Reconciling International Trade Data

Shaar, Karam

Victoria University of Wellington

September 2017

Reconciling International Trade Data

Karam Shaar [a]

Abstract

International trade data are in substantial measurement error. Data reported by some countries mean next to nothing. This study develops an index of trade data quality based on the consistency between a country's claims on bilateral trade and the corresponding claims of the rest of the world from 1962 to 2014. The index takes the relative significance of each partner and data availability into account. We produce a more reliable set of bilateral and total international trade data using the index. Findings include (a) the actual exports of most countries with low data quality are considerably higher than self-reported. (b) Corruption and poor data quality are strongly correlated. (c) Global trade data quality has been deteriorating in the past three decades even though more countries have improved their data quality over time. This is because low-quality reporters have recently increased their share in global trade. (d) China tends to under-report exports and over-report imports. (e) There is only a trivial difference between US self-reported and reconciled data. The same applies to all high-quality reporters. We recommend future studies on trade use our reconciled data.

Keywords: Trade data quality, data discrepancy, trade data reconciliation

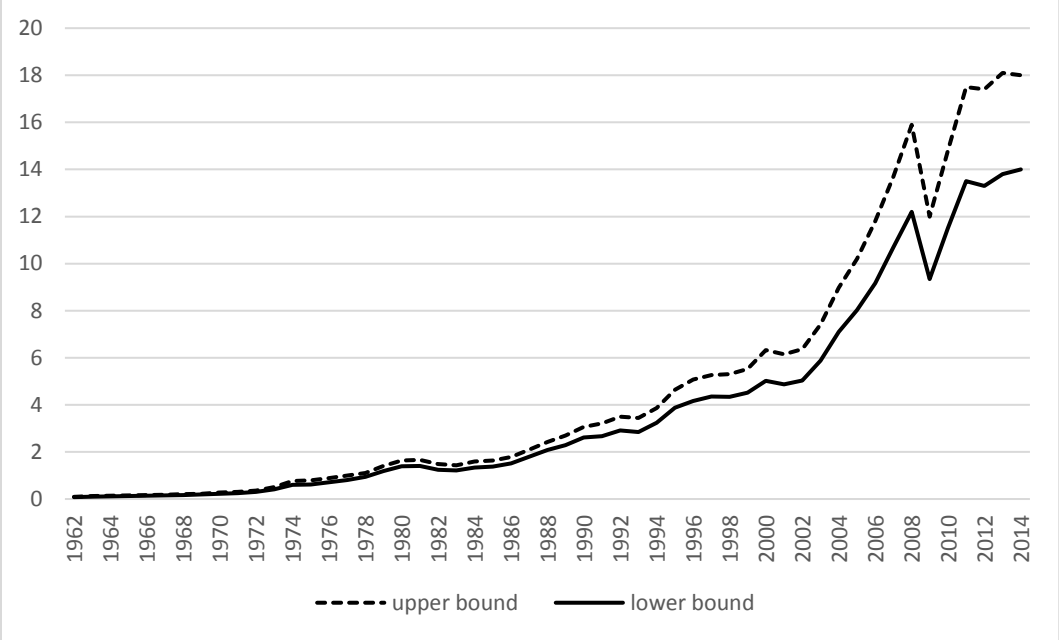
JEL Classification: F1, C02, C18

[a]: Victoria University of Wellington, Wellington 6140, New Zealand, PO box 600, karam.shaar@vuw.ac.nz

1. Introduction

Trade data discrepancy happens when two countries report different values of trade with each other. To have a sense of the magnitude of the issue, Figure 1 plots global trade based on the lower and higher report of every bilateral trade flow available in the period 1962-2014. Which of the two series represents actual global trade? No one knows.

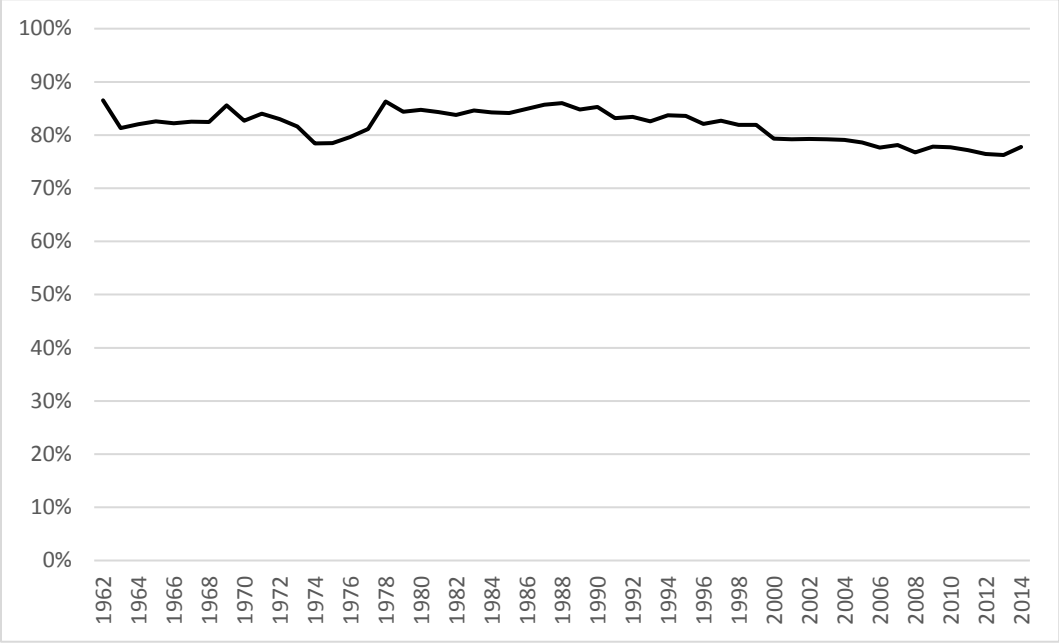
Figure 1: Global trade in trillions of dollars based on higher vs lower reports (1962-2014)



Note: global trade is calculated only for cases in which the same trade flow is reported by both partners, which covers 453,067 flows (906,134 observations). The numbers are calculated using Standard International Trade Classification (SITC), revision 1.

The ratio of the lower to upper bound has been in decline since the early 1990s as Figure 2 shows. That is when global trade started to grow at a substantially faster rate. This clearly indicates that global trade data discrepancy has been increasing over the past three decades. In 2014, the ratio stood at just 78%.

Figure 2: lower bound of global trade divided by the upper bound (1962-2014)



One might expect the value reported by the importer to be greater than the value reported by the exporter since the former usually includes cost, insurance, and freight in the reported value (CIF), while exporters report their trade as free on board (FOB). In contrast, in 36% of the cases in which bilateral trade data are reported by both partners, the value reported by the exporter is, in fact, higher than the value reported to the importer.

Substantial discrepancies can mean a study that uses data reported by a certain partner would reach different conclusions if it used the data reported by the other. Bahmani-Oskooee, Hegerty and Harvey (2013) report "drastically different results for the impact of exchange rate on trade between the US and South Korea depending on which of the two countries reported the data." Such results raise serious questions about the impact of trade data discrepancy on the validity of the empirical research. As stated by Gujarati (2004), "the researcher should always keep in mind that the results of research are only as good as the quality of the data". The impact of data discrepancy on empirical research is widely ignored in the literature. This is reflected by the fact that most studies do not even cite the reporting country but rather cite the data compiler/disseminator, such as the United Nations (UN), International Monetary Fund (IMF), or World Trade Organization (WTO).

The few previous studies explicitly aware of the problem of trade data discrepancy used an informal approach to reduce its impact on their research. Calderon, Chong and Stein (2007) state to "have always relied on the data reported by the country with higher income in the country-pair". This study alternatively provides a set of reconciled international trade data based on the actual quality of each country's data.

Although trade data discrepancy is endemic globally (Ferrantino, Liu and Wang 2012), the topic is widely under-researched in relations other than the US-China-Hong Kong trilateral trade. As shall be seen later, the quality of trade data in many other countries is considerably lower. As far as this paper is concerned, Gehlhar (1996) from Global Trade Analysis Project (GTAP) is the only attempt to reconcile international trade data. GTAP data are reconciled on bilateral commodity level and aggregated later on a regional level only. The data reconciled by the project are available for a fee. As shall be discussed in section 2, we believe GTAP methodology is asymmetric to over and under-reporting and does not control for data unavailability.

