ESSAYS ON MODERN ECONOMIC ISSUES IN INTERNATIONAL TRADE, EXCHANGE RATES AND HOUSING

By

KARAM SHAAR

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy
in Economics.

Victoria University of Wellington
2019

To my wife, Hanna, for her unconditional love...

ABSTRACT

The common theme of the three papers in this thesis is the focus on the impact of data choices on empirical research in Economics. Such choices can be about the source of data; should we source the data from country A or country B in a bilateral trade relation? Is there a way to reconcile the discrepancies in international trade data? In investigating the impact of exchange rate on trade, should we choose high-frequency or low-frequency data? What does the choice of a certain frequency imply for the econometric analysis? In assessing the impact of housing wealth on household consumption, what are the benefits of choosing household-level data? How can we take advantage of aggregate data on house prices to circumvent the endogeneity arising from household-specific confounding factors? This thesis shows that data choices can strongly affect our conclusions regarding several modern economic issues.

The first paper is titled 'Reconciling International Trade Data.' International trade data are filled with discrepancies–where two countries report different values of trade with each other. We develop an index for ranking countries' data quality based on the following notion: the more a country's reports on bilateral trade differ from the corresponding reports of its partners, the more likely it is a low-quality reporter. We calculate the comparative quality for each country's imports and exports separately for every year from 1962 to 2016. We reconcile international trade data through picking the value reported by the country with higher quality in every bilateral flow. The findings include: (a) global trade was under-reported by roughly 5% over the past five years as countries with low data quality under-report both, their imports and exports; (b) erroneous reporting is prevalent among low-quality reporters; (c) importers' data are less likely to be in error; (d) the level of development and corruption are possible determinants of trade data quality; (e) low-quality reporters are 14% more open to trade using reconciled data than using self-reported data (f) China tends to under-report its exports and over-report its imports,

while there is only a small difference between US self-reported and reconciled data. The reconciled trade dataset is made freely available for future studies to use.

The second paper is titled 'Why You Should Use High Frequency Data to Test the Impact of Exchange Rate on Trade.' The paper suggests that testing the impact of exchange rate on trade should be done using high frequency data. Using different data frequencies for identical periods and specifications between the US and Canada, the paper shows that low frequency data suppresses and distorts the evidence of the impact of exchange rate on trade in the short-run and the long-run.

The third paper is titled: 'Housing Leverage and Consumption Expenditure: Evidence from New Zealand Microdata.' The paper investigates how household debt affects the marginal propensity to consume out of housing wealth. The paper uses New Zealand household-level data on spending, income, and debt over the period 2006–2016. The main empirical challenge is to identify exogenous variation in house prices to determine how consumption evolves with movements in household wealth. This identification problem is complicated by the presence of unobserved household characteristics that are correlated with housing wealth. The paper uses a detailed house sale dataset to derive local average house prices and use it as an instrument. The empirical results show that the estimated elasticity of consumption spending to housing wealth is about 0.22%. In dollar terms, the average marginal propensity to consume out of a one-dollar increase in housing wealth is around 2.2 cents. The empirical results confirm that household indebtedness, especially mortgage debt, acts as a drag on consumption spending, not only through the debt overhang channel, but also through influencing the collateral channel of the housing wealth effect.

ACKNOWLEDGEMENTS

Contents

LIST OF TABLES

LIST OF FIGURES

xi

CHAPTER 1:
Preliminaries

This paper-based thesis consists of three papers. The current chapter explains the motivation, the findings, and the structure of each paper.

1.1 Motivation, findings, and structure

The first paper is titled 'Reconciling International Trade Data.' The paper is motivated by the fact that international trade data are filled with discrepancies: where two countries report different values of trade with each other. 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. The impact of data discrepancy on empirical research is widely ignored among data users in the literature. This is reflected by the fact that most studies do not 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 informal approaches 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'. A few studies attempted to reconcile international trade data. Such studies have several issues including the use of biased measures of reporter quality, the subjective choices of acceptable quality thresholds, and not accounting for the role of data availability as an important dimension of reporter quality. Unreconciled data are much more commonly used than reconciled data.

