

Quality Assessment of Trade Data in Malaysia

Dhakhir Abbas Ali^a

Fuadah Johari^b

Mohammad Haji Alias^c

Universiti Sains Islam Malaysia

Abstract: The quality of trade data is essential for any empirical study that uses them to construct a variable representing trade. However, previous studies have largely ignored the issue of trade data discrepancy. This discrepancy occurs when a country's claim on sectoral bilateral trade is different from the value reported by its trade partner. Using a size ratio and correlation coefficient computed from the values reported by both parties of bilateral trades, this study examines trade data discrepancy in 20 trade sectors from 1987 to 2016 between Malaysia and its three major Asian trade partners, i.e. China, Singapore and Japan. We detect alarming discrepancy in Malaysia's trade data. We also find that this discrepancy depends on the trade partner, time and trade sector. This study calls for attention of Malaysian policymakers regarding the possibility of customs mismanagement. We recommend that future empirical studies utilising trade in the analysis to either reconcile the data or to deploy both data sources in the econometric analysis as a robustness check.

Keywords: Data discrepancy, Malaysia, sectoral data, trade

JEL classification: F10, F13, F14

1. Background of the Study

The economy of Malaysia is highly promising. The government launched the New Economic Policy (NEP) in 1971, which was succeeded by the National Development Policy (NDP) in 1990 introduced by then prime minister Mahathir Mohamad and later revised again as the New Economic Model (NEM) in 2010 by then prime minister Najib Tun Razak (Athukorala, 2010). The main goal of the NEM is for Malaysia to achieve high income status by 2020. The NEM envisions doubling GDP per capita by mainly motivating the private sector and encouraging foreign direct investment. This New Economic Model aims to move the Malaysian economy into higher value-added activities in services and industries (Chin, 2010). Malaysia will need better skills, higher competition, a leaner public sector, greater efforts to ensure environmental

^a Faculty of Economics and Muamalat, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia. Email: Dhakhir.a.ali@gmail.com (Corresponding author)

^b Faculty of Economics and Muamalat, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia. Email: fuadah@usim.edu.my

^c Faculty of Economics and Muamalat, Universiti Sains Islam Malaysia Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia. Email: mohdhjalias@usim.edu.my

sustainability, a better knowledge base, and smarter cities, in order to achieve the envisioned NEM (Bahmani-Oskooee & Harvey, 2010).

Another sign of the comparative strength of the Malaysian economy is its rapid economic growth. In each decade from 1961 to 2016, the economy’s average economic growth amounted to 6.5 percent, 7.9 percent, 6.0 percent, 7.2 percent, 4.6 percent and 5.0 percent, respectively (World Bank, 2016). According to figures obtained from the World Bank from 1987 to 2016 as shown in Figure 1, if someone looks at the GDP growth rate in trying to perceive how Malaysia’s economy performed in comparison with the average for the whole world, it can be noticed that Malaysia has performed better in 28 out of 30 years (World Bank, 2016).

As can be seen from the figure, Malaysia’s gross domestic product growth rates have been subjected to multiple dips with two lowest points. First, the Asian financial crisis (1997-1998), which was initially caused by the free float of the Thai baht (Corsetti, Pesenti, & Roubini, 1999). Malaysia recovered from this crisis faster than many of the other countries hit by the same crisis through fixing the exchange rate and imposing strict capital controls over its financial system (Athukorala, 2010). Second, the global financial crisis, which started in the US and spread throughout the world. This crisis was triggered by the bursting of a speculative bubble in the US housing market in 2008 (Athukorala, 2010). However, Malaysia also managed to recover from this crisis through monetary expansion, characterised by low inflation, strong balance of payments, and a healthy banking system. A noteworthy feature of the adjustment process was the remarkable stability of the exchange rate, following a mild depreciation in the first two quarters of 2009 (Goh, Lim, & Sua, 2012).

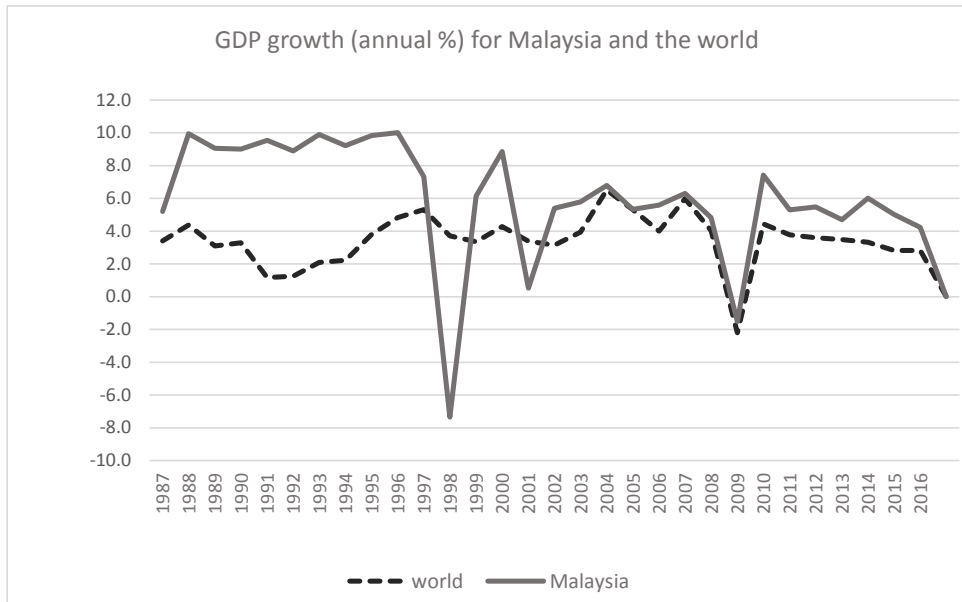


Figure 1. Malaysia’s economic growth compared to the world’s average
 Source: World Development Indicators, World Bank (2016).