We construct Trade Data Quality Index (henceforth, TDQI) through comparing the claims of a certain country on bilateral total trade with the corresponding claims of all of its partners in a certain year. Through this comparison, we produce a measure of weighted average similarity between a country's data and the data reported by the rest of the world. The weights are the trade shares of each partner to reflect their relative significance. We finally arrive at TDQI through accounting for cases in which one of the partners does not report data to control for data unavailability. The index ranges from 0 to 100. It is calculated for the exports and imports separately in every year and for each country. The higher the TDQI, the higher the data quality.

We investigate all available bilateral trade data from 1962 to 2014, which comprises 1,228,492 observations. The main goal of the index is to reconcile international trade data. In any bilateral relation, we pick the data from the partner less likely to be in error, i.e. the partner with higher TDQI. We take the CIF/FOB margin into account. The following link provides the results of TDQI, the reconciled and unreconciled international trade data on bilateral and total levels, as well as the Stata programming codes used in the derivations. https://goo.gl/cGVXDk

Although global trade has never been this complex, imports and exports data quality is improving in more countries than worsening. In global aggregate terms though, the quality has been deteriorating, especially in the past three decades as highlighted in Figure 1. This is due to the increasing importance of low-quality countries in global trade.

We expectedly find that countries with high data quality for imports tend to have high-quality data for exports as well, and vice versa. Most importantly, we show that corruption and trade data quality are strongly correlated, which indicates that corruption in the customs of countries with low data quality might be responsible for the discrepancies in trade data. The causal inference could be investigated using data on commodity level. The paper also finds evidence that negligence in the customs of low-quality reporters can also explain trade data discrepancy.

Using TDQI and the sub-indices it is based on help in tracking who trades what for all countries from 1962 to 2014. The results reveal that the figures reported by low-quality reporters are in many cases substantially different from reconciled data. The paper provides evidence that China's self-reported exports are under-reported, while imports are systematically and considerably over-reported. Most studies in the literature focused on China's exports rather than imports.

In the case of the second largest trader in the world as of 2014, the US, we observe trivial differences between reconciled and unreconciled (self-reported) data. We generally find that exports tend to be under-reported by low-quality reporters, which adds more evidence that corrupt practices such as under-reporting and smuggling might partly explain the discrepancies. We alternatively find that reconciled and self-reported data for high-quality reporters are relatively the same. That is to say, no reconciliation is required in most cases.

Using the reconciled data in a measure that depends on trade in its construction also makes a considerable difference. For example, we calculate trade openness for all countries using reconciled and self-reported data and compare the results for the year 2014. Countries with low data quality are, on average, 11% more open to trade than thought in the literature, which solely depended on self-reported data. Countries with high TDQI have relatively the same trade openness using either data.

The rest of the paper is organized as follows. Section 2 explains the research methodology and the data. Section 3 presents the findings. Section 4 concludes.

2. Data and Methodology

Customs in different countries follow different national classifications to track their external commodity merchandise for purposes of duty imposition and economic analysis. Most countries also report their trade statistics to international institutions such as the UN and World Customs Organization (WCO). These institutions in their turn aim to achieve unanimous scales and definitions for international trade data. Among the most common international trade classifications are the Standard

International Trade Classification (SITC) of the UN and the Harmonized Commodity Description and Coding System, commonly known as the Harmonized System (HS), of the WCO.

HS is a multipurpose international commodity nomenclature. It was adopted in 1983 and entered into force in 1988. HS is revised every 5-7 years. It comprises nearly 5,000 commodity groups in its latest revision in 2012, which theoretically cover almost 16,000 final commodities. Each group is numbered by a six-digit code where an increase in digits reflects higher disaggregation. The groups are arranged in a legal and logical structure depending on the nature of the commodity (Harmonized System (2014).

The UN Statistical Division, Commodity Trade Statistics Database (COMTRADE), maintains the Standard International Trade Classification (SITC). SITC structure is based on the economic functions of commodities at different stages of development. SITC is the oldest international commodity classification system. The first version was released in 1950 and the latest revision (fourth) was in 2008. Both classifications publish bilateral total and commodity trade data.

We observe no considerable differences between bilateral total trade data reported by HS and SITC even in different revisions. The differences are largely on commodity level. For the sake of maximizing the time period, this study retrieves the data from SITC, revision 1, which is the earliest version available online. The annual data extend from 1962 to 2014 and cover all available bilateral total trade. This comes to 1,548,585 observations. The data are in US dollars. Although data are available for years 2015 and 2016, we exclude them from the analysis since some countries have not reported their data yet while some other reports are still subject to revisions by the reporters. The data are publicly available on many online portals. This paper retrieved the data from the World Integrated Trade Solutions (WITS) website, administered by the World Bank.

Although the IMF reports bilateral total trade data in its Direction of Trade (DOT) Database, we use SITC because, unlike the former, no estimates or predictions are made for any observation. Using SITC allows for testing the quality of the data as reported by countries. We compare DOT with SITC data for the whole time period of the study and find the two sets to be identical in most cases with the difference that DOT, in some cases, fills the gaps when the data is not reported by a certain country by its partner's report and uses a CIF/FOB margin of 10% to convert from one type of estimation to the other. It is not clear why only some unreported data are replaced by the partner's claims. DOT also occasionally estimates annual trade values based on available monthly data. WTO does not report bilateral data and the number of reporting countries for aggregate trade is smaller than in SITC and HS.

WTO data also extend over shorter periods. The World Bank retrieves its trade data from WTO, IMF, and the UN.

Table 1: potential causes of discrepancy

Type	Cause	explanation
lack of uniformity in data compilation methodologies between the partners	what does constitute exports and imports?	The treatment of re-exports, goods in transit, inward and outward processing, and re-imports.
	partner attribution	attributing exports to the final known destination or the country of consignment attributing imports to the country of origin or the country of consignment
	geographical definition of a trade partner	example: whether the Virgin Islands are a part of the United States or a separate entity/country.
	conversion to foreign currency	what method is used to express the value of trade in terms of a foreign currency
uniformity in data compilation methodologies	The valuation method of exports and imports	where imports are usually reported on CIF basis exports are usually reported on FOB
	timing effect	shipments are registered at different points of time by both countries since they are registered as they happen
Corruption	mis-invoicing	over-invoicing the value of a shipment to take advantage of certain export support schemes or under-invoicing for duty evasion
	smuggling	not registering a shipment for duty avoidance or due to the illegality of the traded goods by either one of the partners or both
	partner misattribution	attributing trade to another partner (against the stated methodology), to benefit from lower duties
Negligence		

* Source: Author's summary

Corresponding claims on bilateral trade may not be directly comparable for the reasons listed in Table 1. UN Statistical Division regularly publishes its recommendations to individual countries to enhance comparability. SITC distinguishes between total and gross bilateral trade.

Gross exports= total exports+ re-exports

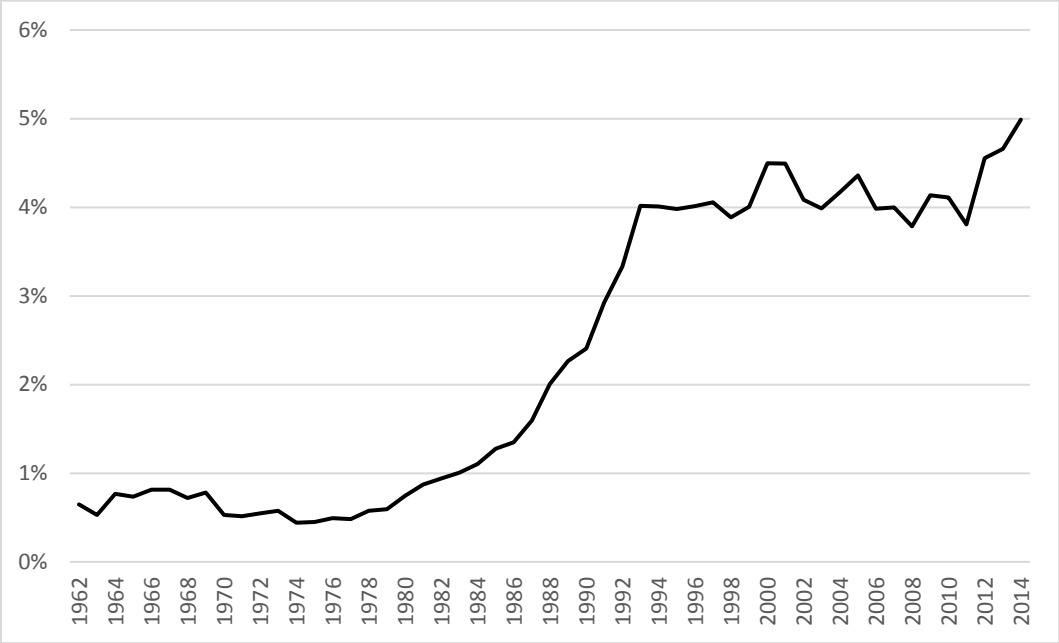
Gross imports= total imports+ re-imports

As defined by the UN Statistical Division (2011), re-imports are imports of domestic goods which were previously recorded as exports (returned). Re-exports are exports of foreign goods which were previously recorded as imports and later re-exported without 'substantial changes'. Assuming no bilateral re-exports and re-imports take place, total and gross bilateral trade data would be identical. For all countries along the period 1962-2014, reported re-imports as a percentage of gross imports is less than 0.5%. To ensure comparability, we assume no re-imports took place during the period of the study. Alternatively, we assume that all countries are equally engaged in re-importing, which allows TDQI to remain comparatively correct, which is what we need as we reconcile international trade data through relying on the claim of the country with higher TDQI in any bilateral relationship.