The paper 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 2016. The index takes data availability and the relative significance of each partner into account. We reconcile international trade data

through picking the value reported by the country with higher data quality index in every bilateral flow. The findings include: (a) global trade was under-reported by roughly 5% over the past five years as countries with lower data quality under-report their trade; (b) erroneous reporting is prevalent among low-quality reporters; (c) importers' data are less likely to be in error; (d) the level of development and corruption are possible determinants of trade data quality; (e) China tends to under-report its exports and over-report its imports, while there is only a small difference between US self-reported and reconciled data.

We make the reconciled trade dataset freely available for future studies to use. The paper is organized as follows: Section 2.1 introduces the paper. Section 2.2 lists the sources of unreconciled international trade data. Section 2.3 summarises the cause of data discrepancy from the literature. Section 2.4 reviews the attempts made at reconciling international trade data and highlights the need for an alternative method. Section 2.5 discusses the raw data and the methodology. Section 2.6 presents the findings. Section 2.7 discusses the robustness of the proposed data quality index. Section 2.8 concludes.

The second paper is titled 'Why You Should Use High Frequency Data to Test the Impact of Exchange Rate on Trade.' The paper was published in the journal of Applied Economics Letters (Shaar and Khaled 2018).

The short paper provides theoretical and empirical evidence that investigating the exchange rate-trade nexus using low frequency data obscures the possibility of a true relation between the two variables. The use of low frequency data also suppresses the short-run relation and turns the variables of interest into I(2) series, which makes investigating the relation more problematic. The paper suggests that testing the impact of exchange rate on trade should be done using high-frequency data. Using different data frequencies for identical periods and specifications between the US and Canada, we show that low frequency data suppresses and distorts the evidence of the impact of exchange rate on trade in the short-run and

the long-run. The paper is organized as follows: Section 3.1 introduces the paper. Section 3.2 explains the research methodology, section 3.3 presents the findings, and section 3.4 concludes.

The third paper is titled 'Housing Leverage and Marginal Propensity to Consume.' The paper was published at the [Reserve Bank of New Zealand](#).

Understanding household consumption spending is crucial for modelling business cycles and designing macroeconomic policy. This paper investigates how household debt affects the marginal propensity to consume out of housing wealth. We use microdata from Statistics New Zealand's 'Household Economic Survey' (HES) to investigate how household leverage affects the marginal propensity to consume (MPC) out of housing wealth. HES data provide detailed information on household spending, income and loans.

Empirically, estimating the effect of housing wealth changes on household expenditure faces two types of endogeneity issues. First, any evidence of an association between housing wealth variations and consumption changes could be driven by unobservable confounding factors such as future income expectations or household preferences. Second, naive regressions with total household spending can suffer from reverse causality, in which high housing-related spending leads to higher property values. We combine HES data with Real Estate Institute of New Zealand (REINZ) micro house price data to address the endogeneity issues that arise from using household-level cross-sectional data.

In the empirical analysis, we first assess the validity of average local house prices as an instrument for individual house prices. The first stage regression suggests that the instrument can explain up to 22% of the variation in individual house prices reported in HES. We then run a benchmark regression of total household expenditure excluding housing-related spending on housing wealth. The IV estimation suggests that using household-level prices leads to downward bias,

which is the result of various causes of endogeneity issues discussed above. The average MPC out of a one-dollar increase in exogenous housing wealth is around 2.2 cents. All regressions control for income, household characteristics, and regional and time fixed effects. We also split non-housing expenditure into durables and non-durables. In line with other studies in the literature, we find that durable consumption is more sensitive to changes in housing wealth than non-durables.

We then focus on the role of household leverage in determining the MPC out of housing wealth. In this analysis, we study how leverage measures, such as the loan-to-house-value ratio (LVR) and the DTI, affect the estimated MPC out of housing wealth. Overall, we find that household leverage weakens the MPC associated with housing. To examine the robustness of these findings, we investigate whether household spending responds differently depending on the age and type of home ownership.

The findings confirm that the consumption of mortgagors is less sensitive to housing wealth as compared to outright homeowners. The regression with an age-housing wealth interaction also shows that the response of younger households to changes in their housing wealth is weaker than the response of older households, which tend to be less leveraged. The present paper is organised as follows: Section 4.1 introduces the paper. Section 4.2 discusses the HES data. This is followed by a discussion of the empirical approach in Section4.3. Section 4.4 presents our empirical results and robustness checks. We conclude in Section 4.5.