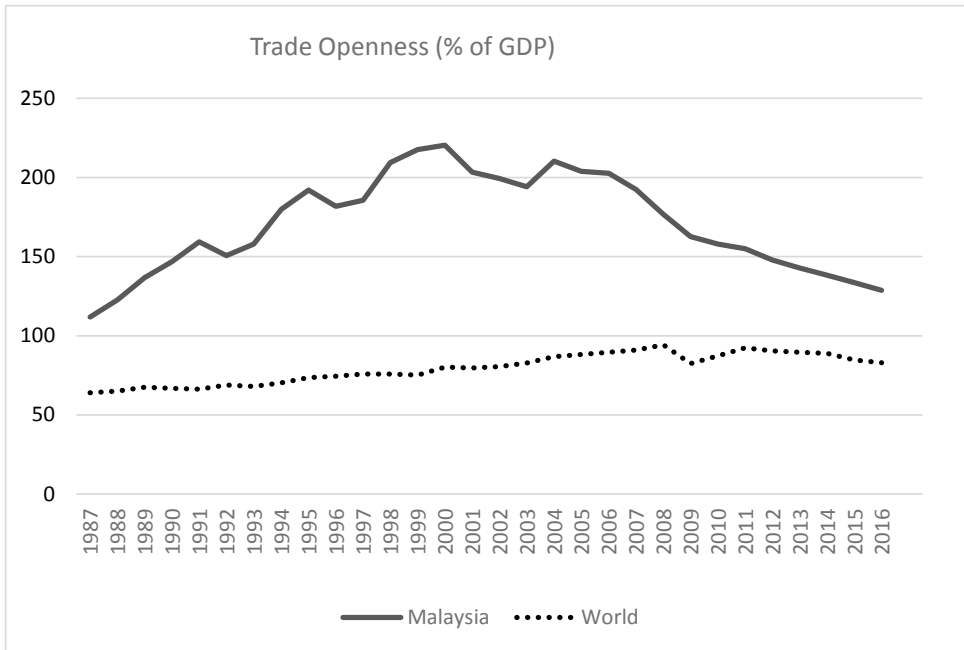


Figure 2. Trade openness in Malaysia (Malaysia's data)

Source: *World Development Indicators*, World Bank (2016).

Malaysia's strong economic growth is largely owed to trade as reflected by its trade openness. This index is defined as the sum of exports and imports of a country divided by its gross domestic product. Therefore, the index is unit free and comparable among countries (World Bank, 2016). Figure 2 depicts Malaysia's trade openness for the period of this study and compares it with the average of the world.

As clearly reflected, Malaysia's trade openness is very high compared with other countries. In fact, out of 248 countries and entities classified by the World Bank's *World Development Indicators* database, Malaysia was ranked as the world's 19th most open trading economy for the year 2016. This very clearly illustrates that Malaysia's policymakers are highly interested in stimulating an export-driven growth of the economy, which also shows the centrality of trade to economic growth.

For analysing Malaysia's aggregate trade at a glance, Figure 3 depicts the total imports and exports annually, for the period 1987-2016 (World Bank, 2016). As the figure clarifies, Malaysia has run trade balance surpluses almost throughout the whole period of 1987-2016. It is clear that the trade sector was not badly affected by the Asian financial crisis. Even imports and exports that began to slow down by the end of 2008 due to the global financial crisis, started to recover by mid-2009.

To shed light on Malaysia's trade partners, Figure 4 ranks trading partners in their order of significance for the year 2016. They are sorted based on the sum of imports and exports with Malaysia. The data is reported by Malaysia (SITC, Rev.1).

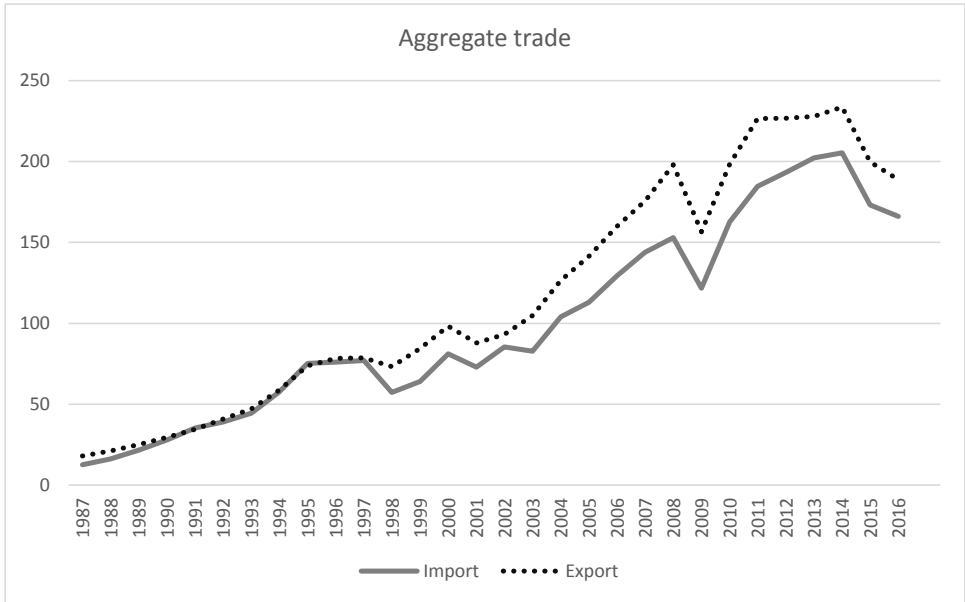


Figure 3. Malaysia’s total annual exports and imports (Malaysia’s data) (USD billion)
 Source: SITC, Revision one, World Integrated Trade Solution, <https://wits.worldbank.org>



Figure 4. Malaysia’s key trading partners in 2016 (Malaysia’s data)
 Source: SITC, Revision one, World Integrated Trade Solution, <https://wits.worldbank.org>

Based on this methodology of ranking Malaysia’s trade partners as explained above, it can be observed that Malaysia’s top ten trading partners are responsible for as high as 69.5 percent of its overall trade with the whole world. Furthermore, the three major Asian trade partners are only responsible for nearly 36.8 percent. It should be stated, however, that the differences among the top three trade partners are not very substantial with 16.2, 12.5 and 8.1 percent for trade with China, Singapore and Japan, respectively. To have a better perception of Malaysia’s trade with each of its three major trade partners, Figures 5, 6 and 7, illustrate Malaysia’s annual exports and imports with each of the three partners, respectively. The figures are reported by Malaysia in billions of USD.

From Figures 5, 6 and 7, the following key findings can be drawn. First, it was not until the early years of the 21st century that China started to play a pivotal role in Malaysia’s external trade. Yet, in nearly one decade, China managed to become Malaysia’s top trading partner. Second, there seems to be a very similar relation between Malaysia’s trade with each of Singapore and Japan. This is reflected by the fact that trade with both countries tends to move in the same direction over time. This similarity indicates that both partners are largely governed by the wider international global economic system. For example, trade with each of Singapore and Japan was clearly affected by the Southeast Asian financial crisis in the late 20th century; however, this does not seem to be applicable for the case of China. The same also applies to some extent to the global financial crisis.



