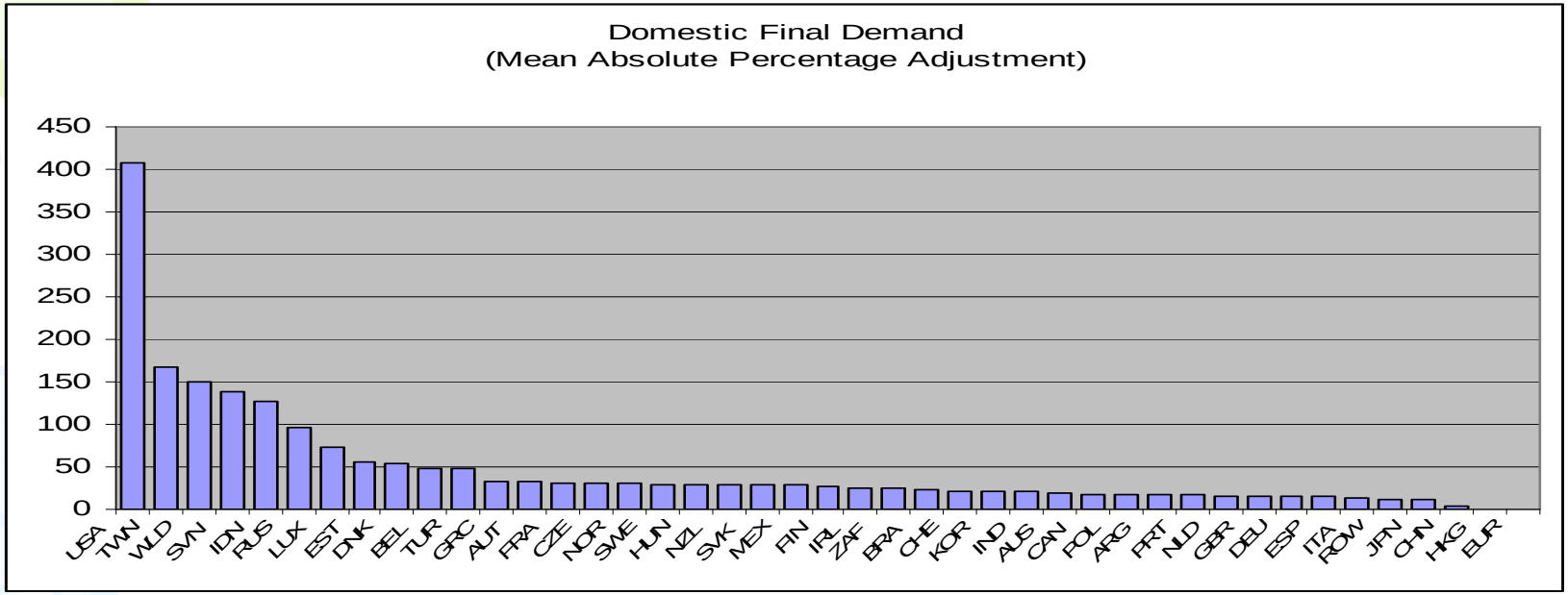

**Bilateral adjustments impact by both the initial discrepancies and data reliability**

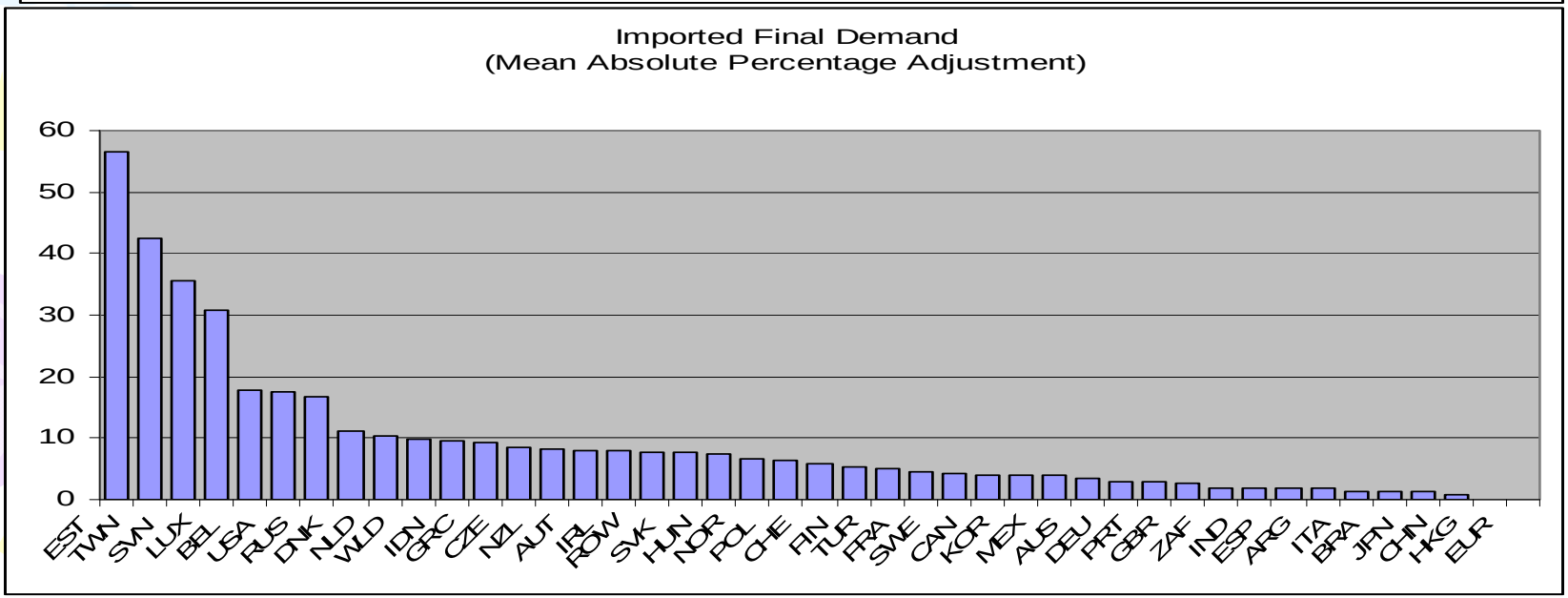
# The degree of adjustment help identify potential problem in the raw data



# The degree of adjustment help identify potential problem in the raw data



# Conclusions

- We developed a three-stage mathematical programming model to reconcile detailed bilateral goods and services trade statistics with individual country's input-output tables to produce an integrate ICIO database for the world.
- We document the major data sources for such data reconciliation excise and their pro and cons.