Global re-exports as a percentage of gross exports along 1962-2014 is almost 2.3%. In 2014, as much as 77% of reported world re-exports went through Hong Kong and the US only, with 53.5% and 23% respectively. According to UN Statistical Division recommendations, a country should attribute its exports to the final known destination and its imports to the country of origin.

This study uses bilateral total data rather than bilateral gross data because economists mostly care about the actual trade from domestic origin rather than re-exports. Total trade also enhances the comparability between corresponding trade flows. Assuming all global re-exports happen through Hong Kong and the US only, and if the rest of the world complies with the recommendations of attribution, TDQI of the imports of Hong Kong and the US would be downwardly biased. TDQI for their exports remains unaffected since re-exports are not included in bilateral total exports but are in fact included in total imports when they enter a country. As we observe later, however, not all countries are capable of identifying the country of origin or the country of final destination. This requires a special treatment for re-exporting countries, especially in the recent three decades in which re-exports as a percentage of global gross exports started to exceed 1%.

Figure 3: Global share of re-exports in gross exports for 1962-2014



As stated in Table 1, the third reason for trade data discrepancy is the geographical definition of a trade partner. To this end, we drop from our analysis any entity or country, which is not considered by its partners as a partner. To ensure comparability, we drop any observation relating to exporting to or importing from them. As listed in Appendix B, most of these are tiny islands. This leaves us with 1,228,492 bilateral trade observations. The share of the exports reported by these entities in global self-reported exports along the whole period of the study is 1.8%. The share, in reality, is lower as some of these entities are double-counted such as Ethiopia (excluding Eritrea), where we still include Ethiopia (including Eritrea) in our analysis. The rest of our results and graphs are all based on the sample after dropping these entities.

The causes of discrepancy related to timing and conversion to foreign currency are expected to play a trivial role. Since the data are annual, the beginning and end of the period are expected balance each other out unless a substantial growth in trade occurred during that period. UN Statistical Division has very clear recommendations on conversion to foreign currency as well.

The other issue to be addressed is the fact that countries, as also recommended by the UN Statistical Division, tend to report their exports based on FOB and their imports based on CIF. This recommendation deters us from directly comparing corresponding bilateral trade data since the importer includes the costs of insurance

and freight in the value of imports and the exporter does not include them in the value of exports. However, the same logic applied to the treatment of re-imports in TDQI applies here since the difference between corresponding flows is expected to be systemic. Assuming all countries are equally likely to trade with the same combination of far and near partners, TDQI remains comparatively correct. We take the issue of CIF-FOB estimation into account when we reconcile the data. The causes of discrepancy related to corruption and negligence are discussed later in section 3.

For a specific year and direction of trade, let N denote the number of all trade partners involved in the calculation of TDQI for country i. For partners $j = 1, \dots, N$, we define:

M_{ij}: import value of i from partner j, reported by i

X_{ij}: export value of i to j, reported by i

M_{ji}: import value of j from i, reported by j

X_{ji}: export value of j to i, reported by j

$M_{i.}$: sum of import values of i from all partners, reported by i

$X_{i.}$: sum of export values of i to all partners, reported by i

The similarity between corresponding trade flows for country i's exports is defined as:

$$S_{ij}^{X} = 1 - \frac{|X_{ij} - M_{ji}|}{X_{ij} + M_{ji}} \qquad [1]$$

The similarity between corresponding trade flows for country i's imports is defined as:

$$S_{ij}^{M} = 1 - \frac{|M_{ij} - X_{ji}|}{M_{ij} + X_{ji}} \qquad [2]$$

For all years in our data, we have 453,067 bilateral trade flows for which there are data reported by both partners (906,134 observations) to calculate the similarity measures. The remaining total bilateral trade observations are reported by one side only, which we treat differently as shown later. The similarity measures are unit-free, and therefore, are comparable over time and across countries without worrying about inflation or trade value. The measures are confined between greater than zero

and one and the higher the similarity is, the more consistent the two claims are. Gehlhar (1996) uses |exports-imports|/imports as a measure of similarity (called Accuracy Level), which ranges from zero to infinity. Their measure is used to decide whether two claims match or not based on an arbitrary threshold of 0.2.

Table 2: Illustrative example of GTAP methodology of measuring similarity

	New Zealand-reported imports from the US (m)	US-reported exports to New Zealand (x)	calculating GTAP measure for NZ imports |x-m|/m	proposed ratio |x-m|/(x+m)
scenario 1	33	40	0.21 (not matched)	0.10
scenario 2	40	33	0.18 (matched)	0.10

Accuracy Level measure is asymmetric as it yields different results when the difference in claims is the same. Therefore, GTAP measure is biased towards over-reporting as compared to under-reporting when calculating the accuracy level for the importer.

Now we turn to calculating the relative significance of each similarity measure between country *i* and each of its trade partners as reflected by the trade shares. Trade shares are unit-free, range from greater than 0 to 1, and add up to unity.

Exports of country *i* to country *j* as a share of *i*'s total exports as reported by *i* is:

$$Sh_{ij}^{X} = \frac{X_{ij}}{X_i} \qquad [3]$$

Imports of country *i* from country *j* as a share of *i*'s total imports as reported by *i*:

$$Sh_{ij}^{M} = \frac{M_{ij}}{M_i} \qquad [4]$$

For country *i*, we calculate the weighted average of the similarities between *i*'s reports on trade and the corresponding values reported by all of its partners *j* to *N*. Average similarities for imports of country *i*:

$$u(S_i^M) = \sum_{j=1}^{N} Sh_{ij}^{M} * S_{ij}^{M} \qquad [5]$$

Average similarities for exports of country *i*:

$$u(S_i^X) = \sum_{j=1}^{N} Sh_{ij}^{X} * S_{ij}^{X} \qquad [6]$$

Average similarities for imports and average similarities for exports for each country in each year range from greater than 0 to 1. The higher the $u(S_i)$, the higher

the average similarity between country *i*'s data and the data reported by its partners. The rationale behind the measure of average similarities as a proxy for data quality is straightforward. A country which generally reports trade values that are inconsistent with what is claimed by its trade partners is assumed to have low data quality.

Each single similarity measure between two trade values is included in the calculation of two average similarities. The $u(S)$ for imports of one country and the $u(S)$ for exports of the other. Because of this, a country with higher data quality gets its average similarity lowered due to trading with a country with lower quality. However, as *j* grows to its maximum *N* in Equations 5 and 6, all countries are assumed to be equally likely to trade with the same combination of countries with high and low data quality. Therefore, although downward biased by itself as a measure of quality, TDQI remains comparatively correct, i.e. compared from a year to another, trade flow to another, and a country to another.

The main concern about the average similarities measure as a proxy for data quality arises when a country does not report all of its bilateral trade data. This would cause the measure to be incorrect due to the bias of the trade shares used in constructing it. Therefore, a country with low data availability might have an artificially high or low average of similarities. GTAP reconciled data does not take this issue into consideration. To account for the impact of data availability, we first develop a Trade Data Availability Index (henceforth, TDAI).

For country i in a specific year, let:

m_i^M: number of countries which reported exports to country i while i did not report any imports from them

n_i^M: number of all countries which reported exports to country i

m_i^X: number of countries which reported imports from country i while i did not report any exports to them

n_i^X: number of all countries which reported imports from country i

Imports TDAI for country i is given as follows:

$$TDAI_i^M = \left(1 - \left(\frac{m_i^M}{n_i^M}\right)\right) * 100 \qquad [7]$$

Exports TDAI for i:

$$TDAI_i^X = \left(1 - \left(\frac{m_i^X}{n_i^X}\right)\right) * 100 \qquad [8]$$

TDAI ranges from 0 to 100. The higher the TDAI, the more available the bilateral trade data of a country. However, not reporting bilateral trade with a certain partner or partners might be the right practice from the home country i because such trade did not actually happen. Assuming all countries are equally susceptible to this scenario, TDAI remains comparatively correct although might be downwardly biased by itself.

