A Three Stage Reconciliation Method to Construct Time Series International Input-output Database (Work in Progress)

Nadim Ahmad and Norihiko Yamano, OECD

Zhi Wang
*United States International Trade Commission**

ASSA/AEA Annual Meeting
San Diego, January 4, 2013

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

Presentation Outline

- Background and Objective
- Basic Ideas of the adjustment framework
  - Use reporter relative reliability index to control the reconciliation process
  - Use constraints bring additional information into the data set during reconciliation
- Data Sources to Test the Method
- Preliminary Results
- Direction of future work

Why we need time series Global IO tables to estimate trade in value-added

- It is a well-known fact that national income accounts record domestic output (transactions) in value added terms while standard trade statistics record trade in gross terms. This shortcoming in official trade statistics and their inconsistency with the system of national account standards has long been recognized by both economists and economic policymakers.
- An accurate assessment of trade in value added has to go beyond a single country's effort, as it requires information on cross-border input-output relationships.
- Direct measurement of value-added trade is extremely difficult, primarily because the information is not available in business record-keeping systems. The most feasible and most promising approaches to developing comprehensive and consistent value-added trade measures that go beyond case studies of individual high-profile products (such as the iPod) have to involve the use of Global Input-Output tables.

Why we need a model to reconcile official statistics and balance Global IO tables

- International trade statistics do not balance at the global level; giving rise to the humorous anecdote of Earth trading with Mars or the Moon. At the national level this can generally be ignored; the perspective being that the inconsistencies are in other country's accounts. But when considering global accounts, these inconsistencies create significant problems
- There are a numbers of attempts to compile global IO tables in recent years (Lenzen et al. (2012), Wang(2011), Wang et al. (2012), Johnson and Noguera (2012)), including the OECD, who plan to launch a database of TIVA indicators on January 19, 2013, which has led to important improvements in the qualities of the estimated global IO tables.
- One important resent development is the publicly available WIOD project/database funded by the EU

Common features in recent attempts to compile global IO tables

- Benchmark to official national accounts estimates of output and final consumption (as not all countries' supply-use tables are necessarily benchmarked to, nor revised in line with, their GDP by expenditure account).
- Assumptions used to allocate imports to users have moved away from the traditional crude "proportionality" assumption and now capture heterogeneities in imports from different sources based on the end use category that is available in trade statistic (UN Broad Economic Category classification);
- A recognition that shares rather than values per se are what matter in official bilateral trade statistics.

Pros in WIOD Data Set

- Time-series benchmarked on National Account data
- the use of supply-use tables as the starting point to integrate trade statistics and derived the final symmetric world IO table
- Construct net tax, trade and transport margin matrices from SUTs at purchasers' prices to basic prices
- 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 is calculated as residual and can become negative for some products
- International transportation margins were assumed produced in the rest of the world by the “Panama assumption” and not linked back to the world economy
- Detailed coverage on the 27 EU member countries, but include only less than 10 developing countries
- No reconciliation procedure based on data reliability are used, statistics from any official source are treated equally, thus mix up high quality data with low quality data

Objectives

- Develop a formal procedure to integrate individual countries' Supply and Use Tables (SUTs) with official trade statistics to estimate a consistent annual inter-country input-output (ICIO) account for the world that benchmarked by official National Account Statistics.
- Implement and test the procedure with real world data from 1995 to 2009. SUTs data from WIOD. National accounts and trade statistics from UN and OECD.
- How much adjustment on individual country's GDP is needed to eliminate "exports to the Moon"? Whether it is within the range that accepted by NSI.

Basic Ideas of the Reconciliation Methods

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

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 Reconciliation Procedure

- Decompose the reconciliation process into three Stages
- Model structure determined by available statistics
- Make full use of all available official statistics and related information
- Reduce model dimension

Stage 1

- Reconcile goods and services trade by products reported in each country's SUTs and total trade statistics in goods and services reported in each countries GDP by expenditure account to produce a set of country and commodity level total exports and imports and CIF margins that satisfy：
  - Global export supply plus international transportation margin equals global import demand for each goods and service group.
  - The use of international transportation services equals net exports of such services from all countries

Stage 2

Reconcile each country's SUTs with the global consistent exports to and import from the world as well as CIF margins from the first stage and fill missing SUT data between benchmark years for countries that do not have annual SUT statistics subject to following Constraints:

- For each industry, total intermediate inputs purchased from all products and all sources (domestic and imported) plus value-added sum up to the industry's gross output;
- For each product, the amount sold to all industries as domestic intermediate inputs plus the amount sold to final users as domestic final goods and services plus the amount of domestic exports equal total product output produced by the industries
- For each product, imported intermediates used by all industries, imported final goods used by all users, and the amount of goods re-exported, sum to total imports, which is fixed at the global consistent gross import demand solved from stage 1;
- For each product, domestic exports plus re-exports equals gross exports, which is fixed at the global consistent level solved from the stage 1;
- The sum of each type final demand by products plus net tax equals the aggregate final demand categories in each country's GDP by expenditure account.

Final Stage

- **Integrate individual country's SUTs with bilateral trade statistics to produce consistent annual global SUTs**
  - Allocate each country's total export and imports by product into its trading partners based on share computed from official bilateral trade statistics.
  - Separate bilateral trade flows of each product into to intermediate and final uses based on detailed trade statistics and OECD extended UN BEC classification.
  - Each country's total exports and imports from the world from stage 1 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. All countries' official trade statistics are fully used.
  - Global SUTs is transferred to industry by industry ICIO tables using "Model D" discussed in Eurostat (2008) similar to WIOD.