- Test the model using OECD IO tables and UN trade data, preliminary results are encouraging and show that the model is feasible and may have great potential in the estimation of an integrated world IO account.

# Unsolved Issues and Future Work

The model is still in its early stage of development, there are many issues remain to be solved

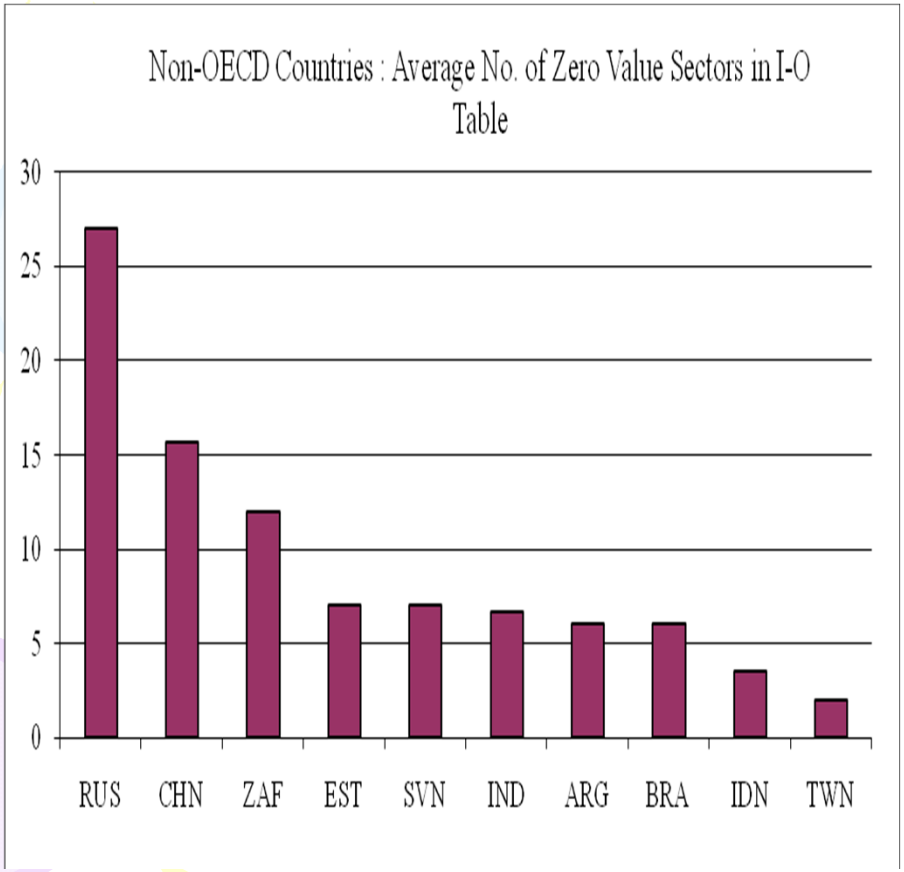
- *Symmetric IO tables or supply and use tables?*
- *More or less industry and commodity groups?*
- *Reliability index for service trade and IO statistics*
- *Re-exports and re-exports-make-ups*
- *Purchases on the domestic territory by non-residents and direct purchases abroad by residents*