Figure 5. Malaysia’s annual imports and exports to China (Malaysia’s data) (USD billion)
 Source: SITC, Revision one, World Integrated Trade Solution, <https://wits.worldbank.org>

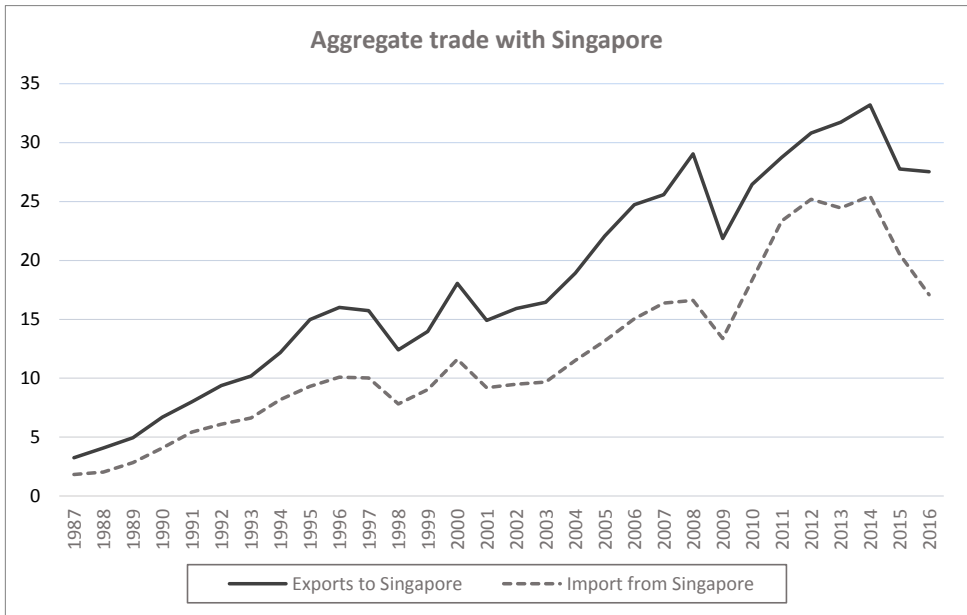


Figure 6. Malaysia’s annual imports and exports to Singapore (Malaysia’s data) (USD billion)
 Source: SITC, Revision one, World Integrated Trade Solution, <https://wits.worldbank.org>

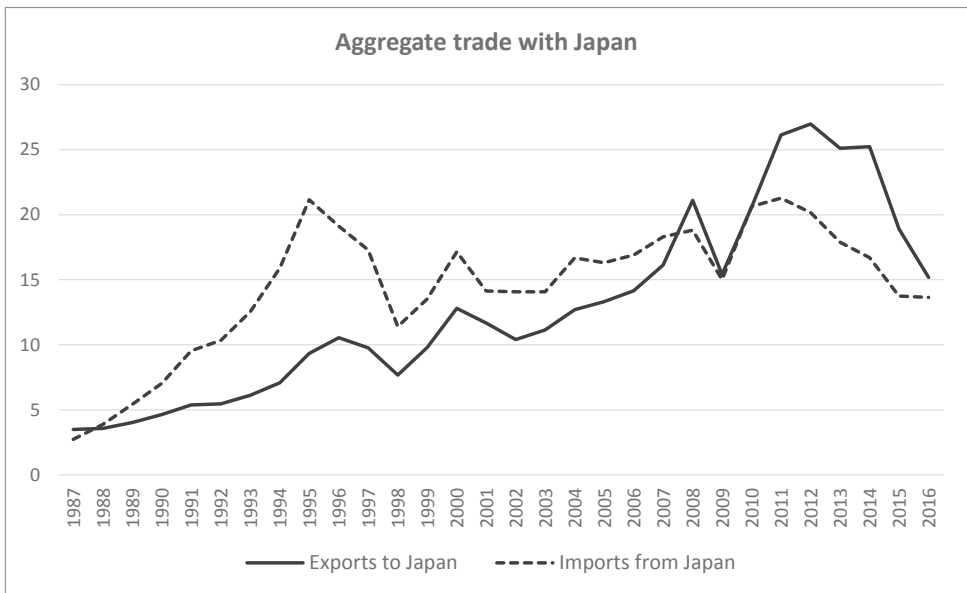


Figure 7. Malaysia’s annual imports and exports to Japan (Malaysia’s data) (USD billion)
 Source: SITC, Revision one, World Integrated Trade Solution, <https://wits.worldbank.org>

In theory, exports originate from country (y) to country (x) and the value of this trade flow is registered by both countries, in country (y) as exports and in country (x) as imports. Hypothetically, the two figures are supposed to be identical. Unfortunately, this is hardly the case (Ferrantino, Liu, & Wang, 2012). The reasons behind such asymmetry in trade data is caused by a combination of many factors, such as the timing of shipment, shipping and insurance costs, classification of goods, re-exports, and so on (Ferrantino & Wang, 2008).

However, given the centrality of trade to Malaysia's economy as illustrated earlier and the crucial dependence of economists on trade data to answer almost any economic query, to what extent should we trust Malaysia's trade data? The central question that this paper tries to address is whether the trade data presented and discussed above is accurate or not. In other words, we aim to check whether Malaysia's claims on trade with its partners equate what its partners claims on trade with it. To achieve this research objective, the study focuses on Malaysia's trade with its three major trade partners, i.e. China, Singapore and Japan to save space.

After this preliminary section, section 2 illustrates the theoretical causes of the problem of trade data discrepancy. Section 3 introduces the measures used to assess the strength of discrepancy. Section 4 presents the empirical findings, while the last section summarises the study and provides some crucial recommendations for policymakers and academicians.

2. Possible Drivers of Trade Data Discrepancy

The reasons why trade figures reported by exporters differ from importers' figures for the typical trade flow may include, at the very least, the following.

2.1 Time of Shipment

Suppose a shipment sails out from Malaysia on December 2013 and reaches China on January 2014, one would think that a higher export is reported in 2013, while a higher import is recorded in 2014. If the shipment is of considerably high value, it could create an issue for the monthly data. For annual data, the impact is assumed to be negligible due to balancing of the differences between the start and the end of a year, unless trade growth is enormous and shipping time lag is lengthy (Hamanaka, 2012).

2.2 Official Definitions of Exports and Imports

During evaluation of the shipment value by customs, three common internationally accepted guidelines are considered. The first is Free Alongside Ship (FAS), which allows registering of the goods' actual cost before they are loaded to the transport means. The second is Free on Board (FOB), which allows registering of goods on FAS as well as adding the costs of loading on-board to the transport means. Finally, Cost, Insurance and Freight (CIF) involves the value of goods on FOB, and allows adding freight and insurance costs (Ferrantino & Wang, 2008; West, 1995).