CHAPTER 2:

Reconciling International Trade Data

2.1 Introduction

When trade flows from country A to country B, A reports its value as exports and B reports it as imports. Therefore, every bilateral trade flow is reported twice. Trade data discrepancy occurs when two countries report different values of trade with each other. To illustrate the scale of discrepancies in international trade data, Figure 2.1 plots global trade through aggregating bilateral trade flows once using the higher reported value and once using the lower reported value. Which of the two possibilities represents actual global trade?

FIGURE 2.1: GLOBAL TRADE BASED ON HIGHER VS LOWER REPORTS IN EVERY BILATERAL RELATION ($US TRILLIONS)

Source: The Standard International Trade Classification (SITC)
Note: all trade data are on Free-on-Board basis (see section 2.3)
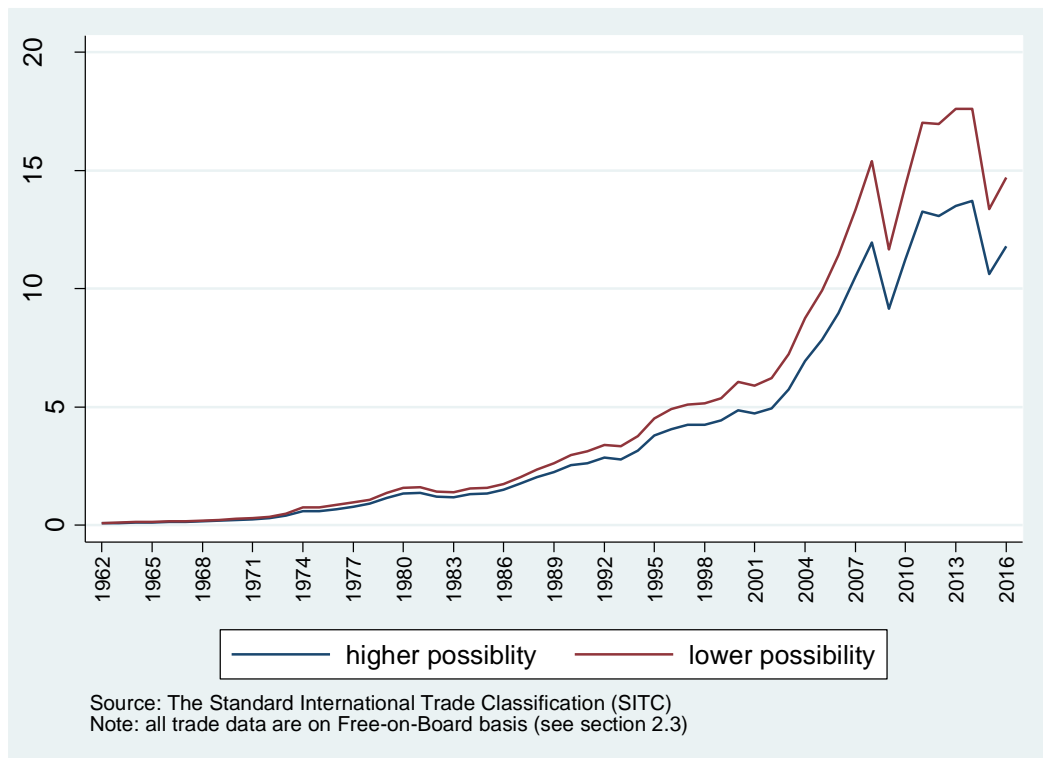
Figure 2.1 shows that global trade in 2016 can be anywhere between $11.8 trillion and $14.7 trillion depending on the claims about bilateral trade that you think are more reliable. Figure 2.2 plots the lower 'possibility' relative to the higher 'possibility': the ratio of possibilities was in decline most of the time since the late 1980s and stood at 0.8 in 2016 (Figure 2.2).