The Reconciliation Method: Objective Function and Reliability Weights

Adjust a given set of initial trade and SUT statistics according to an penalty function to satisfy global accounting and consistence constraints with minimum deviation from both official SUT and bilateral trade statistics.

$$\text{Min } S = \frac{1}{2}\left\{\sum_{s=1}^{G}\sum_{r=1}^{G}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\left(z_{ij}^{sr}-\bar{z}_{ij}^{sr}\right)^2}{wz_{ij}^{sr}}+\sum_{s=1}^{G}\sum_{r=1}^{G}\sum_{i=1}^{N}\sum_{k=1}^{K}\frac{\left(v_{ik}^{sr}-\bar{v}_{ik}^{sr}\right)^2}{wv_{ik}^{sr}}\right\}$$

Full Use of Mirrored Bilateral Trade Statistics: Reliability of reported trade statistics

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

- An indicator of reporter reliability is a measure of how consistency a country reports its trade statistics relative to all its trading partners. It should able to catch the strength and weakness of a country's ability to consistently report its trade for each end use categories in different commodities

Full Use of Mirrored Bilateral Trade Statistics: Estimate reporter specific reliability indexes

- **It is the share of accurately reported trade in total trade for a special end use category in a particular sector (less than 20 percent discrepancies in mirrored data)**
- **All available bilateral trade data in the world were 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
- It will encourage the model to adjust those unreliable initial data more than those reliable ones in the reconciliation process.

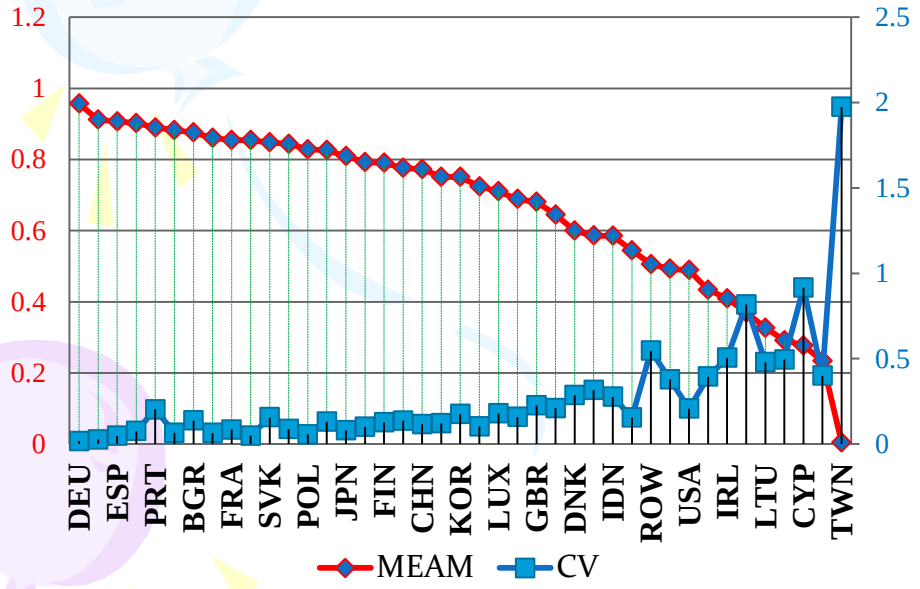
Average Exporter Relative Reliability Index 1995-2007, China

Commodity	MEAN	CV	MIN	MAX
Food and beverages (15)	0.81	0.06	0.71	0.88
Chemicals (24)	0.78	0.12	0.63	0.88
Basic metals (27)	0.72	0.17	0.46	0.89
Wood and products (20)	0.60	0.39	0.29	0.87
Paper and paper products (21)	0.53	0.31	0.22	0.80
Wearing apparel (18)	0.24	0.56	0.04	0.42
Rubber and plastic products (25)	0.14	0.65	0.07	0.39
Auto and Parts (34)	0.09	1.08	0.02	0.36
Leather products (19)	0.09	0.24	0.05	0.14
Electrical machinery (31)	0.07	1.09	0.03	0.33

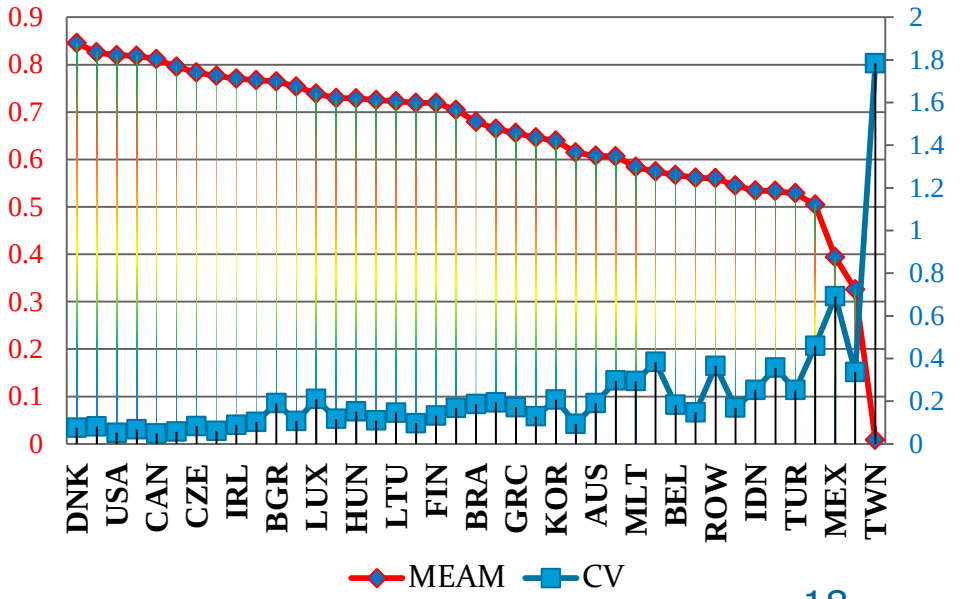
Average Importer Relative Reliability Index 1995-2007, The United States

Commodity	MEAN	CV	MIN	MAX
Auto and Parts (34)	0.95	0.02	0.92	0.97
Wood products (20)	0.92	0.05	0.83	0.97
Machinery and equipment (29)	0.91	0.06	0.77	0.97
Paper and paper products (21)	0.91	0.06	0.79	0.95
Food and beverages (15)	0.85	0.04	0.80	0.90
Textiles (17)	0.55	0.19	0.39	0.71
Wearing apparel (18)	0.54	0.10	0.43	0.61
Tobacco products (16)	0.50	0.34	0.17	0.71
Leather products (19)	0.30	0.34	0.16	0.48
Printed and recorded matter (22)	0.16	0.60	0.05	0.40