Usually, imports are registered by countries on CIF, while exports on either FOB or FAS (Martin, 2013). Freight and insurance costs need to be taken into account for

comparing the mirrored trade flows. A systemic reason for trade data discrepancy could result due to differences in evaluation methods pertaining to goods (Ferrantino & Wang, 2008).

Thus, should there be a larger value recorded by the importing country compared to the partner's corresponding export data, or vice versa, that discrepancy is likely caused by the FOB/CIF margin (Ferrantino & Wang, 2008).

2.3 Third Country Effect

It is broadly thought that it is more likely for importers to know about the country of origin compared to exporters who tend to know more about the final destination country (Ferrantino & Wang, 2008). Should this rule be applicable, one could expect bigger importer's data compared to the exporter's corresponding data. These days, this rule does not necessarily stand true for all. In some trade flows, there are chances where the exporter knows more about the final destination than the importer does about the country of origin. In such a scenario, the exporter's data will likely be greater (Ferrantino & Wang, 2008).

As per the recommendation of the UN Statistical Department, a country should attribute its imports to the country of origin each time, while recording of exports in the country of origin needs to be done as exports to the final known destination. However, under UN mandate, re-exports can be described as goods that are imported by a country (Y) from the country of origin (X), which are meant for re-exporting to the country of final destination (Z) without any significant alteration to the nature of the goods (UN, 2011). 'Goods in transit' can be defined as shipments transported from X to Z via country Y, without shipment unloading, except for substituting the means of transport (Ferrantino & Wang, 2008).

However, in spite of the above-mentioned definitions, in bilateral trade, the role of a third country is still somewhat vague. Exporting of a shipment from X to Y to Z could create a list of possibilities, none of which aligns to the UN recommendations:

- X reports Y as the final destination
- Z reports Y as the country of origin
- X & Y report the shipment as exports to C
- The goods are stored in Y for a span of time, and sold later on to another nation
- Y adds a mark-up to the shipment value prior to re-exporting to Z

The third country remains a growing phenomenon even though several attempts have been made for reconciling international trade data. For many major traders, including the European Union, lack of readily available re-export data discourages the reconciliation attempts (Guo, 2009).

2.4 Misclassification of Goods

In terms of sector-level trade data, it may happen that a shipment is imputed to various groups pertaining to an international commodity classification system applicable for each country even when values are typical and accredited to the true partner, which

were registered by both sides of a bilateral trade. Suppose a good like medical dressings is exported by a Chinese producer to Malaysia, the good would be categorised by China's customs in SITC rev.1 under 54199 'other pharmaceutical goods', while Malaysia would categorise it under 54191 'bandages, etc. coated/impregnated with pharm.pro'. This issue is created only when specific goods or sectors, such as the case in this study, are evaluated by researchers (Ferrantino & Wang, 2008).

2.5 Intentional Mis-Invoicing and Misattribution

In mis-invoicing, the value of a shipment is declared to be either higher (over-invoicing) or lower (under-invoicing) when compared with the real value. Traders can receive numerous incentives to mis-invoice their shipments, e.g., tax and tariff evasion (Hamanaka, 2012).

Mis-attribution occurs when deliberate incorrect declarations are made by traders pertaining to a shipment's origin or destination. The incentives associated with these actions may include reaping benefits of special tariff schemes specific to countries, which allow the products for a specific destination or origin certain drawbacks or duty reductions (Hamanaka, 2012). Discrepancies are created in trade data due to mis-invoicing and mis-attribution for both geographical value and attribution (Ferrantino & Wang, 2008).

2.6 Smuggling

Smuggling can be regarded as the most extreme case of under-invoicing, in which a transaction occurs without being registered. Exporters and/or importers can be involved in smuggling activities. On top of receiving benefits and incentives due to under-invoicing, smuggling can also arise due to a ban of certain types of trade applicable to some countries, for example, pornography, arms, intellectual property, certain drugs, antiques, etc. (Ferrantino & Wang, 2008).

Vital data error persists in bilateral trade if customs on both sides fail to report smuggled goods. When the smuggling activities of goods are registered by one side in a bilateral trade, it results in discrepancy of trade data. In most cases, this occurs when goods are illegal only for one partner, due to unequal degree of law enforcement amongst partners, or when smugglers perform mis-attribution of origin or destination (Ferrantino & Wang, 2008).

3. Research Methodology

The trade data set that is employed in this study is gathered on an annual basis for over 26 years starting from 1987 to 2016. The Standard International Trade Classification (SITC), revision one, two digits level is employed for the compilation of the investigated trade sectors. Based on availability of data, the analysis includes the 20 common sectors traded between Malaysia and its three key trade partners, as amassed from the World Bank's World Integrated Trade Solution database (Appendix A). It should be noted that this study assesses trade data discrepancy for both exports and imports for each sector

Table 1. Number of investigated sectors for data discrepancy in each flow and trade partner

Trade flow	Number of sectors
Malaysia's exports to Singapore	20
Malaysia's imports from Singapore	20
Malaysia's exports to China	20
Malaysia's imports from China	20
Malaysia's exports to Japan	20
Malaysia's imports from Japan	20
Total number of flows investigated for discrepancy	120
Total number of reports, two mirrored reports to assess discrepancy	120 x 2 = 240

with each of Malaysia's partners. That is, trade data discrepancy is assessed for 240 trade flows. Table 1 clarifies the distribution of sectors and partners.

The objective of the study is achieved through using the simple statistical tools of Pearson's product-moment correlation coefficient and size ratio (Malaysia's report/partner's report). Calculating the quantities stated above allows this study to answer the questions related to:

- Trade partner effect: Does sectoral trade data discrepancy vary among different pairs of countries? (i.e., Malaysia-China, Malaysia-Singapore, Malaysia-Japan)
- Time effect: Does sectoral trade data discrepancy change over time?
- Sector effect: Does sectoral trade data discrepancy vary among sectors?

3.1 Pearson Product-Moment Correlation Coefficient

As this study is not interested in capturing the causes of data discrepancy, it is not needed to run a regression to assess the causality or dependence among variables. However, it is plausible to use Pearson's correlation function as an indication for temporal similarity among reports and the ability of one report to explain the values of the other (Derrick, Bates, & Dufek, 1994).