Since not reporting any bilateral data, while the partner does, indicates a similarity of 0, the trade data quality of country i equals:

$$TDQI_i^X = u(S_i^X) * TDAI_i^X \qquad [9]$$

$$TDQI_i^M = u(S_i^M) * TDAI_i^M \qquad [10]$$

As average similarities range from above 0 to 1 and TDAI ranges from 0 to 100, the product of the two ranges from 0 to 100. Additionally, since TDAI and the average of similarities are both downward biased, yet comparatively correct, TDQI should also be understood only comparatively. The measure is a function of not only the quality of available data but also the level of data availability. The higher the index, the higher the quality of a country's trade data.

TDQI can simply tell us which partner in any bilateral trade relation is more likely to be more accurate, which serves as the basis of data reconciliation. In any bilateral trade relation, we reconcile the discrepancy in claims through picking the data from the source with higher TDQI. Following Gehlhar (1996), we do not use the average of claims weighted by TDQI as a basis of reconciliation because the data reported by some low-quality reporters are next to meaningless.

When reconciling a bilateral flow using the exporter's claim, we use the value with no change. When we reconcile using the importer's claim, we convert the value from CIF to FOB. This is to ensure the reconciled data are consistent on the basis of estimation (all FOB). To this end, we use a conversion factor of 10% following the Direction of Trade data disseminated by the IMF. To further improve the quality of reconciled data, we encourage future studies to estimate relation-specific CIF/FOB margins rather than using a unanimous 10% for all relations.

The problem of using TDQI for reconciling trade data happens due to re-exportation. As discussed earlier, most reported global re-exports happen through Hong Kong and the US with more than half global re-exports going through the former only. The impact of re-exports on TDQI depends on whether the country of origin, the re-exporting country and the country of final destination comply with UN Statistics Division recommendations of partner attribution. Otherwise, it is very hard to determine who traded what. The relatively high TDQI for exports and imports for the US (around 90%) along the period of the study indicates that the countries of final destination, mostly Canada and Mexico, correctly attribute their imports to the country of origin. This could be because the US properly informs them about the true origin of the goods.

Hong Kong's TDQI for imports is also high. The problem of using TDQI for data reconciliation for Hong Kong happens at the side of exports as some high-quality reporters wrongly report Hong Kong as the country of origin, while in reality, the goods originate from China. This might happen due to the change in goods ownership once in Hong Kong and before re-exportation (Ferrantino, Liu and Wang 2012). Since this information might be only available to Hong Kong, it is hard for its partners to identify the correct country of origin.

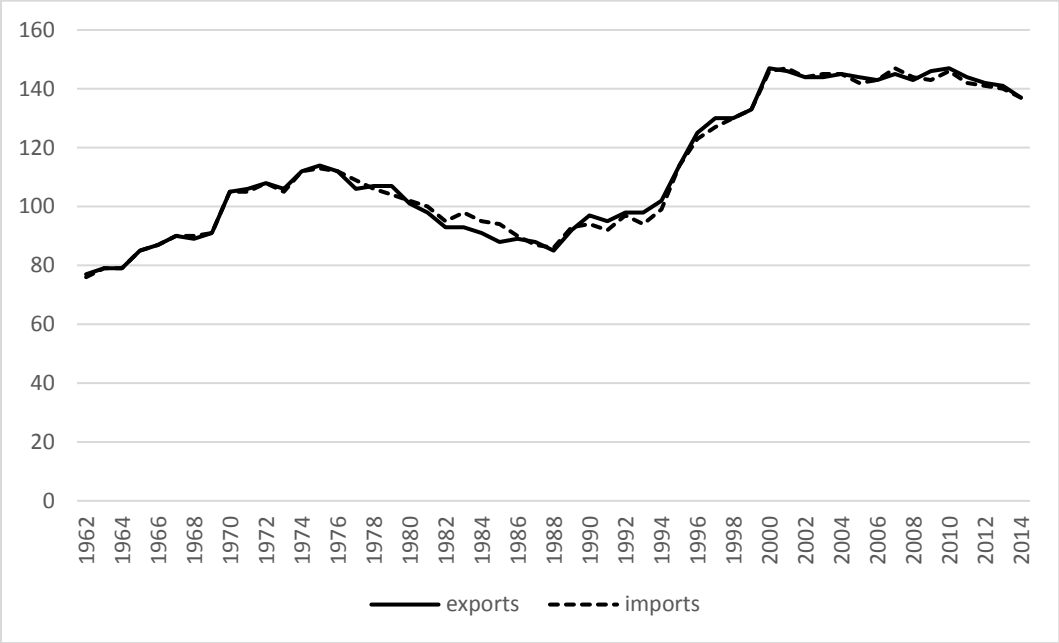
This is why we believe Hong Kong's exports data are more accurate than the data reported by importers from her. Even when these importers have a higher TDQI for their imports. Hong Kong's exports are the only case where we do not reconcile the data using the partners' imports but using Hong Kong's exports regardless of TDQI score. In other words, this is the only case where we believe a country is right and the rest of the world is wrong.

The uncertainty surrounding the trade of re-exporting countries and their partners using unreconciled data is largely transferred to our reconciled data as well. This is because TDQI fails to address the role of re-exports given the available data and the ambiguity of partner attribution practices in each country. The uncertainty is extended to countries known to be active in re-exportation despite not reporting to SITC such as the Netherlands and Singapore.

3. Results and discussions

The number of countries reporting their bilateral trade increased over time. Where 77 countries reported bilateral exports to at least one partner in 1962, the number rose to 137 in 2014. For imports, the number rose from 76 to 137 over the same period. This growth has plateaued in the last one and a half decades.

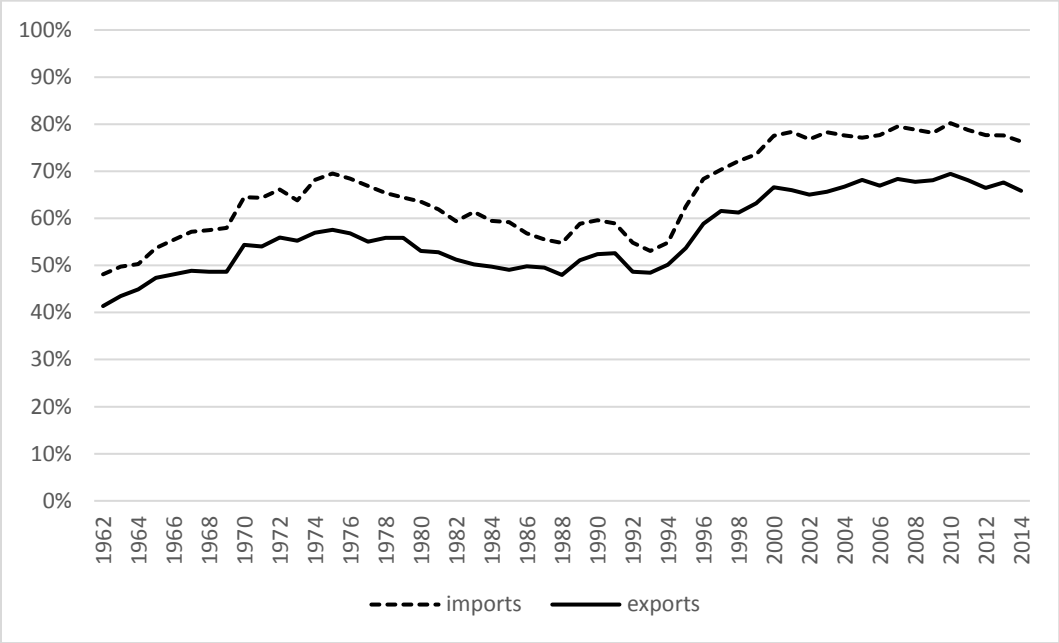
Figure 4: Number of countries reporting bilateral trade to at least one partner



Note that some of these countries cease to exist while some others became countries after 1962.

Figure 5 calculates the unweighted average of TDAI for all countries in a given year for exports and imports. TDAI for imports has been systematically higher than for exports for the whole period.