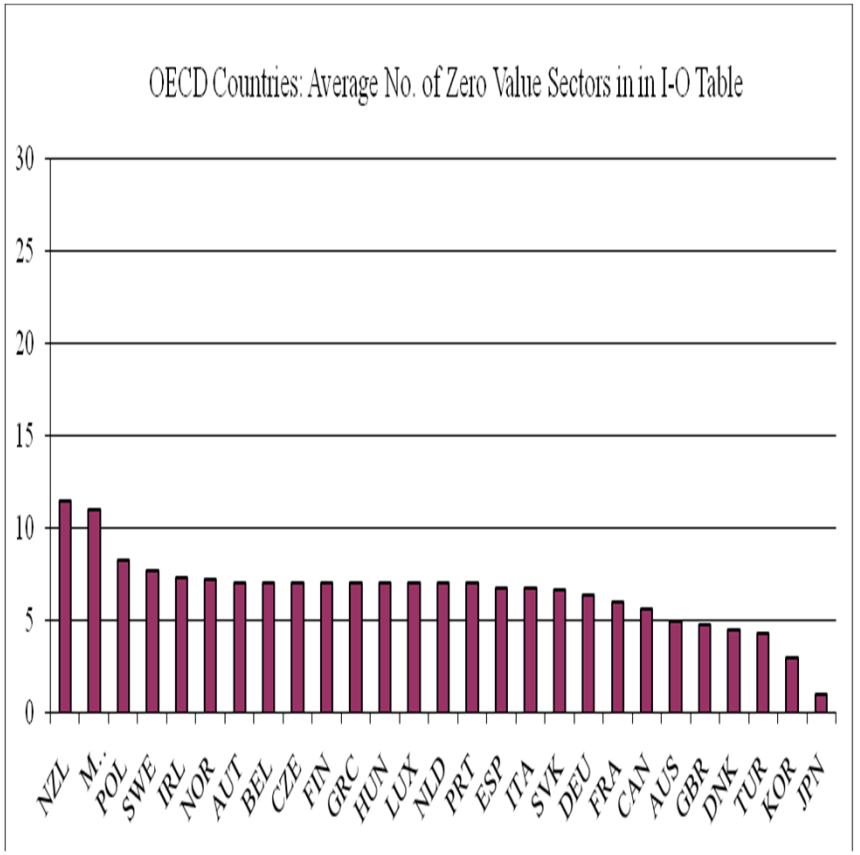
# Why Supply and Use tables are Better?

- Easy benchmark to National Account statistics
- Has both industry and commodity dimension that consistent to the UN standard of system of national account
- More close to actual data collected by NSI and avoid errors inherent in the assumptions imposed when transferring SUTs to symmetric IO tables
- NSI put major part of their resources into compile national account statistics to obtain better GDP estimates

# More disaggregation may not always better when detailed data are not available

## Missing sectors in current version OECD IO tables



# Pros in WIOD Data Set

- Time-series benchmarked on National Account data
- Construct net tax, trade and transport margin matrices from SUTs at purchasers' prices to basic prices
- Improved allocation of imports to end use category
- From current price to constant price tables, both national and international (PPP) deflator will be used
- Closely linked with EU KLEMS and World KLEMS to obtain better and detailed capital types and labor skill levels breakdown, could widely used by the entire economic profession, not only IO and CGE modelers

# Cons in WIOD Data Set

- Only use import statistics as trade data, abandon all information in export statistics
- Exports to Rest of the world are calculated as residual and can become negative for some products
- No reconciliation procedure based on data reliability will be used, statistics from any official source are treated equally (mix up high quality data with low quality data)
- Detailed coverage on the 27 EU member countries, but include only less than 10 developing countries

# Thanks for your attention!

- Other comments/questions?

- My E-mail Address:
  Zhi.Wang@USITC.GOV