International Conference on Business, Finance and Economics 2018

ISBN	978-81-933584-5-0	**VOL**	01
Website	www.coreconferences.com	**eMail**	mail@coreconferences.com
Received	15 – January – 2018	**Accepted**	10 - February – 2018
Article ID	CoreConferences011	**eAID**	CoreConferences.2018.011

Distortions in Trade Statistics Revisited: Data and Empirical Issues

Sho Haneda[1], Akihiro Yogata[2], Naohiko Ijiri[3]
[1]School of Social Welfare, Tokyo University of Social Welfare
[2]Graduate School of Economics, Nihon University
[3]College of Economics, Nihon University

Abstract: *This paper aims to quantify distortions due to "data updating problem" in trade statistics that might cause biases in econometric analyses and policy evaluation. Using the UN Comtrade database over the period of 2005-2015, three main results are clarified. Firstly, the paper finds differences between old and updated data regarding the number of transactions as well as trade values. Secondly, the degree of distortions significantly differs among countries, even within OECD countries. Finally, estimation results indicate that the coefficient on independent variables can be changed because of the data updating problem in econometric analyses. As a policy implication, it should be noted that a replication of econometric results of previous studies requires exactly the same data. Thus, it may be required that UN Comtrade and other statistic offices keep old data in their website.*

JEL Classifications: F13, F14, F23

Key words: Trade statistics, Distortions, Data updating, UN Comtrade, HS classification

1. Introduction

Quality and availability of trade data have been crucial for economists and policy makers, especially in econometric analyses policy evaluation. A variety of trade statistics are published and can be freely downloaded in order to conduct empirical studies. However, previous studies basically use the "snapshot" of trade statistics in their researches. In this research, trade data obtained from the United Nations (herein UN) Comrade are employed to check the biases caused by the problem of data updating. The data updating problem refers to the issue such that the values of the first released data differ drastically from those of updated data. In order to fill the gap, this paper intends to quantify the distortions in trade statistics by comparing old and updated data in UN Comtrade. The aims of this study is to quantify distortions in official trade data published by international organizations. The rest of this paper is organized as follows. The section 2 introduces the methodology for a data manipulation

2. Methodology

This section explains the methodology for constructing databases that can identify the biases in trade statistics due to the updating problem.

2-1 Explanation of HS System

The trade data are reported from member countries to the UN according to the rule of International Merchandise Trade

Statistics: Concepts and Definitions (hereinafter IMTSCD). However, the statistical reports submitted by some countries should be modified because reporting system of these countries tends to differ from that of IMTSCD. To publish trade statistics of member countries, the UN basically revises the reported data in accordance with their rule.

Trade values are reported in two manners, which are Free On Board (FOB) for export values as well as Cost, Insurance and Freight (CIF) for import values. The former only includes the value of transactions while the latter additionally contains the cost of insurance and transportations.

Trade data are published based on Harmonised Commodity Description and Coding System (HS), Standard International Trade Classification (SITC) and Broad Economic Categories (BEC) classifications. Since the HS has more than 5,000 product IDs, this paper employ the HS classification for checking distortions. The HS system was developed by World Custom Organization (herein WCO). The classification has been updated every five years so far. It is organized by 4 levels, which are Section, Chapter, Heading and Subheading. The 6-digit level HS system is internationally common while each country has own schedule beyond 6-digit.

2-2 Data Coverage

a) Country

In this paper, 10 reporter countries and all partner are included. Since the period of data collection is from April 2017 to October 2017, this study only has a relatively small number of transactions.

b) Product

The research is based on HS 6-digit level and all products are covered.

c) Period

As we only obtained the data for 10 countries, target periods are 2005-2015.

d) Classification

As it is mentioned that the HS classification has been revising its codes every five years, this paper covers the following versions: HS1996, 6-digits HS2002, 6-digits HS2007, 6-digits, HS2012, 6-digits and HS2017, 6-digits.

2-3 Calculation Method

In order to investigate the distortions in trade reports due to the data updating problem, the paper calculates three types of variables (see table 1). Firstly, it defines the transaction 3 as the **REVEALED transaction** since trade value exists only in the updated data. Secondly, the paper uses the transaction 2 as the data for the **HIDDEN transaction** because it only appears in the old data. Finally, the transaction 4 is included in the **REVISED transaction** as both old and updated data have it while trade values differ between them.

Table 1: Definitions of each variable

Transactions	Flow	Reporter	Partner	Period	HS code	Trade value (Old data)	Trade value (Updated data data)
1	Export	Japan	UK	2014	871310	107,469	107,469
2	Export	Japan	UK	2014	871320	5,523	N/A
3	Import	Japan	UK	2014	831790	N/A	112,563
4	Import	Japan	UK	2014	871310	112,325	201,252

Note: These are defined by authors.