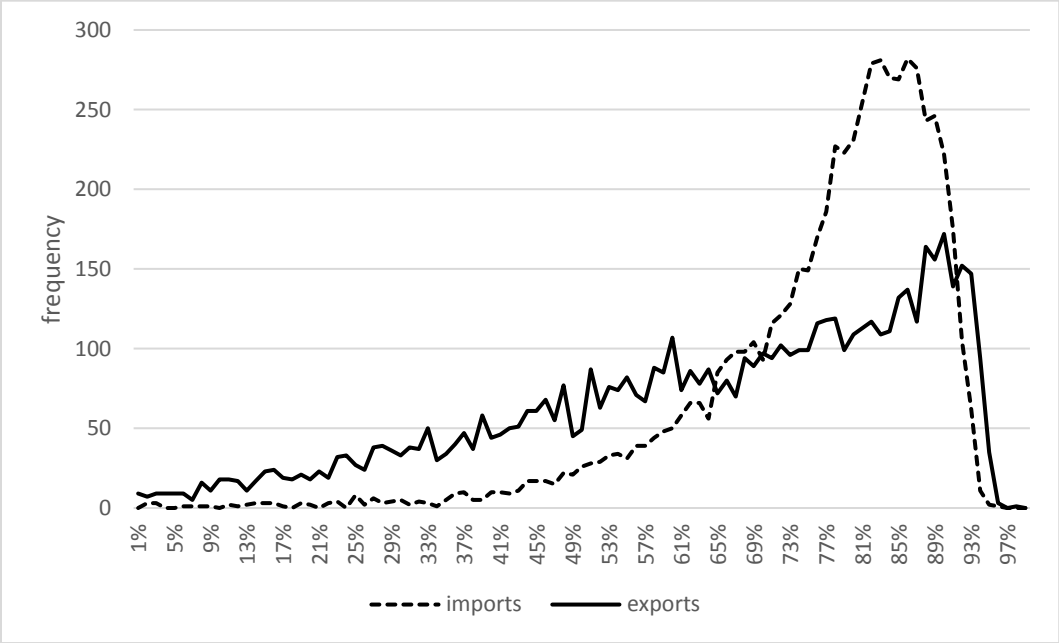
Figure 5: Annual average of TDAI for all countries



This indicates that countries generally care more about determining the origin of their imports than the destination of their exports, probably for taxing purposes. Alternatively, this could be because it is easier to identify the origin of imports than to identify the final destination of exports.

Figure 6 plots the histograms of TDQI for exports and imports for all countries and years. TDQI for exports is less left-skewed and has a lower kurtosis. Both histograms have roughly the same standard deviation of 0.36.

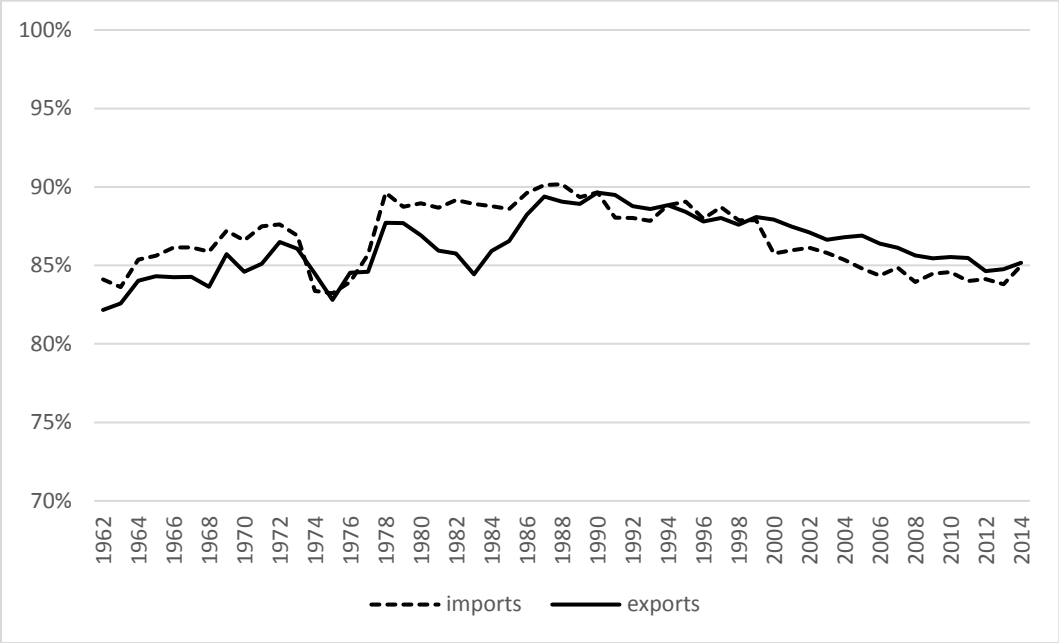
Figure 6: distribution of TDQI for all countries with non-zero TDQI (all years)



TDQI for exports averages 47% while TDQI for imports averages 56%. This difference is because TDQI is based on TDAI, where the latter is higher for imports than for exports as shown earlier. This implies that it is more likely to pick the claim reported by the importer when reconciling the data.

Average TDQI for imports and exports are roughly the same if we take the relative share of each country in global trade into account, i.e. calculating the weighted average of TDQI for all countries in the world.

Figure 7: annual average of TDQI for all countries weighted by trade shares in global trade



The trade shares used in calculating the weighted averages are using self-reported data. Using partner-reported trade shares yields very similar results.

Calculating global trade data quality using weighted average TDQI is largely in line with the results obtained using the method of lower-upper bound ratio explained in Figure 2. We, therefore, conclude that global trade data quality has been deteriorating starting from the early 1990s, which coincides with the substantial increase in global trade growth rate.

This deterioration in quality is not true using the unweighted average of TDQI for all countries in a given year. To capture this, for each country, we calculate the difference between TDQI in the last year it appeared on our list of countries and the first year. For the case of exports, 77 countries had a positive difference while 51 had a negative difference. Therefore, the quality of exports of more countries improved than worsened over time. Similarly, on the side of imports, the number of countries with quality improvement stood at 76 compared with 55 negative differences.

Combining these numbers with the conclusion from Figure 7 means that the countries which have worsened in terms of TDQI have increased their trade shares over time. In summary, global trade data quality has been worsening starting from the 1990s although more countries have improved than worsened over time.

To take a closer look at the results, Table 1 in Appendix A lists TDQI and TDAI for exports and imports for the last year covered in this study, 2014. The table is sorted by TDQI for exports from highest to lowest. Since corruption is one of the causes of trade data discrepancy, the table also lists the score of each country on Corruption Perceptions Index released by Transparency International every year. The higher the index, the less corrupt the country is. As stated in Table 1, this paper defines corruption as smuggling, partner misattribution, and mis-invoicing.

Table 3: Correlations between TDQI, TDAI, and Corruption Perceptions Index for the year 2014

	TDAI exports	TDAI imports	TDQI exports	TDQI imports	Corruption Perceptions Index
TDAI exports	1				
TDAI imports	92%	1			
TDQI exports	94%	82%	1		
TDQI imports	93%	96%	87%	1	
Corruption Perceptions Index	56%	46%	57%	53%	1

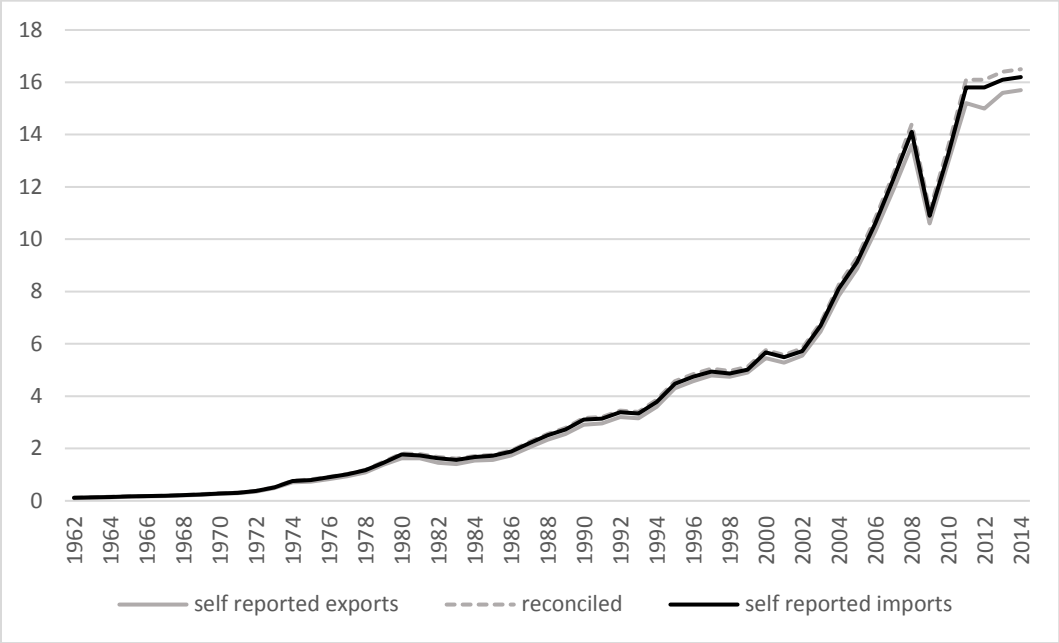
All correlations are statistically significant at 1%

Correlations between TDAI and TDQI for exports and imports are all above 80%. These strong correlations are expected as a country which maintains a good record of its exports is likely to do the same for its imports as well, and vice versa.