FIGURE 2.2 : LOWER POSSIBILITY OF GLOBAL TRADE DIVIDED BY THE HIGHER POSSIBILITY
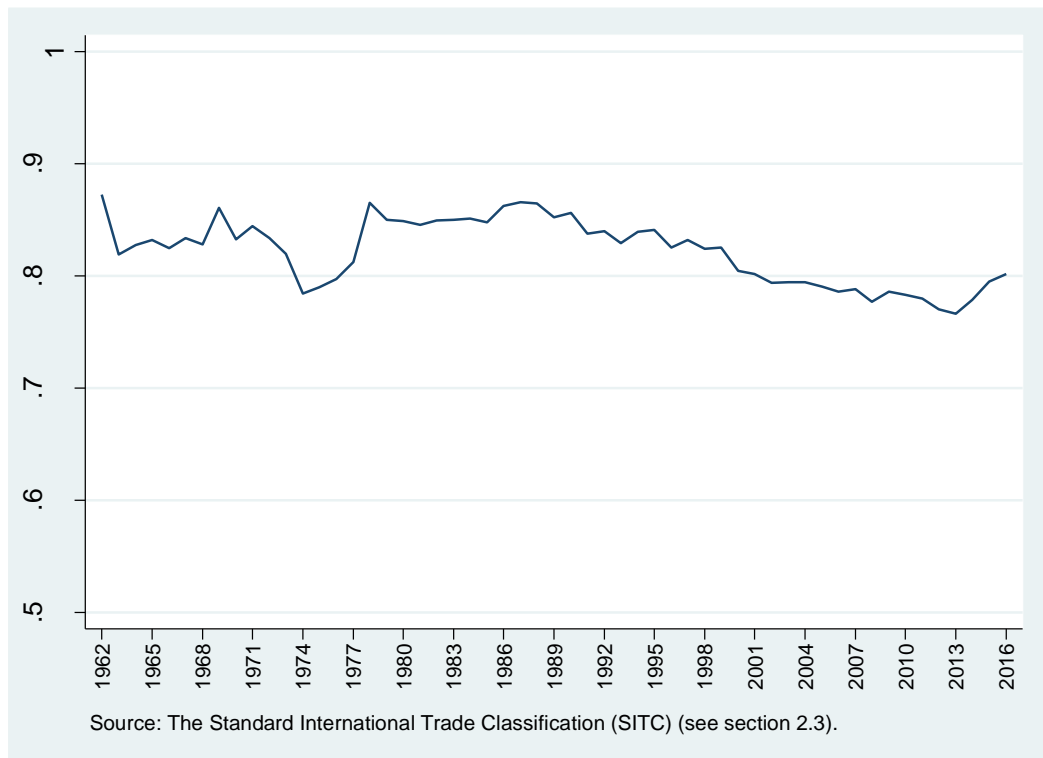
Source: The Standard International Trade Classification (SITC) (see section 2.3).

Substantial discrepancies can mean a study that uses data reported by a certain partner would reach a different conclusion than if it used the data reported by the other. For example, Bahmani-Oskooee et al. (2013) report 'drastically different results for the impact of the exchange rate on trade between the US and South Korea

depending on which of the two countries reported the data.'[1] 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.'

Four organizations attempted to reconcile trade data for a large number of countries and over an extended period of time. We review the methodologies followed by the four institutions in section 2.4 and highlight the need for an improved method. The main issues with the literature are the use of biased measures of reporter quality, the subjective choices of acceptable quality thresholds, and not accounting for the role of data availability as an important dimension of reporter quality.

We construct a Trade Data Quality Index (TDQI) based on the following notion: the more a country's reports on bilateral trade differ from the corresponding reports of its partners, the more likely it is a low-quality reporter. TDQI accounts for insurance and freight costs, the relative significance of each trade flow, and the cases where a country does not report data to begin with. The index ranges from 0 to 100 and it is calculated for each country's imports and exports separately in every year from 1962 to 2016. TDQI investigates the quality of more than 98% of global trade, which comprises around 1.3 million observations.

When two countries report different values for the same trade flow, one of the two values must be closer to the truth. TDQI is used for reconciling trade data. In any trade flow, we pick the value reported by the partner less likely to be in error: the partner with higher TDQI. We take the costs of insurance and transportation into

---

[1] Bahmani-Oskooee et al. (2013) examined the impact of exchange rates on US-Korea trade in 96 different industries and showed that for the same period of time and using the same econometric specifications, 'cointegration analysis produces drastically different results depending on the choice of the reporting country. Nonparametric analysis shows that the Korean imports are more sensitive to real exchange rate fluctuations than US exports...'.

consideration depending on whether we use the value reported by the exporter or the importer as the more reliable value for reconciliation. The reconciliation does not invent new numbers: the reconciled data come from whichever partner the rest of the world deem more reliable.