Average Exporter Relative Reliability Index Basic Metal (ISIC 27), 1995-2007



Average Importer Relative Reliability Index Basic Metal (ISIC 27), 1995-2007



Full Use of Mirrored Bilateral Trade Statistics: Initial value and constraints

- **Combine mirror trade data into initial estimates using reliability indexes as weights**

relsh(c,s,r,t,”exp”) = RIX(c,s,t)/(RIX(c,s,t)+RIM(c,r,t);

relsh(cc,s,r,t,"imp") = RIM(c,r,t)/(RIX(c,s,t)+RIM(c,r,t);

tfl.l(c,s,r,t) = relsh(c,s,r,t,"exp")*trdx(c,s,r,t)+relsh(c,s,r,t,"imp")*trdm(c,s,r,t);

- **Constraints that contain solution in a reasonable range**

MINEQtrd(c,s,r,t)$(ord(s) ne ord(r))..
tfl(c,s,r,t) + minadjtrd(c,s,r,t)=G= MIN(trdx(c,s,r,t),trdm(c,s,r,t));

MAXEQtrd(c,s,r,t)$(ord(s) ne ord(r))..
tfl(c,s,r,t) - minadjtrd(c,s,r,t)=L= Max(trdx(c,s,r,t),trdm(c,s,r,t));

The Adjustment Method:

Sufficient constraints with meaningful information

--- GlobalSUT.gms(6844246) 17459 Mb
--- LOOPS tsl = 2008
--- 5,043,499 rows 8,292,779 columns 25,320,581 non-zeroes
--- 5,842,171 nl-code 1,861,041 nl-non-zeroes
--- GlobalSUT.gms(6844246) 17408 Mb
--- Executing CPLEX: elapsed 16:38:57.598

S O L V E S U M M A R Y

MODEL BAL OBJECTIVE SI
TYPE QCP DIRECTION MINIMIZE
SOLVER CPLEX FROM LINE 6844246

**** SOLVER STATUS 1 Normal Completion
**** MODEL STATUS 1 Optimal
**** OBJECTIVE VALUE 2493580.7377

Data Source

- We use country total goods/services exports to/imports from the world (both in FOB) from official *National Account Estimates (*GDP by expenditure*)* as aggregate controls.
- 40 individual country annual SUTs (1995 to 2009) from WIOD, SUT for Rest of the World estimated by OECD.
- Bilateral goods trade statistics from OECD BTDIxE database; bilateral service trade data from OECD EBTSI dataset.
- There were significant differences in values among different sources (UNCTAD, IMF's IFS and BOP database, WITS-COMTRADE, and the OECD ITCS), we choose control from official National Account statistics has good reasons.

Total goods and services trade of major trading countries from NA is very close to data reported by other sources

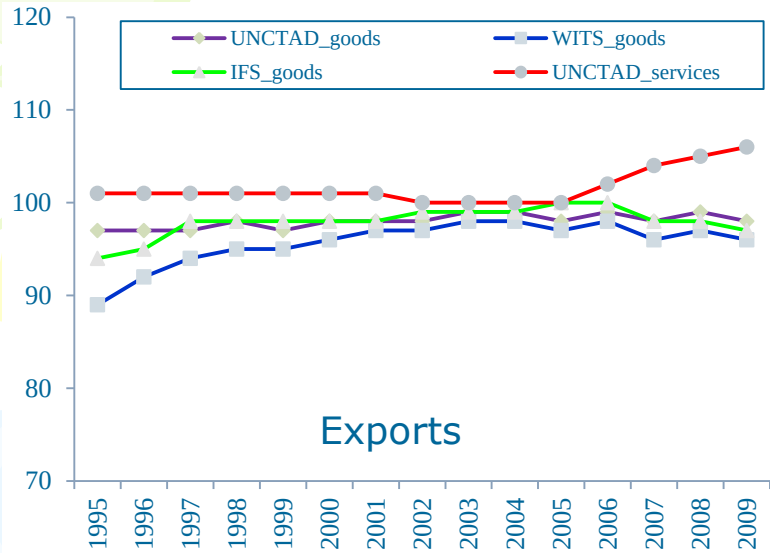
as percent of NA data by sources

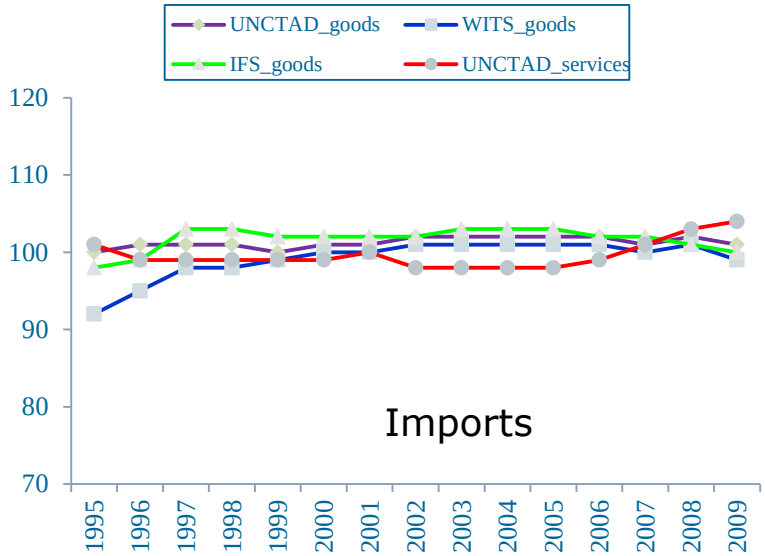
Reporter	Source	Exports								Imports							
		1995	1997	1999	2001	2003	2005	2007	2009	1995	1997	1999	2001	2003	2005	2007	2009
China	UNCTAD	102	100	100	100	100	100	100	100	107	104	104	105	105	105	106	105
	WITS	102	100	100	100	100	100	100	100	107	104	104	105	105	105	106	105
	IFS	102	100	100	100	100	100	100	100	107	104	104	105	105	105	106	105
Japan	UNCTAD	104	103	104	105	105	105	106	107	113	110	110	111	111	109	109	110
	WITS	104	103	104	105	105	105	106	107	113	110	110	111	111	109	109	110
	IFS	104	103	104	105	105	105	104	106	113	110	111	111	111	109	108	110
Germany	UNCTAD	100	100	100	100	100	99	99	98	101	102	100	101	100	99	99	98
	WITS	100	100	100	100	100	99	99	97	101	102	100	101	100	99	99	97
	IFS	100	100	100	100	100	99	98	97	101	102	100	101	100	99	99	98
United States	UNCTAD	100	100	99	100	100	100	101	99	102	101	101	101	102	102	102	101
	WITS	100	100	99	100	100	100	101	99	102	101	101	101	102	102	102	100
	IFS	100	100	100	100	100	100	101	99	102	102	101	101	102	102	102	101