The correlations between TDQI for exports and imports with corruption index are 57% and 53% respectively. These positive and statistically significant correlations suggest that corruption might be largely responsible for trade data discrepancy. As shall be seen later, a causal relation might be supported by the fact that low-quality reporters tend to underreport their exports, suggesting smuggling or tax evasion. Someone can investigate the relationship between tariffs and discrepancy using commodity level data to address the causal connection properly.

Countries on top of the list in Table 1 in the Appendix are mostly developed and transparent. On the other hand, countries with low data quality are, in general, the ones where corruption is endemic and border controls are largely non-existent. These findings provide more evidence to the conclusion of Yeats (1990) in his paper: “On the accuracy of economic observations: Do sub-Saharan trade statistics mean anything?”. The answer to Yeats’s question is: no. This paper adds that this is not peculiar to Sub-Saharan countries and provides a solution for the discrepancies through reconciliation. Following the methodology stated in section 2, we use TDQI measures to reconcile international trade data. Figure 8 plots global trade using self-reported and reconciled data.

Figure 8: Global trade using self-reported and reconciled data in trillions of dollars



The data used in calculating self-reported exports and imports are summed whether the partner reported trade or not.

Almost along the whole period, reconciled data are higher than self-reported imports and exports even though the reconciled data are reported on FOB rather than CIF. This indicates that global trade is under-estimated. More apparent differences exist on a country level.

To capture how the reconciled and self-reported data differ on a country level, we calculate the ratio of the two for every total trade flow in every year. That is a country's trade with the rest of the world. We calculate the descriptive statistics of the ratio for the top and bottom 50% of countries in terms of TDQI.

Table 4: median and (standard deviation) of reconciled-unreconciled ratio

	Exports	Imports
High TDQI	1.03 (0.13)	0.99 (0.08)
low TDQI	1.23 (3.46)	0.97 (0.9)

The ratio is calculated before adjusting for CIF/FOB margin to have a clearer picture. The total number of observations is around 6100. The standard deviation is based on the median 95% of the observations (trimmed) to avoid the impact of outliers.

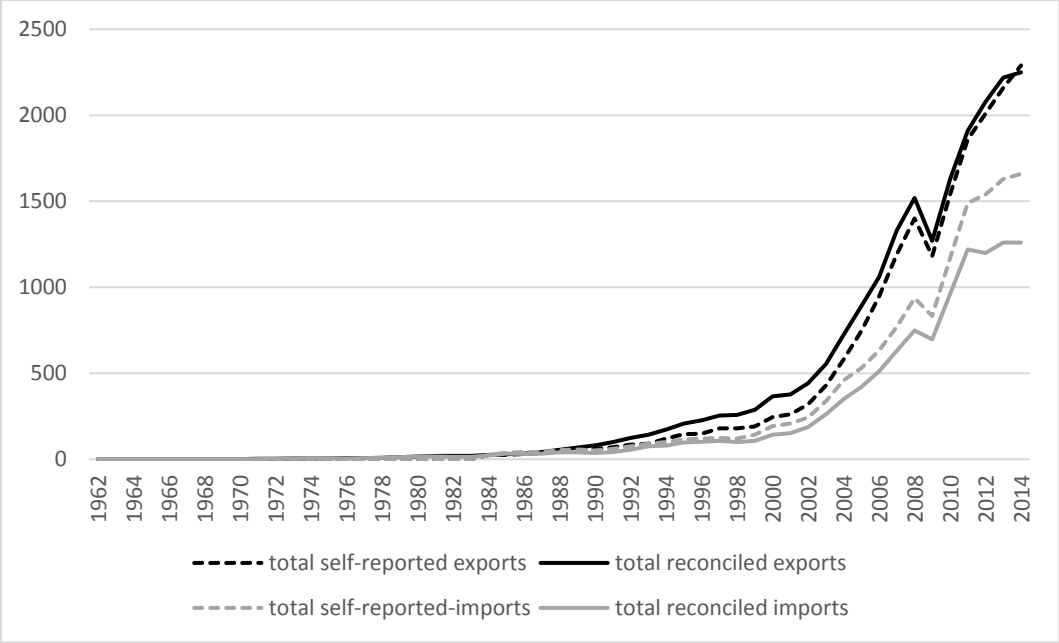
On average, high-quality reporters have their data unchanged after reconciliation. That is to say, their data are used to reconcile their partners' data. This holds for

most countries as the low standard deviations of roughly 0.1 indicate. On the other hand, low-quality reporters tend to strongly under-report their exports. Additionally, there are significant differences in the spread of the distribution of the ratio among low-quality reporters.

The fact that low-quality reporters tend to strongly under-report their exports adds more evidence that corruption can explain a share of the discrepancy as this indicates smuggling or under-invoicing. However, the large standard deviations associated with the ratio also indicate that there is an idiosyncratic component in trade data discrepancy, which implies negligence.

Using either reconciled or self-reported data, the largest two traders in the world, as of 2014, in terms of the sum of exports and imports are China followed by the US. To have a better understanding of the reconciled data we focus now on these two countries.

Figure 9: China's total exports and imports in billions of dollars



China started reporting data to SITC on 1984. Our reconciled data starts from 1962

China's exports are generally under-reported although the gap has been closing recently. This is not related to the fact that a share of China's exports goes through Hong Kong since the plotted data are for total trade, i.e. China and the rest of the world combined. The less investigated issue in the literature is China's data

discrepancy on the side of imports. China's reconciled imports are considerably smaller than self-reported corresponding values. Where FOB adjustment can explain 10% of the difference, the rest is because China tends to over-report its imports from many high-quality reporters. The following table shows the difference in claims on China's imports for the year 2014.

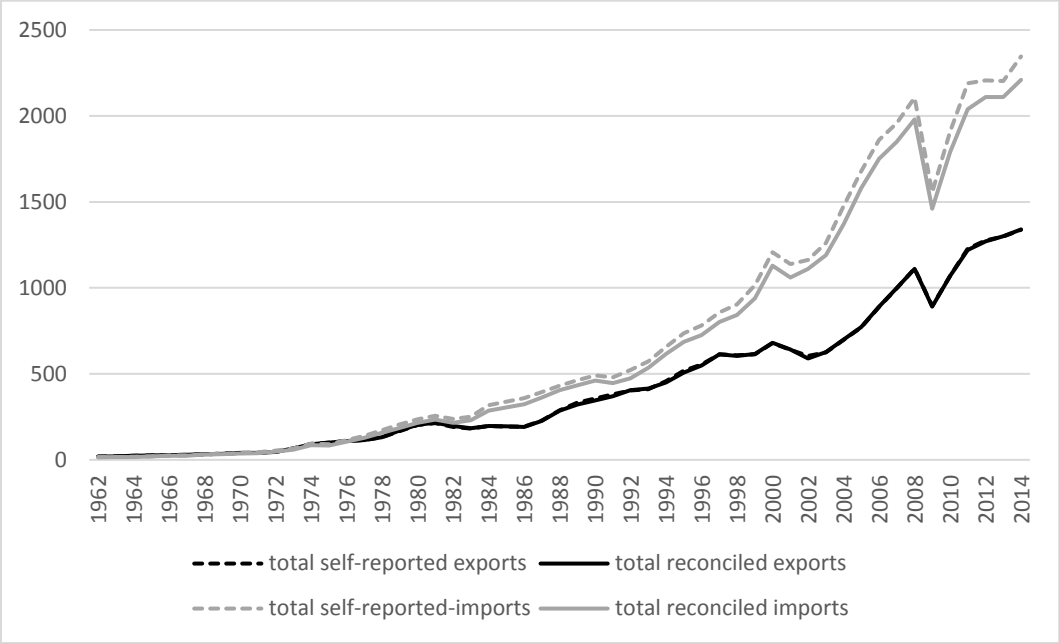
Table 5: China's claims on imports vs its partners' claims on exports to her in billions of dollars in 2014

Partner	partner's TDQI for exports in 2014	China's TDQI for imports in 2014	partner's claim	China's claim	reconciled bilateral trade (FOB)	more reliable reporter	partner claim/China's claim
Korea, Rep.	91%	79%	145.3	190.1	145.3	Korea, Rep.	0.76
Japan	91%	79%	126.2	162.8	126.2	Japan	0.78
United States	92%	79%	115.2	159.8	115.2	United States	0.72
Germany	95%	79%	99.2	105	99.2	Germany	0.94
Australia	82%	79%	70.1	97.7	70.1	Australia	0.72
Malaysia	86%	79%	28.2	55.7	28.2	Malaysia	0.51
Brazil	87%	79%	40.6	51.7	40.6	Brazil	0.79
Saudi Arabia	46%	79%	7.1	48.5	43.7	China	0.15
South Africa	82%	79%	8.7	44.6	8.7	South Africa	0.19
Russian Federation	74%	79%	37.4	41.6	37.5	China	0.90

The list includes the top 10 exporters to China (sorted by China's reports)

US exports remain relatively unchanged after reconciliation since US TDQI is comparatively high. That is, no reconciliation is actually required in most cases. Note that the difference between self-reported imports and reconciled imports is largely due to the conversion of imports from CIF to FOB.