For example, suppose Malaysia's report pertaining to exports to China for sector i is represented by x , and y denotes China's imports from Malaysia that China had reported, which is applicable to the same sector in a typical period, Pearson's correlation coefficient for the two reports can be regarded as the covariance pertaining to the two variables that is divided by the product of their standard deviations (Lee Rodgers & Nicewander, 1988), as represented below:

$$r_{x,y} = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2 (y - \bar{y})^2}}$$

The range of coefficient of Pearson's correlation (r) lies between 1 and -1, in which 1 denotes existence of a strong positive linear association between y and x (suggesting that both reports possess high similarity), and -1 denotes the presence of a strong inverse linear association while 0 signifies no linear association (Derrick et al., 1994).

3.2 Reports Ratio

Even though the correlation coefficient can explain the similarity seen amongst reports to some extent, the question regarding differences in size cannot be answered. In the case of exports, the ratio between Malaysia's exports to country X as reported by Malaysia, to X's imports from Malaysia as reported by X for the same sector and year, is considered by the researcher to obtain a number that is unit free. Any result found to be equivalent to unity signifies identical reports with no discrepancy. However, if the figure exceeds unity, it shows overestimation by Malaysia pertaining to its exports to X (or underestimation by X pertaining to its imports from Malaysia). If the value is smaller than one, it shows overestimation by X pertaining to its imports from Malaysia (McDonald, 1985).

Inversely, the ratio between imports by Malaysia from X as reported by Malaysia, to exports by X to Malaysia as reported by X, is taken into account by the researcher to obtain a number that is not only unit free but also possesses the opposite interpretation. A number equalling 1 signifies identical reports with no discrepancy. However, should the figure exceeds 1, it denotes overestimation by Malaysia pertaining to its imports from X, and a value less than 1 signifies overestimation by X pertaining to its exports to Malaysia.

3.3 Descriptive Statistics

As the amount of data is humongous as well as distributed amongst various trade partners of Malaysia, the tools pertaining to standard deviation, descriptive statistics of arithmetic mean and frequency description that consider kurtosis and skewness are employed to summarise the results showcasing the two discrepancy measures. The summary on discrepancy provides a detailed description for the effects of trade partner, time and sector on the strength of discrepancy.

4. Empirical Results

Section 4.1 presents the results of dynamic discrepancy among Malaysia and each of its trade partners, while section 4.2 presents the results for the sector effect.

4.1 Trade Data Discrepancy: Time Effect

Time effect results are meant to assess the discrepancy of trade data as reported by Malaysia and each of her trade partners over time. The results of the statistical tools of Pearson product-moment correlation coefficient and size ratio explained in section 3 are presented in Tables 2 and 3, respectively.

Since each bilateral trade flow between Malaysia and each of her trade partners is disaggregated among 20 different sectors, each of the figures reported in the first part of Table 2 is the mean of size ratios for all 20 sectors. Based on the descriptive statistics the following conclusions can be drawn:

- On average, trade data discrepancy in Malaysia's data is most apparent in imports from China (4.41) and least prevalent in exports to Singapore (0.96).

Table 2. Annual size ratio

Year	MYS imports from CHN	MYS imports from SGP	MYS imports from JPN	MYS exports to CHN	MYS exports to SGP	MYS exports to JPN	Average
1987	1.40	0.36	3.54	0.87	1.24	1.41	1.47
1988	1.52	0.31	2.56	0.69	0.98	1.30	1.23
1989	2.15	0.38	2.25	2.53	1.04	1.20	1.59
1990	2.21	0.50	2.48	0.66	1.04	1.28	1.36
1991	1.76	0.48	1.96	2.62	0.96	1.14	1.49
1992	2.08	0.64	1.38	1.34	1.00	0.97	1.24
1993	4.17	0.45	2.07	2.84	0.77	1.02	1.89
1994	2.31	0.34	2.42	1.08	0.81	1.03	1.33
1995	1.80	0.32	1.86	0.89	0.92	1.02	1.14
1996	1.39	0.36	1.61	0.87	0.86	1.05	1.02
1997	2.21	0.37	1.42	1.15	0.85	0.95	1.16
1998	1.67	0.36	1.45	1.23	0.84	1.05	1.10
1999	3.14	0.37	1.27	0.69	0.83	1.00	1.22
2000	4.18	0.41	1.49	0.63	0.82	0.93	1.41
2001	51.14	0.41	1.59	0.70	0.82	1.02	9.28
2002	1.57	0.46	1.32	0.75	0.84	1.02	0.99
2003	2.06	0.39	1.24	0.83	0.86	1.03	1.07
2004	4.46	0.39	1.28	0.81	0.87	1.08	1.48
2005	1.99	0.43	1.32	0.88	0.88	1.04	1.09
2006	3.90	0.42	1.37	1.04	0.99	1.08	1.47
2007	2.06	0.48	1.91	0.97	1.03	1.14	1.27
2008	8.94	0.42	2.99	0.90	0.95	1.11	2.55
2009	6.29	0.51	1.48	0.92	1.05	1.30	1.93
2010	1.25	0.53	2.38	0.85	1.04	1.30	1.22
2011	1.77	0.62	1.76	0.95	1.05	1.22	1.23
2012	1.02	0.72	1.75	0.86	1.29	1.19	1.14
2013	1.73	0.52	1.98	1.20	1.17	1.21	1.30
2014	8.16	0.52	1.99	0.99	1.12	1.01	2.30
2015	1.17	0.56	1.66	1.05	0.94	1.04	1.07
2016	2.77	0.54	1.59	1.04	0.91	1.07	1.32
Descriptive statistics							
Mean	4.41	0.45	1.85	1.09	0.96	1.11	1.64
Standard deviation	9.04	0.10	0.55	0.56	0.13	0.12	1.49
Skewness	5.09	0.82	1.32	2.37	0.82	0.81	4.99
Kurtosis	26.96	0.43	1.83	4.91	0.25	-0.20	26.19

Note: MYS – Malaysia, CHN – China, SGP – Singapore, JPN – Japan.

- On average, Malaysia's exports data is much less problematic than imports. This notion applies to Malaysia's trade with each of its major three Asian trade partners.
- As the measure of standard deviation shows, discrepancy in Malaysia's imports from China is highly inconsistent from one year to another with a value of 9.04. Malaysia's imports and exports to Singapore and exports to Japan are highly consistent and no major difference exists from one year to another with values of 0.1, 0.13 and 0.12, respectively.
- The positive skewness in all trade flows indicates that the values for most years are less than the mean. This fact draws attention that the figures for mean are biased and do not reflect the reality of general discrepancy accurately due to the presence of many extreme values (Jorion, 2007). In other words, the higher the skewness, the higher the bias in the mean measure of discrepancy.
- The same fact is also supported by the high kurtosis, which also indicates the existence of extreme differences from one year to another (Jorion, 2007).
- Overall, using all four descriptive statistical measures, Malaysia's trade data discrepancy is least apparent with Japan, but more apparent with each of China and Singapore.