Reporter	Source	Exports								Imports							
		1995	1997	1999	2001	2003	2005	2007	2009	1995	1997	1999	2001	2003	2005	2007	2009
China	UNCTAD	88	100	100	100	100	100	100	100	89	100	100	100	100	100	100	100
	BOP	88	100	100	100	100	100	100	98	89	100	100	100	100	100	100	100
	OECD	100	100	100	100	100	100	100	100*	100	100	100	100	100	100	100	100*
Japan	UNCTAD	124	128	130	131	130	131	137	138	111	116	117	117	118	117	120	123
	BOP	124	128	130	131	130	131	137	139	111	116	117	117	118	117	120	123
	OECD	124	128	130	131	122	126	134	131*	110	111	114	116	115	117	115	116*
Germany	UNCTAD	100	99	95	99	104	105	103	102	100	100	98	98	101	101	101	102
	BOP	100	99	95	100	104	105	106	109	100	100	101	101	101	101	101	102
	OECD	100	99	96	100	104	104	101	96*	101	102	102	101	101	99	99	101*
United States	UNCTAD	95	95	95	94	95	96	96	98	97	97	97	96	98	98	98	101
	BOP	95	95	95	94	95	96	98	98	97	97	97	96	98	98	97	99
	OECD	96	96	97	97	97	98	101	100*	97	97	97	96	98	99	99	99*

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

Total world trade from Official National Account is very close to data reported by other sources





Total world imports in NA are measured by FOB and very close to total world exports Share of imports over exports by data sources

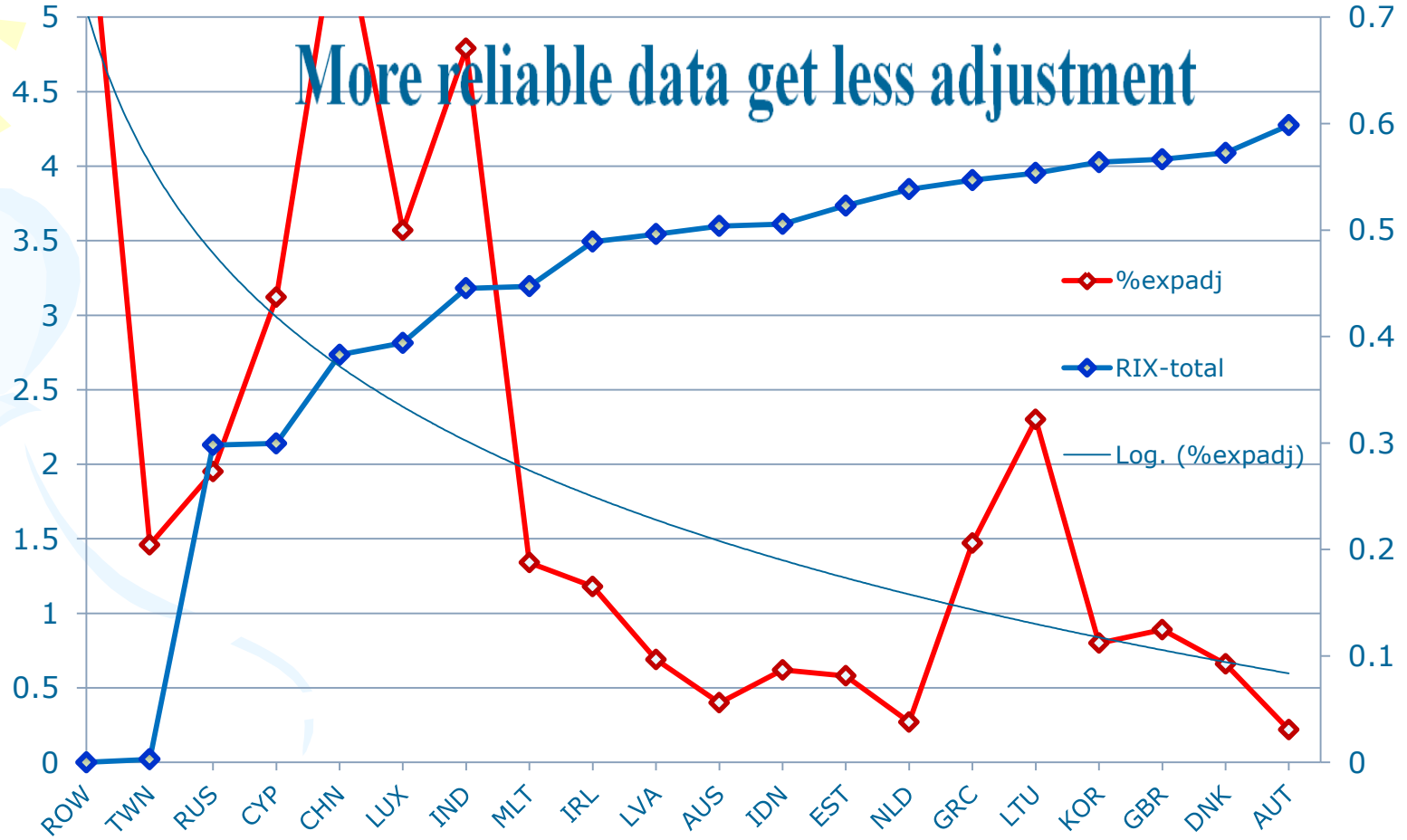
	Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
UN (total)		**98**	**98**	**98**	**99**	**99**	**100**	**100**	**99**	**99**	**99**	**99**	**98**	**98**	**98**	**98**
UN	Goods	98	98	97	98	99	99	100	99	99	99	99	99	98	99	99
UNCTAD		101	101	101	101	102	103	103	102	102	102	103	101	101	102	101
IFS		102	103	102	102	103	103	104	103	103	103	102	101	102	102	101
BOP		97	98	99	100	101	103	103	103	103	103	103	103	103	101	101
BOP2		97	98	98	97	99	101	100	99	99	99	100	100	100	101	100
WITS		101	102	102	101	103	103	103	102	102	103	103	102	103	103	102
UN	Services	100	100	100	100	100	100	100	100	100	98	98	96	96	96	97
UNCTAD		100	99	98	98	98	98	99	98	98	96	96	94	93	93	94
BOP		100	98	97	97	97	96	97	96	96	94	92	90	88	94	94