Figure 10: US total exports and imports in billions of dollars



Using unreconciled data does not only affect the findings of a study that uses trade data directly as done by Bahmani-Oskooee, Hegerty and Harvey (2013). Unreconciled low-quality data is also expected to affect a study that uses trade data indirectly, say in an index that depends on trade in its construction. For example, in the following we calculate the widely used trade openness index, using both reconciled and self-reported data for the year 2014. The index is defined as (exports+imports)/GDP. We use the average of TDQI for exports and imports for each country to calculate the conditional means of trade openness.

Table 6: Trade Openness Index calculated using reconciled and unreconciled data in 2014

		average trade openness
unreconciled data	Average TDQI>0.75	0.66 (76)
	Average TDQI<0.75	0.63 (60)
reconciled data	Average TDQI>0.75	0.64 (76)
	Average TDQI<0.75	0.74 (60)

GDP data are in real USD retrieved from the World Bank. Numbers in brackets refer to the number of countries upon which the mean is based.

Trade openness for countries with high data quality is, on average, the same using reconciled or self-reported data. For the countries with lower data quality, however,

using self-reported data would result in significantly lowered trade openness. Out of 136 countries in 2014, 62 of them fall into that category. The difference of 11% in trade openness is expected to be even larger since self-reported imports tend to be based on CIF while reconciled data are based on FOB.

4. Conclusion

The issue of trade data discrepancy has long been ignored, which has a profound impact on the reliability of economic analysis. This study constructed an index that measures the quality of trade data using a multi-mirror technique in which a country's bilateral data are compared with the data reported by all of its partners to measure the similarity. This index is used in producing a reconciled, higher quality, set of international trade data. The results are available for public access using the following link: https://goo.gl/cGVXDk

The main shortcoming of the data reconciled in this paper is the uncertainty surrounding the reconciled exports of countries engaged in re-exportation.

This research can be extended in three different directions. First, to calculate relation-specific CIF/FOB margins using commodity level data rather than the unanimous 10% used in this paper. This increases the accuracy of reconciled data. Second, to use commodity-level trade and tariff data to infer to what extent corruption can cause data discrepancy. Third, to measure TDQI on commodity level and compare it with the results reported in this paper to capture the role of commodity misclassification in data discrepancy.

Acknowledgment

I am thankful to Victoria Doctoral Scholarship for supporting my research. This work has also benefited from the technical help and comments of Mostafa Karimi, Jonathan Chiu, Robert Kirkby, and Harold Cuffe. I would like to also thank Mark Gehlhar from GTAP project for his helpful comments.

References

2011. *International Merchandise Trade Statistics- Concepts and Definitions* (Department of Economic and Social Affairs, Statistical Division, United Nations, No. 52, Series M, New York).

Bahmani-Oskooee, Mohsen, Scott Hegerty, and Hanafiah Harvey, 2013. Exchange-rate sensitivity of commodity trade flows: Does the choice of reporting country affect the empirical estimates?, *The Journal of International Trade & Economic Development* 22, 1183-1213.

Calderon, Cesar, Alberto Chong, and Ernesto Stein, 2007. Trade intensity and business cycle synchronization: Are developing countries any different?, *Journal of international Economics* 71, 2-21.

Ferrantino, Michael J, Xuepeng Liu, and Zhi Wang, 2012. Evasion behaviors of exporters and importers: Evidence from the US–China trade data discrepancy, *Journal of international Economics* 86, 141-157.

Gehlhar, Mark, 1996. Reconciling bilateral trade data for use in GTAP, *GTAP Technical Papers*, 11.

Gujarati, Damodar, 2004. *Basic Econometrics, Fourth Edition* (McGraw Hill, New York, United States).

"The Harmonized Commodity Description and Coding System", 2014. http://www.wcoomd.org/en/topics/nomenclature/overview/what-is-the-harmonized-system.aspx.

Yeats, Alexander J, 1990. On the accuracy of economic observations: Do sub-Saharan trade statistics mean anything?, *The World Bank Economic Review* 4, 135-156.

Appendix A

Table 1: TDAI and TDQI for exports and imports as well as Corruption Perceptions Index for 2014

Reporter	TDAI exports	TDAI imports	TDQI exports	TDQI imports	Corruption perceptions index
Germany	100%	100%	95%	88%	79
Canada	100%	100%	95%	88%	81
Italy	100%	100%	95%	93%	43
Spain	100%	100%	94%	90%	60
France	100%	100%	94%	84%	69
Poland	100%	100%	94%	83%	61
Austria	100%	100%	94%	85%	72
Bulgaria	99%	97%	94%	83%	43
United Kingdom	100%	100%	93%	93%	78
Portugal	100%	98%	93%	90%	63
New Zealand	99%	100%	93%	84%	91
Finland	100%	99%	93%	89%	89
Greece	99%	99%	92%	84%	43
Slovenia	99%	100%	92%	88%	58
Mexico	99%	100%	92%	87%	35
United States	100%	100%	92%	90%	74
Denmark	100%	100%	92%	95%	92
Hungary	100%	99%	92%	93%	54
Czech Republic	100%	100%	91%	84%	51
Sweden	100%	100%	91%	92%	87
Romania	99%	100%	91%	95%	43
Korea, Rep.	100%	100%	91%	79%	55
Argentina	98%	98%	91%	87%	34
Japan	100%	100%	91%	78%	76
Thailand	100%	100%	90%	81%	38
Indonesia	99%	100%	89%	85%	34
Switzerland	100%	100%	89%	86%	86
Ireland	100%	100%	89%	88%	74
Ukraine	98%	100%	88%	89%	26
Serbia, FR(Serbia/Montenegro)	96%	100%	87%	76%	41
Brazil	99%	100%	87%	93%	43
Slovak Republic	98%	100%	87%	79%	50
Sri Lanka	98%	98%	87%	77%	38
Netherlands	100%	100%	86%	87%	83
Pakistan	97%	98%	86%	66%	29
China	100%	100%	86%	79%	36
India	100%	100%	86%	77%	38
Belgium	100%	100%	86%	90%	76
Malaysia	99%	99%	86%	80%	52

Reporter	TDAI exports	TDAI imports	TDQI exports	TDQI imports	Corruption perceptions index
Luxembourg	99%	99%	86%	86%	82
Norway	100%	99%	85%	81%	86
Peru	94%	100%	85%	91%	38
Chile	93%	99%	85%	92%	73
Dominican Republic	91%	98%	85%	90%	32
Qatar	91%	96%	84%	78%	69
Estonia	96%	100%	84%	87%	69
Morocco	93%	99%	84%	81%	39
Israel	99%	99%	84%	79%	60
Croatia	95%	100%	83%	92%	48
Turkey	99%	100%	83%	88%	45
Macedonia, FYR	90%	99%	83%	79%	45
South Africa	100%	100%	82%	85%	44
Nepal	88%	100%	82%	79%	29
Philippines	100%	97%	82%	69%	38
Tunisia	92%	100%	82%	88%	40
Australia	98%	98%	82%	91%	80
Oman	89%	92%	81%	64%	45
Latvia	100%	98%	81%	84%	55
Egypt, Arab Rep.	96%	96%	81%	80%	37
Lebanon	94%	99%	78%	91%	27
Colombia	95%	98%	77%	90%	37
Vietnam	87%	87%	77%	75%	31
Jordan	91%	96%	76%	76%	49
Madagascar	84%	98%	76%	68%	28
Belarus	98%	95%	75%	81%	31
Bosnia and Herzegovina	81%	99%	75%	77%	39
Uruguay	91%	99%	74%	86%	73
Russian Federation	98%	100%	74%	89%	27
Lithuania	99%	96%	74%	78%	58
Ecuador	88%	99%	74%	86%	33
Singapore	100%	100%	74%	82%	84
Nigeria	89%	97%	73%	81%	27
Guatemala	82%	96%	73%	80%	32
Malta	94%	97%	72%	72%	55
El Salvador	78%	98%	70%	79%	39
Paraguay	89%	98%	69%	77%	24
Cote d'Ivoire	84%	97%	68%	86%	32
Bolivia	74%	98%	68%	75%	35
Nicaragua	77%	100%	67%	85%	28
Algeria	79%	100%	67%	90%	36
Uganda	83%	98%	66%	55%	26