3. Stylized Facts: Distortions in Trade Statistics

This section summarises the characteristics of our results using UN Comtrade database. Main findings are threefold. Firstly, the paper finds the gap between first released data and updated data regarding the number of transactions. Secondly, the degree of the difference significantly differs among countries, even within developed economies. Thirdly, REVISED transactions may be the main source of biases in econometric analyses. Finally, the value of total trade seems not to change substantially after the updates.

4. Econometric Analysis

In order to achieve the purpose, this paper uses two types of dependent variables in the empirical analysis. The one is the value of exports and imports in old data and the other is those in updated data. The paper employs these variables to conduct our empirical analysis.

There are two steps in this section. Firstly, the paper estimates the gravity equation of international trade by Ordinary Least Square (OLS) with robust standard errors. Secondly, this study checks the similarity of coefficients from different OLS regressions, which are regression for old data and that for new data, using the Chi-squared test.

In the first stage of the empirical section, the baseline specification is:

$$lnTrade_{ijklt} = \beta_1 lnGDP_{jt} + \beta_2 lnDISTANCE_{ij} + \beta_3 Contiguity_{ij} + \beta_4 LANGUAGE_{ij} + \beta_5 COLONY_{ij} + \beta_6 COMColony_{ij} + \varepsilon_{ijt} \quad (1)$$

where *i, j, k, l* and *t* denote reporter, partner, trade flow, product and year respectively. In addition, *Trade* and *DISTANCE* are defined as value of trade as well as physical distance between two countries. *Contiguity* is a dummy variable for country-pairs that have a common border. *LANGUAGE* is equal to 1 if a country-pair shares the same language and 0 otherwise. *Colony* and *COMColony* are dummy variable for county-pairs which had a colonial relationship between them and that for country-pairs which have a common coloniser. Finally, ε is the error term.

As we explained above, one of our aims is to test the difference between old and new trade data. In order to examine the similarity (or difference) of coefficients from them, the paper uses *suest* command and Chi-squared test in STATA. This study checks the similarity using REVISED transactions for each exporter in order to quantify the bias caused by data updating issues. The results are summarised in table C for exports and D for imports.

Main findings are twofold. Firstly, it is clarified that there might be the difference in the coefficients between the results from old data and those from new data. Secondly, the degree of the variance depends on trade flow (export and import), reporter, and independent variables.

5 Concluding Remarks

This paper finds that the data updating problem exists, even for OECD countries. Also, the issue can cause the biases in the quantitative analysis.

As a policy implication, it should be noted that replication of the results of previous studies needs exactly the same data. Thus, it might be required that UN Comtrade and other statistic centres keep old data in their website. Furthermore, an international harmonisation of data collection and revision methods in the world need to be considered for the TRUE policy evaluation.

Further studies can target other economic variables such as FDI, income, employments, etc. In addition, the ranking of the quality of statistics of each country may be important for future research.

Main References

1. Beja, E. L. (2008). Estimating Trade Mis-invoicing from China: 2000–2005'. *China and World Economy*, 16, 1, 82–92.

2. Buehn, A. and Eichler, S. (2011). Trade Misinvoicing: The Dark Side of World Trade. *The World Economy*, 2011, 1263-1287.
3. Fisman, R. and Wei, S.-J. (2004). Tax Rates and Tax Evasion: Evidence from ''Missing Imports'' in China. *Journal of Political Economy,* 112, 2, 471–96.
4. Fisman, R. and Wei, S.-J. (2007). The Smuggling of Art, and the Art of Smuggling: Uncovering the Illicit Trade in Cultural Property and Antiques. *NBER Working Paper No. 13446.*
5. IMF (1993). A Guide to Direction of Trade Statistics (Washington, DC: IMF).
6. Kellenberg, D and Levinson, A. (2016). MISREPORTING TRADE: TARIFF EVASION, CORRUPTION, AND AUDITING STANDARDS. *NBER Working Paper No. 22593*
7. Kumakura M. (2011). The Characteristics and Usage of UN Comtrade. *IDE JETRO Research report 2010-2-03.*

Appendix

Table A: Target countries and date of updates

Country	Year	Publication Note	First Publication Date	Last Publication Date
Brazil	2007	Full revision	2008.03.03	2017.05.23
	2008	Full revision	2009.04.20	2017.05.23
	2009	Full revision	2010.01.15	2017.05.23
	2010	Full revision	2011.02.02	2017.05.16
	2011	Full revision	2012.02.01	2017.05.23
	2012	Full revision	2013.02.07	2017.05.23
	2013	Full revision	2014.01.17	2017.05.24
	2014	Full revision	2015.01.21	2017.05.24
	2015	Full revision	2016.02.24	2017.05.25
Canada	2014	Full revision	2015.02.12	2017.04.26
	2015	Full revision	2016.02.23	2017.06.27
China	2014	Full revision	2015.05.07	2017.06.28
	2015	Full revision	2016.05.25	2017.06.29
Estonia	2015	Full revision	2016.02.13	2017.09.06
Finland	2005	Full revision	2006.05.24	2017.03.02
Malta	2015	Full revision	2016.06.28	2017.05.15
Mexico	2015	Full revision	2016.03.01	2017.10.06
Netherlands	2015	Full revision	2016.05.20	2017.09.13
Rep. of Korea	2015	Full revision	2016.09.23	2017.05.08
Sweden	2015	Full revision	2016.03.17	2017.05.09

Source: UN, Comtrade.