Mean Absolute Percentage Adjustment

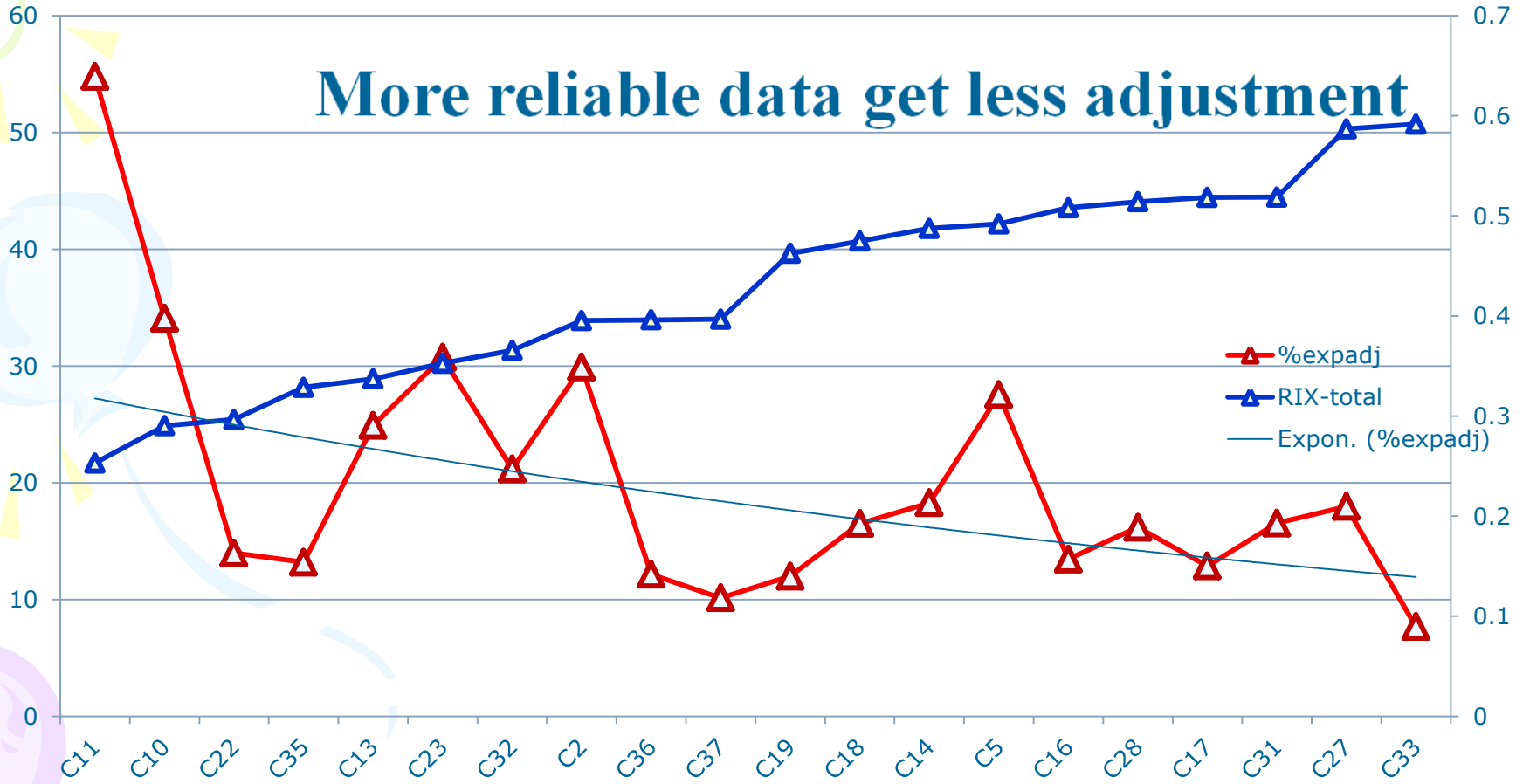
- Measurement of adjustment from official statistics: only the proportionate deviation and not the absolute deviation that matters

$$MAPA^{C} = \frac{100 \bullet \sum_{t=1995}^{2009} \sum_{r=1}^{g} | \overline{X}_{rt} - X0_{rt} |}{\sum_{t=1995}^{2009} \sum_{r=1}^{g} X0_{rt}}$$