We find that over the last five years, global trade using self-reported exports was 6.3% lower than using reconciled data; while imports were 3.3% lower. On average, global trade is 4.8% higher using reconciled data. The differences are because low-quality reporters under-report their trade (imports and exports) although imports data tend to be more reliable. Erroneous reporting is also prevalent among low-quality reporters. Reconciled and self-reported data for high-quality reporters are similar for two reasons. First, in trade with a low-quality reporter, the value is picked from the country with higher quality: self-reported data are the source for reconciliation. Second, in trade between high-quality reporters, the magnitude of the discrepancies is small.

We show that several factors can explain trade data quality, such as the size of the economy, the level of development, and corruption in the public sector. Although global trade has never been more complex, import and export data quality improved in most countries over the period of the study. For global aggregate trade, however, the quality has been deteriorating over most of the past two decades because of the increasing importance of low-quality reporters.[2] Where China tends to under-report its exports and over-report its imports, we observe trivial differences between reconciled and self-reported data for the US.

TDQI materially improves international trade data quality. We recommend future studies on international trade use the data reconciled in this study as the differences

---

[2] Importance is measured in terms of the trade share.

between the reconciled and unreconciled data have important implications for our understanding of international trade. For illustration, we calculate trade openness for every country in 2016, once using the data reconciled here and once using self-reported data.[3] Low-quality reporters are, on average, 14% more open to trade using reconciled data. Countries with medium and high-quality data have roughly the same trade openness using reconciled and unreconciled data.[4] The literature largely depends on self-reported trade data.[5]

The following link provides the results of TDQI, the reconciled and unreconciled international trade data on a bilateral level and for a country's trade with the world combined. It also provides the Stata programming codes used in the calculations. https://goo.gl/cGVXDk

2.2 Sources of unreconciled international trade data

Most countries report their commodity trade to international institutions. The institutions, in their turn, aim to achieve unanimous scales and definitions for international trade through adopting commodity classifications and promoting standard reporting practices. Two institutions maintain the most popular commodity classifications. First, the UN Statistics Division, which maintains the Standard International Trade Classification (SITC).[6] The SITC is based on the economic functions of commodities at different stages of development. It is the oldest international commodity classification: the first version was released in 1950;

---

[3] Trade openness is defined as (exports+imports)/GDP.
[4] We break the reporters into three groups (high, medium, and low-quality) based on the average quality of their imports and exports data.
[5] To confirm this, we reviewed the top 30 papers in terms of citations on Scopus search engine, where the papers were published since 2015 and contained 'international trade' as a keyword. 29% of the papers discuss the issue of data discrepancy or use a data source that reconciles trade data discrepancies.
[6] The SITC is a part of the Commodity Trade Statistics Database (COMTRADE).

the earliest online version is available from 1962. Second, the World Customs Organization maintains the Harmonized System (HS). The HS is a multipurpose classification that was implemented in 1988. There are correspondence tables between the two classifications (de Saint Vaulry, 2008).

The HS and the SITC were revised several times to account for the changes in the nature of the traded goods. Both classifications publish bilateral data on commodity as well as aggregate levels. Self-reported commodity trade does not necessarily add up to self-reported aggregate trade.[7] There are marginal differences between bilateral aggregate data published by the HS and the SITC, even across different revisions: the differences are on the commodity level, while this study is concerned with aggregate trade.

The IMF publishes bilateral aggregate trade in its Direction of Trade (DOT) Database. We compare DOT with SITC/HS data for the period 1962–2016 and find the data sets identical in most cases. DOT, occasionally, replaces missing values with partners' corresponding values; it is not clear under what circumstances the replacements are made. DOT uses a 10% share of the reported value to account for insurance and transportation when converting between the data reported by the exporter and the importer to replace missing values.[8] DOT also estimates annual trade values based on available monthly reports. The WTO does not report bilateral trade and the number of reporting countries is smaller than in the SITC and the HS. The World Bank retrieves its data from the WTO, the IMF, and COMTRADE.