Table B: The change in the value of total trade

		The value of trade (billion US dollar)			Total changes in the value of trade (billion US dollar)					
Country	Year	Export	Import	Total trade	Export	(%)	Import	(%)	Total trade	(%)
Brazil	2007	161	121	281	-0.006	0.00%	0.002	0.00%	-0.004	0.00%
	2008	153	135	288	**4.766**	**3.12%**	-0.201	-0.15%	**4.565**	**1.58%**
	2009	153	128	281	0.001	0.00%	0.008	0.01%	0.008	0.00%
	2010	202	182	384	**4.542**	**2.25%**	1.224	0.67%	**5.766**	**1.50%**
	2011	256	226	482	-0.068	-0.03%	0.020	0.01%	-0.048	-0.01%
	2012	243	223	466	-0.019	-0.01%	0.408	0.18%	0.388	0.08%
	2013	242	240	482	**4.145**	**1.71%**	0.860	0.36%	**5.005**	**1.04%**
	2014	225	229	454	0.067	0.03%	0.005	0.00%	0.072	0.02%
	2015	191	171	363	**2.450**	**1.28%**	0.684	0.40%	3.134	0.86%
Canada	2014	475	463	938	1.602	0.34%	0.087	0.02%	1.689	0.18%
	2015	409	420	828	0.081	0.02%	0.601	0.14%	0.682	0.08%
China	2014	2,342	1,959	4,302	-0.016	0.00%	1.142	0.06%	1.126	0.03%
	2015	2,273	1,680	3,953	-8.415	-0.37%	-2.194	-0.13%	-10.608	-0.27%
Estonia	2015	14	16	30	0.006	0.04%	-0.005	-0.03%	0.001	0.00%
Finland	2005	65	58	124	0.013	0.02%	0.002	0.00%	0.015	0.01%
Malta	2015	4	7	11	**1.335**	**34.09%**	**0.995**	**14.66%**	**2.330**	**21.77%**
Mexico	2015	381	395	776	-0.095	-0.02%	-0.029	-0.01%	-0.124	-0.02%
Netherlands	2015	474	425	899	0.027	0.01%	0.005	0.00%	0.033	0.00%
Rep. of Korea	2015	527	436	963	-0.079	-0.02%	-0.024	-0.01%	-0.103	-0.01%
Sweden	2015	140	138	278	-0.138	-0.10%	0.267	0.19%	0.129	0.05%

Source: UN Comtrade.

Note: Total change in the value of trade is defined as the following: The sum of the value of REVEALED trade - The sum of the value of HIDDEN trade + the sum of the value of REVISED trade.

Table C: The similarity of coefficients for exports

Export							
		Variable					
Country	Year	GDP	Distance	Contiguity	ommon language	Colony	Common colony
Brazil	2007						
	2008						
	2009				***		
	2010	***	***				
	2011	***	***	***	***	***	
	2012				***		
	2013				***		
	2014						
	2015				***		
Canada	2014		***	***		**	
	2015		*				
China	2014		***		***		
	2015	***	***	**	***		
Estonia	2015						
Finland	2005						
Malta	2015		**			***	
Mexico	2015					**	
Netherlands	2015		**	*			
Rep. of Korea	2015		**				
Sweden	2015	***		***	**	**	

Note: ***, ** and * denote that the equality of coefficients can be rejected by Chi-squared test at 1%, 5% and 10% level of significance respectively.

Table D: The similarity of coefficients for imports

Import							
		Variable					
Country	Year	GDP	Distance	Contiguity	Common language	Colony	Common colony
Brazil	2007	***	**		***	***	
	2008		***	***			
	2009	**	**	***	***		
	2010	***	***	***	***	***	
	2011	***	*	*	*		
	2012	**					
	2013	***					
	2014						
	2015						
Canada	2014						
	2015						
China	2014	***				***	
	2015	*					
Estonia	2015					***	
Finland	2005						
Malta	2015	***	***		***	***	***
Mexico	2015						
Netherlands	2015	***				***	
Rep. of Korea	2015						
Sweden	2015				**	**	

Note: ***, ** and * denote that the equality of coefficients can be rejected by Chi-squared test at 1%, 5% and 10% level of significance respectively.