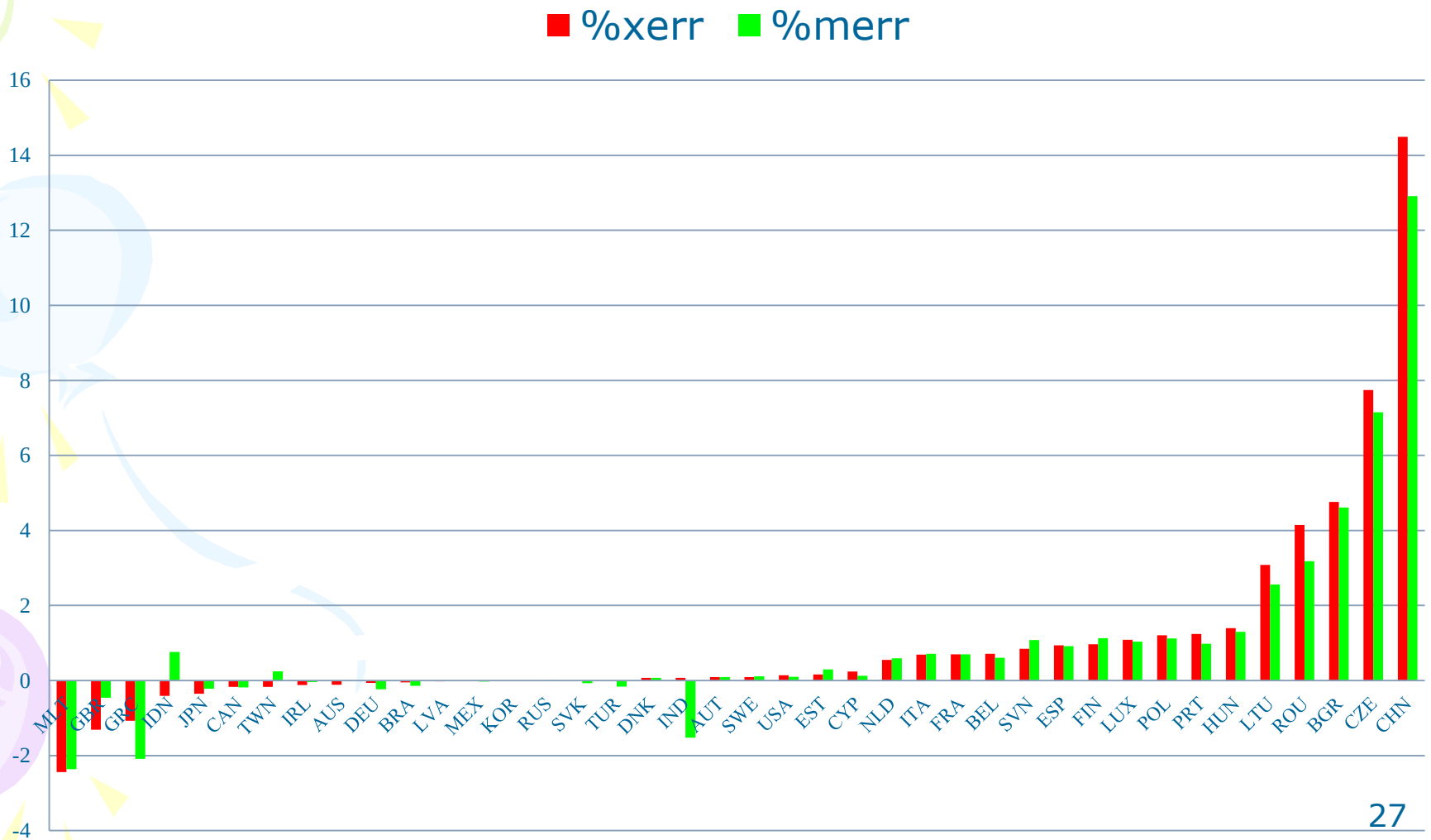
Reporter Reliability and Mean Absolute Percentage Adjustment of Total Exports by WIOD Countries, 1995 -2009



Reporter Reliability and Mean Absolute Percentage Adjustment of World Goods exports by WIOD product, 1995 -2009