---

[7] Due to confidentiality, countries may not report some of their detailed trade.
[8] See section 2.3 for more details on export and import valuation methods.

2.3 Causes of trade data discrepancy

Several studies investigated why some countries might report different values of trade with each other. Yeats (1990) compared sub-Saharan trade with the rest of the world using SITC data and detected large discrepancies, stating that mis-invoicing and smuggling are 'apparently responsible.' Confirming the logic of the Trade Data Quality Index proposed in this study, Yeats (1990) shows that although major disparities exist in data on trade with developed countries, the average differences in intra-African trade statistics are substantially larger. Finally, the paper suggests that sub-Saharan African data should not be relied on to indicate 'the level, composition, or even direction and trends in African trade.'

West (1995) focused on US-China trade and claimed that most of the discrepancies in their data are due to two factors. First, re-exportation through Hong Kong. Second, the price mark-ups added by merchants in Hong-Kong after the products had left China. Several studies later on focused on US-China trade.

Makhoul and Otterstrom (1998) examined the discrepancies in the Direction of Trade database, managed by the IMF, and found that taking insurance and freight into account when comparing the data of importers and exporters can only explain a small portion of trade data discrepancies. Something, we echo here in this study.

Vincent (2004) used econometric analysis of bilateral trade statistics for Romania and other European countries and found that measurement error, shipment lags, and intentional under-reporting all play a role in explaining discrepancies for two different types of sawn wood.

Guo (2010) analysed the asymmetries in trade between China and its top five trading partners during the period 1992–2008. The paper lists several reasons for the discrepancies such as the different pricing systems between imports and

exports, the different trade systems among countries, re-exporting, and re-importing. At odds with what we aim to show, Guo (2010) concluded that it is difficult to find a systematic way to correct asymmetries in international trade statistics among countries.

Wang, Gehlhar, and Yao (2010) proposed a mathematical programming model to reconcile trade data discrepancies in the presence of a third re-exporting country. However, the proposed method requires initialization with detailed estimates of bilateral trade flows and data reliability indices, which are not observable from the data.

Ferrantino et al. (2012) investigated the direct trade between the US and China and found 'strong statistical evidence' of under-reporting exports at the Chinese border to avoid paying value-added tax (VAT). We also confirm this finding. The under-reporting of Chinese exports to the US accounted for approximately two-thirds of the discrepancy over the period 2002-2008.

Simola (2012) assessed trade data quality in Russia and showed that while some discrepancies can be explained by the misclassification of commodities, some discrepancies apparently reflect deliberate misreporting.

Several studies in the literature also showed that trade data discrepancy can be explained by favourable bilateral trade agreements, export schemes, tariffs, and corruption. For example, Fisman and Wei (2009) showed that bilateral trade discrepancies are highly correlated with the corruption level of the exporting country as measured by commonly used survey-based indices, and that this correlation is stronger for artefact-rich countries.

Based on the studies reviewed above, we break the cause of trade data discrepancies into four categories:

TABLE 2.1: POTENTIAL CAUSES OF TRADE DATA DISCREPANCY

Type	Cause	Explanation
lack of uniformity in data compilation methodologies between the partners	what constitutes imports and exports?	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 trading 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 imports and exports	imports are usually reported on cost-insurance-freight (CIF) basis; exports are usually reported as free-on-board (FOB)
	timing effect	shipments are registered at different points of time by both countries since they are registered as they happen (shipment lags)
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
erroneous reporting		failure to value the goods correctly or to attribute the flow to the right partner due to negligence

Source: Author's summary of the literature.

2.4 Previous efforts in reconciling trade data

The literature documents multiple attempts investigating the causes of discrepancies among a handful of countries (see section 2.3). The US-Hong Kong-China trilateral trade received most of the attention. Other countries, such as Mexico, the US, and Canada, set up commissions to reconcile the discrepancies with each other. The review in this section is concerned with attempts at reconciling trade data internationally: reconciliation for a large number of countries and over an extended period of time. Note that all such attempts reconciled the data on the commodity-level; our attempt reconciles aggregate trade only.[9]

In summary, the main problems with the literature are: the use of biased measures of reporter quality, the subjective choice of acceptable quality thresholds, and not accounting for the role of data availability as a dimension of reporter quality.