Each of the figures reported in the first part of Table 3 is the correlation between the values of all 20 sectors as reported by Malaysia and each of her trade partners in a year. Based on the descriptive statistics the following conclusions can be drawn:

- On average, temporal similarity is highest in Malaysia's imports from Japan and exports to Singapore with 0.99. In other words, the reports of one side can highly explain the reports of the other, i.e. the difference among reports is systematic.
- Average annual correlation is the lowest for Malaysia's exports and imports with China with 0.93 and 0.95, respectively.
- On average, there is no clear difference whether Malaysia's exports or imports are more problematic in terms of temporal similarity.
- As the measure of standard deviation shows, discrepancy in Malaysia's exports to China is highly variant with a value of 0.07.
- Malaysia's imports from Japan, as well as the imports and exports to Singapore, are highly consistent and no major difference exists from one year to another with the values of standard deviation of 0.01, 0.02 and 0.02, respectively.
- The negative skewness in all trade flows indicates that the values of most years are higher than the mean (Jorion, 2007). In other words, temporal similarity in most of the years is better than the mean.
- Positive kurtosis of Malaysia's exports indicates that most of the annual correlation values are clustered around the mean, i.e. the distribution of values has fat tails (leptokurtic distribution). This indicates as well that the values of annual correlation are more consistent from one year to another. The opposite applies to Malaysia's imports from each of its trade partners (Jorion, 2007).
- Overall, using all four descriptive statistical measures, no clear conclusion can be made regarding whether Malaysia's imports or exports are more temporally similar with the reports of its partners.

Table 3. Annual correlation

Year	MYS imports from CHN	MYS imports from SGP	MYS imports from JPN	MYS exports to CHN	MYS exports to SGP	MYS exports to JPN	Average
1987	0.83	0.95	1.00	0.98	0.94	0.77	0.91
1988	0.88	0.94	1.00	0.88	0.96	0.79	0.91
1989	0.88	0.93	1.00	0.92	0.98	0.85	0.93
1990	0.90	0.94	1.00	0.97	0.98	0.94	0.95
1991	0.94	0.93	1.00	1.00	0.98	0.97	0.97
1992	0.91	0.93	1.00	1.00	0.98	0.98	0.97
1993	0.91	0.96	1.00	0.96	0.99	0.98	0.97
1994	0.85	0.96	1.00	0.98	1.00	0.99	0.96
1995	0.90	0.98	1.00	0.96	1.00	0.99	0.97
1996	0.94	0.99	1.00	0.91	1.00	0.99	0.97
1997	0.97	0.99	1.00	0.88	1.00	0.97	0.97
1998	0.99	1.00	1.00	0.90	1.00	0.98	0.98
1999	0.99	1.00	1.00	0.93	1.00	1.00	0.98
2000	0.99	1.00	1.00	0.96	1.00	1.00	0.99
2001	0.99	1.00	1.00	0.90	1.00	0.99	0.98
2002	0.99	1.00	0.99	0.93	1.00	0.99	0.98
2003	0.99	1.00	0.98	0.96	1.00	1.00	0.99
2004	0.99	0.99	0.99	0.93	1.00	1.00	0.98
2005	0.98	0.99	0.99	0.95	1.00	1.00	0.99
2006	0.99	0.99	0.99	0.94	1.00	1.00	0.98
2007	1.00	0.99	0.98	0.93	1.00	1.00	0.98
2008	0.99	0.98	0.99	0.59	0.95	0.99	0.91
2009	0.97	0.99	0.98	0.97	0.99	0.97	0.98
2010	0.99	0.99	0.98	0.88	0.99	1.00	0.97
2011	0.98	0.99	0.98	0.94	1.00	1.00	0.98
2012	0.96	0.99	0.98	0.96	1.00	1.00	0.98
2013	0.94	0.99	0.98	0.96	1.00	1.00	0.98
2014	0.96	0.99	0.98	0.99	1.00	0.99	0.99
2015	0.97	1.00	0.99	0.98	1.00	1.00	0.99
2016	0.98	1.00	0.99	0.98	1.00	1.00	0.99
Descriptive statistics							
Mean	0.95	0.98	0.99	0.93	0.99	0.97	0.97
Standard deviation	0.05	0.02	0.01	0.07	0.02	0.06	0.02
Skewness	-1.14	-1.20	-0.39	-3.74	-2.20	-2.72	-1.75
Kurtosis	0.27	-0.19	-1.46	17.30	4.27	6.63	2.07

Note: MYS – Malaysia, CHN – China, SGP – Singapore, JPN – Japan.

4.2 Trade Data Discrepancy: Sector Effect

Sector effect results are meant to assess the discrepancy of trade data as reported by Malaysia and each of her trade partners over sectors. The statistical tools of Pearson product-moment correlation coefficient and size ratio explained in section 4.1 are presented in Tables 4 and 5, respectively.

Since each bilateral trade flow between Malaysia and each of her trade partners is reported over 26 years, each of the figures reported in the first part of Table 4 is the mean of size ratios for all years in one trade sector. Based on the descriptive statistics the following conclusions can be drawn:

Table 4. Sectoral size ratio

Sector	MYS imports from CHN	MYS imports from SGP	MYS imports from JPN	MYS exports to CHN	MYS exports to SGP	MYS exports to JPN	Average
26	1.67	0.47	1.31	0.99	1.85	0.20	1.08
28	66.87	0.86	9.06	0.46	0.59	0.25	13.01
42	1.01	0.62	2.61	0.89	1.35	0.04	1.09
51	1.28	0.53	1.21	0.84	0.91	0.12	0.81
55	1.05	0.38	1.28	0.93	0.92	0.22	0.80
58	1.16	0.61	1.40	0.51	0.92	0.30	0.82
59	0.71	0.39	1.14	0.52	0.93	0.30	0.66
62	0.84	0.36	0.99	0.52	0.86	0.13	0.62
63	1.46	0.42	3.69	0.84	0.83	0.05	1.22
64	1.05	0.57	1.20	0.63	0.83	0.14	0.74
65	1.17	0.17	1.07	0.62	0.58	0.12	0.62
66	0.78	0.47	1.15	0.75	0.81	0.82	0.80
67	1.19	0.18	1.31	2.31	1.00	0.57	1.09
68	1.91	0.36	1.38	1.22	1.01	0.32	1.03
69	0.88	0.61	1.58	1.01	0.91	0.57	0.93
71	1.35	0.35	1.23	1.02	0.81	0.28	0.84
72	1.46	0.51	1.35	0.59	0.71	0.62	0.87
73	0.68	0.42	1.06	3.62	1.77	1.28	1.47
82	0.79	0.32	1.27	3.01	0.79	0.35	1.09
89	0.89	0.44	1.61	0.61	0.79	0.18	0.75
Descriptive statistics							
Mean	4.41	0.45	1.85	1.09	0.96	0.34	1.52
Standard deviation	14.70	0.16	1.81	0.87	0.33	0.30	2.71
Skewness	4.47	0.49	3.74	2.13	1.85	1.85	4.43
Kurtosis	19.98	1.23	14.84	3.83	3.14	3.91	19.71