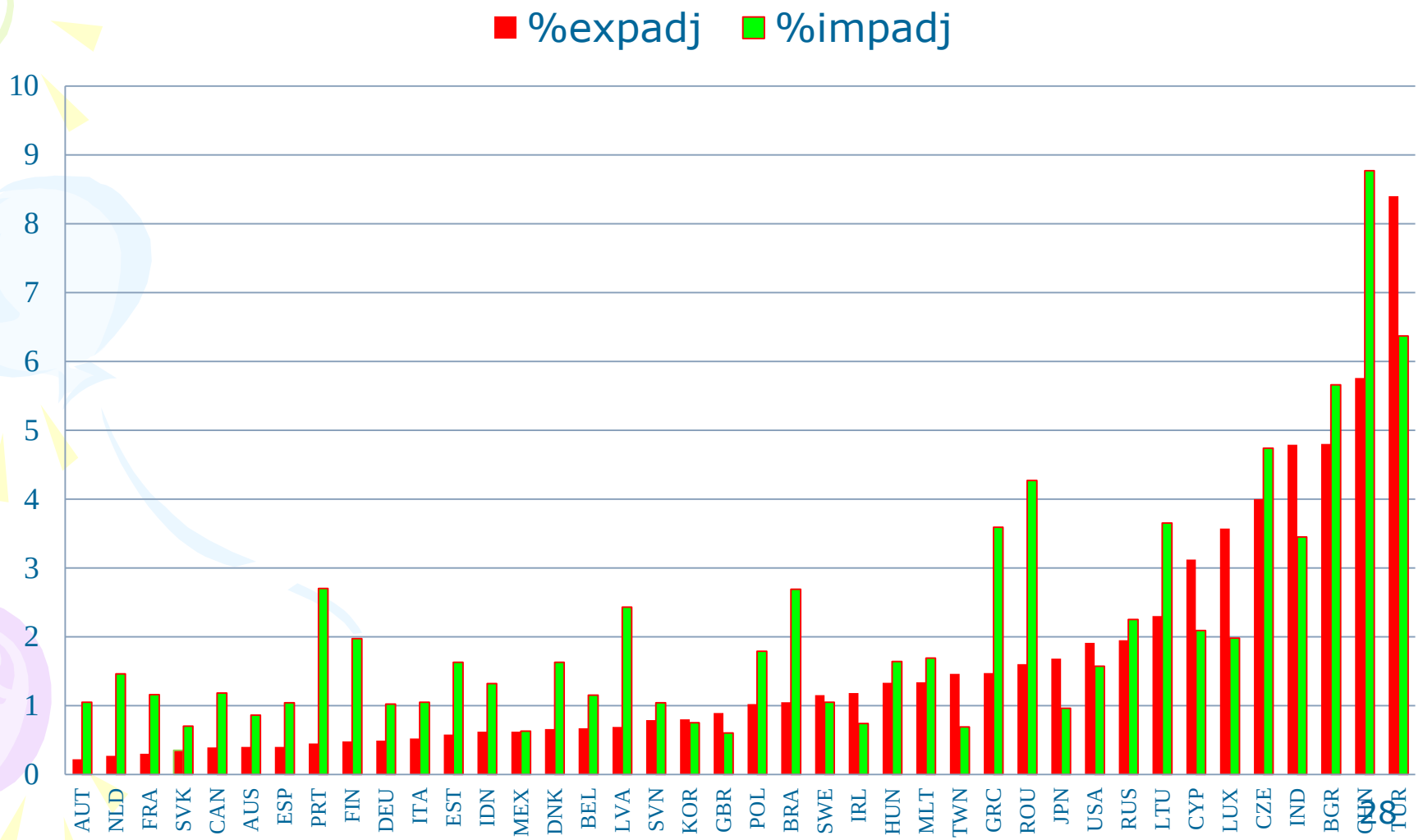


Initial difference of between WIOD SUT Exports/imports and Total Trade Statistics in National Account, 1995-2009



The adjustment for total trade Statistics Reported in National Account

1995-2009



Conclusions

- We developed a three-stage mathematical programming model to reconcile detailed bilateral goods and services trade statistics with individual country's SUTs to produce an global consistent SUT database that benchmarked to each country's GDP by expenditure account with minimum derivation from official SUT and trade statistics.
- We test the model using WIOD individual country SUT and UN national account and trade data, preliminary results show that the model is feasible and impose global consistency will make no significant changes on NSI's reported GDP and other major aggregate national account statistics in the balanced global SUT database. Therefore, the model may have great potential in the estimation of an integrated global ICIO account.

Thanks for your attention!

- Other comments/questions?

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

Why Supply and Use tables are Better?

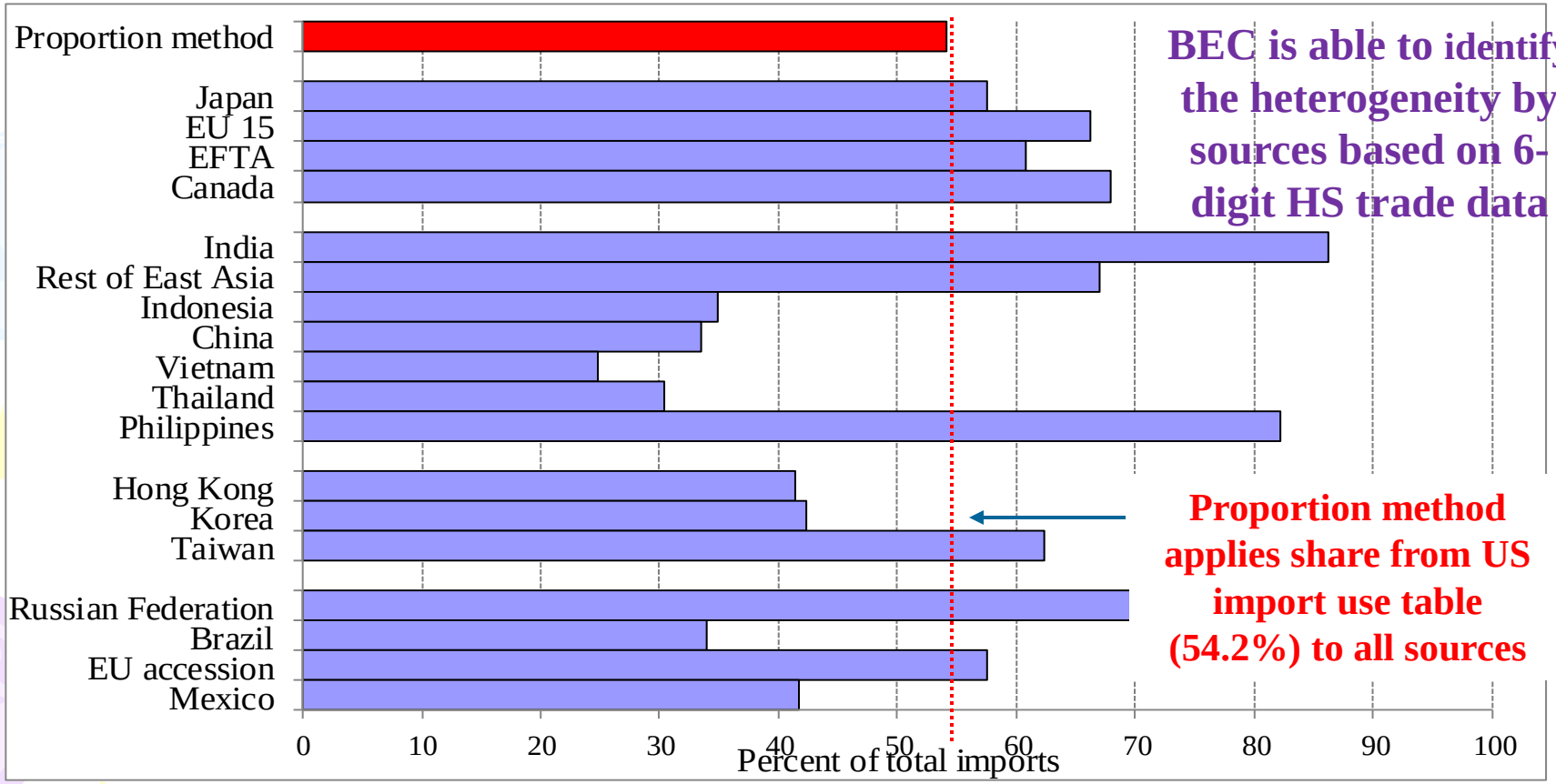
- Easy benchmark to National Accounts statistics
- Has both industry and commodity dimension that consistent to the UN standard of system of national account
- More close to actual data collected by NSI and avoid errors inherent in the assumptions imposed when transferring SUTs to symmetric IO tables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