2.4.1 Global Trade Analysis Project (GTAP)

Gehlhar (1996) reconciles trade data for GTAP starting from the SITC commodity-level and later aggregates the reconciled bilateral flows into different regions. GTAP data are proprietary.

The logic of reconciliation is similar to the one adopted in this study. First, use the overall similarity between a country's reports and the corresponding reports of all of its partners as a measure of reporter quality. Second, reconcile every bilateral flow through picking the value from the reporter with higher quality.

Gehlhar (1996) uses the following measure of similarity between two corresponding values: |export-import|/import. The problem with the measure is

---

[9] Our method can be later extended to investigate the quality on the commodity-level.

the asymmetry to whether the importer or the exporter misreport bilateral trade: the measure yields different results when the difference in reports is the same. The asymmetry means that the measure is biased; the direction of the bias depends on whether the importer or the exporter over or under-reported trade. Section 2.6 shows that low-quality reporters tend to systematically under-report their trade, which highlights the importance of a symmetric measure of similarity.

Gehlhar (1996) uses a threshold of 0.2 for the measure of similarity to classify every bilateral flow into 'matched' and 'not matched'. Reporter quality is then defined as the sum of the total value of matched flows as a share of total self-reported trade.[10]

We use a symmetric measure of similarity and follow a ranking method to avoid setting a subjective threshold. Table 2.2 provides an illustrative example.

TABLE 2.2: ILLUSTRATIVE EXAMPLE OF DIFFERENT METHODOLOGIES OF MEASURING THE SIMILARITY BETWEEN TWO CORRESPONDING VALUES

	New Zealand-reported imports from the US (m)	US-reported exports to New Zealand (x)	Gehlhar (1996) similarity measure $|x-m|/m$	proposed similarity measure $|x-m|/(x+m)$
scenario 1	33	40	0.21 (not matched)	0.10
scenario 2	40	33	0.18 (matched)	0.10

As in the case of other studies in the literature, Gehlhar (1996) does not account for the role of data availability in measuring quality. This paper argues that reporter quality should be a function of available and unavailable data. A country should not be considered a high-quality reporter because it reported a few values consistent

---

[10] What we call 'reporter quality' is called 'reporter accuracy' in Gehlhar (1996). As his study is on commodity-level, the measured accuracy varies by commodity, country, and flow direction (export vs import).

with partners' reports together with many no-trade values that differ substantially from partner's reports.

2.4.2 The National Bureau of Economic Research-the United Nations (NBER-UN)

Feenstra et al. (2005) reconcile the data for the NBER-UN database. Country-specific knowledge about reporting practices over time is used to reconcile the discrepancies in the SITC data on a case-by-case basis. Such knowledge relates to the treatment of re-exports, the geographic boundaries of the partner country, and so on.

However, Feenstra et al. (2005) give primacy to the importers' reports, whenever available, assuming they are more accurate. If the importer's report is not available for a country-pair, the corresponding exporter report is used instead. We argue that primacy should not be given to either side. Although TDQI confirms that importers' reports are more accurate on average, the index also highlights the numerous cases where the opposite is true. The reconciled value should be picked from whichever partner has a higher data quality. As section 2.6 shows, throughout the period of the study, exporters' data are used 43% of the time for reconciling bilateral discrepancies, while importers' data are used for the remaining 57% of the cases.

The case-by-case nature of the reconciliation means it takes a long time to reconcile the data. Timely reconciliation is important, especially for policymakers. NBER-UN data are currently available until the year 2000.

2.4.3 Centre d'Etudes Prospectives et d'Informations Internationales (CEPII)

Two studies from CEPII attempted to reconcile commodity trade. de Saint Vaulry (2008) reconciles the data for the period 1967–2005 using the STIC data. The study breaks each country into four quadrants based on the timeliness as well as the

quality of reporting. The reconciliation depends on the quadrants of the partners and it is achieved using the weighted-average of reports. However, the weights are determined by the author. For example, if a country in quadrant x traded with another in quadrant y, a weighted average of reports of one-third to two-thirds is used to reconcile the discrepancy. The results are likely to be sensitive to the choice of averaging weights.