Reporter	TDAI exports	TDAI imports	TDQI exports	TDQI imports	Corruption perceptions index
Mauritius	75%	99%	66%	84%	54
Cameroon	74%	96%	65%	82%	27
Tanzania	89%	100%	64%	67%	31
Bahrain	76%	100%	63%	52%	49
Iceland	75%	95%	62%	79%	79
Cambodia	84%	89%	62%	63%	21
Guyana	70%	88%	59%	77%	30
Mongolia	60%	97%	59%	87%	39
Montenegro	80%	99%	58%	87%	42
Georgia	74%	88%	57%	75%	52
Azerbaijan	79%	94%	57%	69%	29
Senegal	83%	95%	56%	76%	43
Guinea	71%	94%	55%	75%	25
Namibia	84%	95%	55%	77%	49
Botswana	78%	100%	54%	94%	63
Fiji	66%	96%	54%	66%	
Costa Rica	84%	100%	53%	87%	54
Moldova	76%	99%	53%	76%	35
Kazakhstan	87%	100%	53%	88%	29
Armenia	71%	98%	52%	69%	37
Cyprus	93%	96%	51%	79%	63
Albania	71%	100%	50%	86%	33
Malawi	71%	98%	48%	57%	33
Honduras	69%	93%	47%	72%	29
United Arab Emirates	99%	100%	47%	84%	70
Brunei	52%	85%	47%	63%	
Saudi Arabia	91%	97%	46%	84%	49
Jamaica	55%	86%	46%	78%	38
Benin	73%	97%	41%	68%	39
Maldives	50%	90%	40%	62%	
Mozambique	76%	97%	40%	64%	31
Andorra	72%	92%	38%	79%	
Congo, Rep.	72%	87%	37%	63%	23
Yemen	66%	97%	35%	80%	19
Togo	59%	87%	34%	38%	29
Kuwait	83%	100%	34%	84%	44
Suriname	68%	88%	33%	79%	36
Mauritania	50%	93%	33%	42%	30
Angola	36%	98%	32%	75%	19
Niger	53%	92%	31%	61%	35
Belize	32%	90%	30%	74%	
Panama	53%	74%	30%	38%	37

Reporter	TDAI exports	TDAI imports	TDQI exports	TDQI imports	Corruption perceptions index
Hong Kong, China	98%	100%	27%	86%	74
Zambia	74%	93%	23%	73%	38
Burkina Faso	72%	93%	21%	62%	38
Samoa	30%	93%	20%	65%	52
Burundi	57%	86%	20%	49%	20
Zimbabwe	61%	98%	17%	63%	21
Rwanda	69%	99%	16%	61%	49
Gambia, The	39%	90%	12%	43%	29
Central African Republic	20%	73%	11%	44%	24
Greenland	12%	96%	10%	86%	
Sierra Leone	40%	70%	10%	27%	31
Iraq	13%	73%	10%	54%	16
Macao	33%	95%	7%	58%	
Afghanistan	5%	7%	3%	4%	12
Bangladesh	0%	0%	0%	0%	25
Bhutan	0%	0%	0%	0%	65
Chad	0%	0%	0%	0%	22
Comoros	0%	0%	0%	0%	26
Congo, Dem. Rep.	0%	0%	0%	0%	22
Cuba	0%	0%	0%	0%	46
Djibouti	0%	0%	0%	0%	34
East Timor	0%	0%	0%	0%	28
Equatorial Guinea	0%	0%	0%	0%	
Eritrea	0%	0%	0%	0%	18
Gabon	0%	0%	0%	0%	37
Ghana	0%	0%	0%	0%	48
Guinea-Bissau	0%	0%	0%	0%	19
Haiti	0%	0%	0%	0%	19
Iran, Islamic Rep.	0%	0%	0%	0%	27
Kenya	0%	0%	0%	0%	25
Korea, Dem. Rep.	0%	0%	0%	0%	8
Kyrgyz Republic	0%	0%	0%	0%	27
Lao PDR	0%	0%	0%	0%	25
Lesotho	0%	0%	0%	0%	49
Liberia	0%	0%	0%	0%	37
Libya	0%	0%	0%	0%	18
Mali	0%	0%	0%	0%	32
Myanmar	0%	0%	0%	0%	21
Papua New Guinea	0%	0%	0%	0%	25
Somalia	0%	0%	0%	0%	8
South Sudan	0%	0%	0%	0%	15
Sudan	0%	0%	0%	0%	11

Reporter	TDAI exports	TDAI imports	TDQI exports	TDQI imports	Corruption perceptions index
Swaziland	0%	0%	0%	0%	43
Syrian Arab Republic	0%	0%	0%	0%	20
Tajikistan	0%	0%	0%	0%	23
Turkmenistan	0%	0%	0%	0%	17
Uzbekistan	0%	0%	0%	0%	18
Vanuatu	0%	0%	0%	0%	
Venezuela	0%	0%	0%	0%	19

Appendix B

The list of dropped reporters along the whole period of the study is:

Aruba, Anguilla, Netherlands Antilles, Antigua and Barbuda, Bahamas, The, Bermuda, Barbados, Cook Islands, Cape Verde, Cayman Islands, Dominica, Faeroe Islands, Guadeloupe, Grenada, Kiribati, Saint Kitts-Nevis-Anguilla-Aru, St. Kitts and Nevis, St. Lucia, Montserrat, Martinique, Mayotte, New Caledonia, Other Asia, nes, Palau, Occ.Pal.Terr, French Polynesia, Reunion, Solomon Islands, Saint Pierre and Miquelon, Sao Tome and Principe, Seychelles, Turks and Caicos Isl., Tonga, Trinidad and Tobago, Tuvalu, St. Vincent and the Grenadines, Virgin Islands (U.S.), Wallis and Futura Isl., and Ethiopia (excludes Eritrea). Note that we include Ethiopia (including Eritrea).

The list of dropped partners along the whole period of the study is:

Anguilla, Netherlands Antilles, Antigua and Barbuda, Bahamas, The, Bermuda, Barbados, Bunkers, Curaçao, Cayman Islands, Dominica, Free Zones, Gibraltar, Guadeloupe, Grenada, St. Kitts and Nevis, St. Lucia, Montserrat, Martinique, Other Asia, nes, Sao Tome and Principe, Turks and Caicos Isl., Trinidad and Tobago, Unspecified, St. Vincent and the Grenadines, British Virgin Islands, World, Marshall Islands, Special Categories, Saint Pierre and Miquelon, Holy See, Aruba, American Samoa, Antarctica, Fr. So. Ant. Tr, Bouvet Island, Cocos (Keeling) Islands, Cook Islands, Cape Verde, Christmas Island, Western Sahara, Falkland Island, Faeroe Islands, Micronesia, Fed. Sts., Guam, Heard Island and McDonald Isla, British Indian Ocean Ter., Kiribati, Northern Mariana Islands, Mayotte, New Caledonia, Norfolk Island, Niue, Nauru, Pitcairn, Palau, Occ.Pal.Terr, French Polynesia, Reunion, South Georgia and the South Sa, Saint Helena, Solomon Islands, San Marino, Seychelles, Tokelau, Tonga, Tuvalu, United States Minor Outlying I, Virgin Islands (U.S.), Wallis and Futura Isl., Fm Rhod Nyas, Neutral Zone, Us Msc.Pac.I, Saint Kitts-Nevis-Anguilla-Aru, Fm Panama Cz, Ryukyu Is,

Pacific Islands, Saint Barthélemy, Bonaire, Saint Maarten (Dutch part), Br. Antr. Terr, SIKKIM, and Ethiopia (excludes Eritrea). Note that we include Ethiopia (including Eritrea).