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

Why BEC is Better than Proportional Assumption

Intermediate share of U.S. electronic machinery imports, by source, 2004



What end-use classifications can help

- Intermediate goods identified from gross trade flows are the row sum of each block matrix Z_{sr} in the IO flow matrix Z.
- End use classification such as BEC distinguishes intermediate inputs from final goods in imports from each source in each sector, improve the accuracy of IO coefficients in ICIO table by giving better row total control for each block IO coefficient matrix A_{sr}.
- End-use classifications can identify the heterogeneity of intermediate inputs entering the importing country from different sources, thus is better than the alternative: Proportional method assumes the intermediate share in imports from each source country are the same so it bias the total (include indirect) value-added estimates for each source country, even at aggregate level.

What end-use classifications can't help

- Still have to assume proportionality to allocate intermediate inputs to each industry *within the importing country*
  - Required data not reported by most national statistical agencies
- Industry-level estimates of value-added trade based on such IRIO table may be unreliable with unknown biases, despite their theoretical tractability
- To improve the sector level results,
  - Current end use classifications need to be extended to dual use products and services trade.
  - methods need to be developed to properly distribute imports to domestic users: link firm character data from survey or economic census and customs transaction level statistics.

Why Adjusted Estimates are Better?

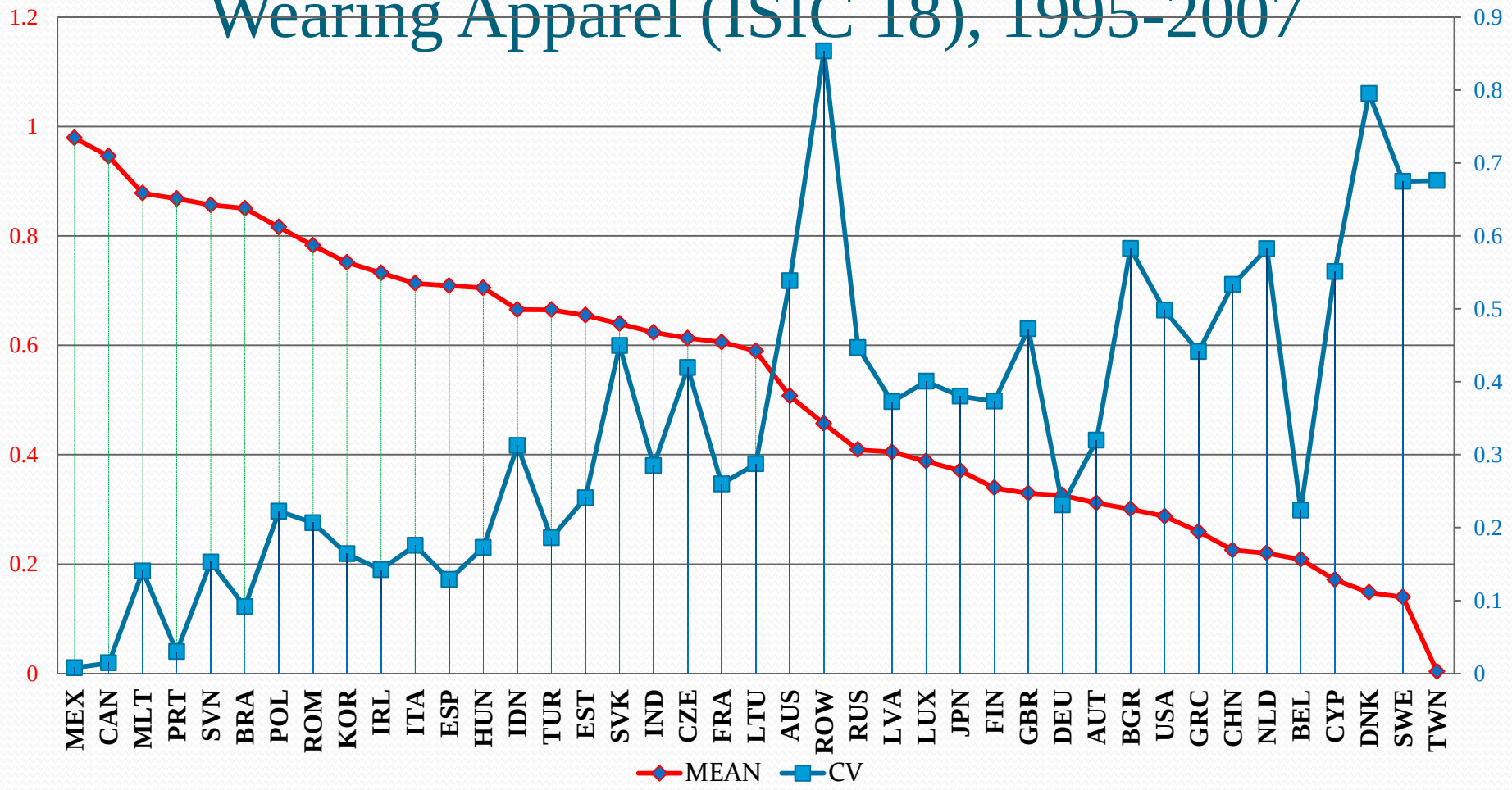
- **D0**: Initial estimates
- **W** variance matrix of initial estimates ,
- **A** coefficient matrix of all linear constraints AD* = 0
- The BLUE :

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

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

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

Average Exporter Relative Reliability Index Wearing Apparel (ISIC 18), 1995-2007



Average Importer Relative Reliability Index Wearing Apparel (ISIC 18), 1995-2007