Note: MYS – Malaysia, CHN – China, SGP – Singapore, JPN – Japan.

- Malaysia's reports on imports from China and Japan in some sectors are very different from the reports of her partners. For example, Malaysia's reported value of imports in metalliferous ores and metal scrap (code 28) from China and Japan is 66.87 and 9.06 times higher, respectively, than the value reported by her partners for the period 1987-2016. However, this does not apply to all sectors.
- As mentioned earlier, Malaysia's imports seem to be much more troubled than her exports in terms of size discrepancy. For example, the size ratio for Malaysia's imports from China is 4.41, while it is only 1.09 for exports.
- As the measure of standard deviation reveals, discrepancy in Malaysia's imports from China is highly variant from one sector to another with a value of 14.70. It is least variant in Malaysia's imports from Singapore with 0.16.
- The positive skewness in all trade flows indicates that the values of most sectors are less than the mean.
- Positive kurtosis of Malaysia's exports and imports indicates that most of the sectoral size ratios are clustered around the mean, i.e. the distribution of values has fat tails.
- Overall, using descriptive statistics, no definite conclusion can be made regarding whether Malaysia's imports or exports are more problematic, or whether the discrepancy is more apparent with a specific trade partner.
- As the last column of the table shows, regardless of the trade partner, year, and direction, data discrepancy is highly related to some sectors rather than others.

Each of the figures reported in the first part of Table 5 is the correlation between the values for all 26 years as reported by Malaysia and each of her trade partners for a sector. The descriptive statistics reveal that:

- On average, temporal similarity among sectors differs from one trade flow to another. It is highest in Malaysia's imports from China with 0.93 and lowest in imports from Singapore with 0.77.
- Average annual correlation is the lowest for Malaysia's imports from China and exports to Japan with 0.93 and 0.92, respectively.
- As the measure of standard deviation shows, discrepancy in Malaysia's imports from Singapore is highly variant with a standard deviation coefficient of 0.30.
- The negative skewness in all trade flows and sectors shows that the values of sectoral correlation in most sectors are higher than the mean. (Jorion, 2007). In other words, temporal similarity in most of the years is better than the mean.
- Positive kurtosis of Malaysia's exports and imports indicates that most of the sectoral correlation values are clustered around the mean. This implies that the values of sectoral correlation are consistent from one sector to another (Jorion, 2007).
- Using all four descriptive statistics, it seems that Malaysia's temporal similarity is highest with China and lowest with Singapore.

Table 5. Sectoral correlation

Sector	MYS imports from CHN	MYS imports from SGP	MYS imports from JPN	MYS exports to CHN	MYS exports to SGP	MYS exports to JPN	Average
26	0.98	-0.35	0.94	0.59	0.13	0.96	0.54
28	0.70	0.97	0.66	0.72	0.83	0.67	0.76
42	0.65	0.75	0.49	0.99	0.88	1.00	0.79
51	0.99	0.98	0.92	1.00	0.97	0.98	0.97
55	0.94	0.99	0.95	0.99	0.95	1.00	0.97
58	0.99	0.99	0.98	0.96	0.99	1.00	0.99
59	0.99	0.93	0.91	0.92	0.97	0.90	0.94
62	0.95	0.90	0.88	0.99	0.95	0.94	0.94
63	0.95	0.73	0.56	0.92	0.95	1.00	0.85
64	0.98	0.67	0.97	0.86	0.99	0.98	0.91
65	0.94	0.69	0.89	0.81	0.94	0.98	0.87
66	0.91	0.89	0.95	0.99	0.99	0.55	0.88
67	1.00	0.75	0.98	0.96	0.98	0.93	0.93
68	0.98	0.44	0.99	0.55	0.92	0.95	0.80
69	0.98	0.94	0.95	0.95	0.99	0.90	0.95
71	0.99	0.93	0.98	0.96	0.98	0.93	0.96
72	0.99	0.84	0.97	0.98	0.92	0.98	0.95
73	0.96	0.73	0.95	0.95	0.54	0.93	0.84
82	0.85	0.89	0.91	0.81	0.97	0.96	0.90
89	0.96	0.80	0.74	0.87	0.93	0.95	0.88
Descriptive statistics							
Mean	0.93	0.77	0.88	0.89	0.89	0.92	0.88
Standard deviation	0.10	0.30	0.15	0.13	0.21	0.11	0.10
Skewness	-2.30	-3.08	-1.80	-1.60	-3.20	-2.67	-2.06
Kurtosis	4.68	11.13	2.25	1.84	10.66	7.01	5.65

Note: MYS – Malaysia, CHN – China, SGP – Singapore, JPN – Japan.

5. Conclusion

Trade is crucial for Malaysia's economic growth and to almost any academic study in the wide field of economics. This study aims to assess the quality of trade data in Malaysia. To this end, after a review of Malaysia's economy, the study lists the theoretical reasons for trade data discrepancy. The utilised statistical tools are able to evaluate the discrepancy over time, trade partner and trade sector.

In terms of size difference among data reports, as assessed yearly and regardless of the sector:

- Discrepancy is highest in imports from China and lowest in exports to Singapore.

- Malaysia's exports data is much less problematic than imports.
- Discrepancy in imports from China is highly inconsistent from one year to another.
- Discrepancy in imports from Singapore and exports and imports from Japan are highly consistent.
- Overall, discrepancy is least apparent with Japan.

In terms of size differences among reports, as assessed on sectors and regardless of the year:

- Discrepancy is highly sector-dependent in imports from China and Japan.
- Imports are more troubled than exports.
- Discrepancy in imports from China is highly variant from one sector to another.
- Discrepancy is least variant in Malaysia's imports from Singapore.

In terms of temporal similarity among data reports, as assessed yearly and regardless of the sector:

- Similarity is highest in Malaysia's imports from Japan and exports to Singapore.
- Annual correlation is the lowest for Malaysia's exports and imports with China.
- It is not clear whether Malaysia's exports or imports are more problematic.
- Discrepancy in Malaysia's exports to China is highly variant.
- No conclusion can be made on whether Malaysia's imports or exports are more temporally similar with the reports of its partners.

In terms of temporal similarity among data reports, as assessed on sectors and regardless of the year:

- Similarity is highest in imports from China and lowest in imports from Singapore.
- Average annual correlation is the lowest for exports and imports with China.
- Trade with China is comparatively least problematic.
- Discrepancy in Malaysia's exports to China is highly variant among sectors.
- Overall, Malaysia's temporal similarity is highest with China and lowest with Singapore.
- Above all, findings of this study regarding the high levels of trade data discrepancy should be very alarming for the government of Malaysia.

As an implication for economic policymakers in Malaysia, academic papers, such as Ferrantino et al. (2012) suggested the high discrepancy as a sign of possible customs malpractice, i.e. tax evasion or avoidance. Assuming that the Malaysian side in bilateral trade with any of China, Singapore, or Japan is responsible for the discrepancy due to mismanagement, investigating and detecting the causes of the discrepancy may save the government vast amounts of money in unpaid tariffs and taxes.

However, it is beyond the scope of this study to confirm whether Malaysia is responsible for trade data discrepancy or not. This finding shows how policymakers are dependent on the results of the literature whereby the data reported may have caused unfavourable results on the economy.

For academics, any empirical economic study using trade data in Malaysia as an independent or dependent variable may benefit from this study, which assesses the state of sectoral trade data discrepancy between Malaysia and her three major trade partners and explains the nature of this discrepancy over time, partner and sector.

This study recommends future economic studies to consider investigating trade data for discrepancy before using in empirical analysis whenever applicable. If considerable discrepancy is detected, the author may be compelled to follow one of the following two choices. First, to reconcile the data through accounting for the causes of discrepancy one by one as discussed in section 2. However, this highly desirable practice may not be always applicable due to data availability.

Therefore, researchers can follow our second recommendation of applying the concept of mutual confirmation as an indication for reliability and as a robustness check for the results of any econometric test which includes trade as one of the variable of interest. In other words, the researcher can consider the outcome of an analysis reliable only when two conditions are met: (a) the regression's coefficient estimates are statistically similar utilising one or the other data source; and (b) the estimates bear the same sign and have equal or somewhat similar magnitudes. If the outcomes are alike in utilising either report individually, the researcher can presume that the true impact of independent variable X on the dependent variable Y exists somewhere between the two outcomes. The precise impacts cannot be attained with the existence of trade data discrepancy.

References

- Athukorala, P.-c. (2010). *Malaysian economy in three crises* (Working Papers in Trade and Development, No. 2010/12). Canberra: Australian National University.
- Bahmani-Oskooee, M., & Harvey, H. (2010). The J-curve: Malaysia versus her major trading partners. *Applied Economics*, 42(9), 1067-1076. <https://doi.org/10.1080/00036840701721158>
- Chin, J. (2010). MRCB, Pos in focus after PM unveils new economic model. *The Edge*. Retrieved from <https://www.theedgemarkets.com/article/mrcb-pos-focus-after-pm-unveils-new-economic-model-0>
- Corsetti, G., Pesenti, P., & Roubini, N. (1999). What caused the Asian currency and financial crisis? *Japan and the World Economy*, 11(3), 305-373. [https://doi.org/10.1016/S0922-1425\(99\)00019-5](https://doi.org/10.1016/S0922-1425(99)00019-5)
- Derrick, T.R., Bates, B.T., & Dufek, J.S. (1994). Evaluation of time-series data sets using the Pearson product-moment correlation coefficient. *Medicine and Science in Sports and Exercise*, 26(7), 919-928.
- Ferrantino, M.J., Liu, X., & Wang, Z. (2012). Evasion behaviors of exporters and importers: Evidence from the US-China trade data discrepancy. *Journal of International Economics*, 86(1), 141-157. <https://doi.org/10.1016/j.jinteco.2011.08.006>
- Ferrantino, M.J., & Wang, Z. (2008). Accounting for discrepancies in bilateral trade: The case of China, Hong Kong, and the United States. *China Economic Review*, 19(3), 502-520. <https://doi.org/10.1016/j.chieco.2008.02.002>
- Goh, S., Lim, M.H., & Sua, T.Y. (2012). The policy responses and implications of the global financial crisis in Asia: A case study for Malaysia. *European Research Studies Journal*, 15(1), 55-70.

- Guo, D. (2009). *Mirror statistics of international trade in manufacturing goods: The case of China* (Research and Statistics Branch, Working Paper, No 19/2009) Vienna: United Nations Industrial Development Organization.
- Hamanaka, S. (2012). Whose trade statistics are correct? Multiple mirror comparison techniques: A test case of Cambodia. *Journal of Economic Policy Reform*, 15(1), 33-56. <https://doi.org/10.1080/17487870.2012.657827>
- Jorion, P. (2007). *Financial risk manager handbook* (Vol. 406). Hoboken, New Jersey: John Wiley & Sons, 31-60.
- Lee Rodgers, J., & Nicewander, W.A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1), 59-66.
- Martin, M.F. (2013). What's the difference?—Comparing US and Chinese trade data (Congressional Research Service report). Washington, DC: Congressional Research Service
- McDonald, D.C. (1985). *Trade data discrepancies and the incentive to smuggle: An empirical analysis* (Divergences dans les données sur les échanges commerciaux et attrait de la contrebande: analyse empirique) (Discrepancias de los datos sobre el comercio e incentivo al contrabando: Análisis empírico). Staff Papers (International Monetary Fund), 32(4), 668-692.
- SITC. (Rev.1). Standard International Trade Classification, Revision one. World Integrated Trade Solution, World Bank. Retrieved from <https://wits.worldbank.org>
- UN. (2011). *International merchandise trade statistics: Concepts and definitions 2010*. New York, NY: Department of Economic and Social Affairs, Statistics Division, United Nations.
- West, L.A. (1995). *Reconciling China's trade statistics* (IPC Staff paper, No. 76). Washington, DC: International Programs Center, Population Division, US Bureau of the Census.
- World Bank. (2016). *World development indicators 2016*. Washington, DC: Author.