Gaulier and Zignago (2010) also reconcile the data for CEPII in what is called BACI database. Reporter quality is estimated using a regression that makes strong assumptions about the distribution of the discrepancies. For instance, the discrepancies of reported data from the truth are assumed to follow a multiplicative log-normal form with a mean of zero. We argue that the difference between reported and actual trade is not likely to be centred around zero (actual trade is likely to be greater due to under-invoicing, and smuggling [see Figures 2.10 and 2.11]).

Most importantly, BACI, like du Saint Vaulry (2008), uses the quality-weighted average of reports to reconcile every flow. The problem with this approach is that the reconciled value in a certain trade flow will end up worse than the report of one of the two partners. When two countries report different values for the same trade flow, one of the two values must be closer to the truth. As done by Gehlhar (1996), it is safer to pick the report from whichever partner with higher quality to ensure the reconciled data are an improvement over self-reported data. As the literature shows and TDQI confirms, the values reported by some low-quality reporters mean next to nothing; such reporters should not be given any weight in the reconciliation process. Finally, for the sake of credibility, it is important not to be seen as making up new numbers for international trade.

2.4.4 Organization for Economic Co-operation and Development (OECD)

Fortanier (2016) reconciles the data for the OECD using the HS from 2007 to 2016. Reporter quality is defined as the 'share of bilateral trade for which the absolute difference with the corresponding trade data is less than or equal to 10% of the sum of these two value flows.'[11] Therefore, Fortanier (2016) uses the same similarly measure followed here. However, like Gehlhar (1996), Fortainer (2016) applies an arbitrary threshold for classifying every two corresponding reports into 'matched' and 'not matched'. As argued earlier, the findings are possibly sensitive to the choice of the threshold.

Most importantly, as in the case of Gaulier and Zignago (2010), the study uses the quality-weighted average of reports in reconciling every bilateral flow. As discussed in section 2.4.3, we argue that this method might end up worsening the quality of trade data than improving it. Finally, Fortainer (2016), like other studies in the literature does not account for the role of data availability in the measurement of reporter quality.

2.5 Data and Methodology

In order to assess data quality, we need to rely on raw data where the publishing international institution made no changes to the data reported by countries. The dataset also needs to be comprehensive. The HS and the SITC meet these two criteria (refer to section 2.2 for more details). Since there are only small differences between bilateral aggregate trade data published by the HS and the SITC, we use the data from the SITC, revision 1, to maximize the period of the study.

---

[11] What we call 'data quality index' is called 'asymmetry measure' in Fortanier (2016).

The annual data extend from 1962 to 2016 and cover all available bilateral aggregate trade; this comes to 1,632,245 observations. The data are in US dollars and exclude the years 2017–2018 since some countries have not reported their trade yet while some other reports are still subject to revisions by the reporting countries.[12] We downloaded the data from the World Integrated Trade Solutions (WITS) website, administered by the World Bank. Several other online portals offer free access to SITC records.

The UN Statistics Division regularly publishes its recommendations to individual countries on preferred reporting practices to enhance comparability. The 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 Statistics Division ('International Merchandise Trade Statistics- Concepts and Definitions ', 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 re-exports and re-imports take place, total and gross bilateral trade data would be identical. We use 'total' rather than 'gross' bilateral data in constructing TDQI because to focus on actual trade from domestic origin; this is in line with de Saint Vaulry (2008). 'Total' trade also enhances the comparability between corresponding trade values.

---

[12] Countries can update their reported trade as more information become available.

For all countries over the period 1962–2016, reported re-imports as a percentage of gross imports is less than 0.5%. To ensure comparability among corresponding trade flows, we assume all countries are equally engaged in re-importing. This allows TDQI to remain comparatively correct, which is what the index aims to achieve as TDQI is used for ranking purposes.

Global re-exports as a percentage of gross exports over 1962–2016 is almost 2.3%. In 2016, as much as 85% of reported global re-exports went through Hong Kong and the US, with 57.2% and 27.9% respectively. According to the UN Statistics Division recommendations, a country should attribute its exports to the final known destination and its imports to the country of origin. 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 total exports but included in total imports when they enter a country.

Not all countries, however, are capable of identifying the country of origin or the country of final destination. As discussed below, 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 increase (Figure 2.3).

FIGURE 2.3: THE SHARE OF RE-EXPORTS IN GROSS EXPORTS GLOBALLY
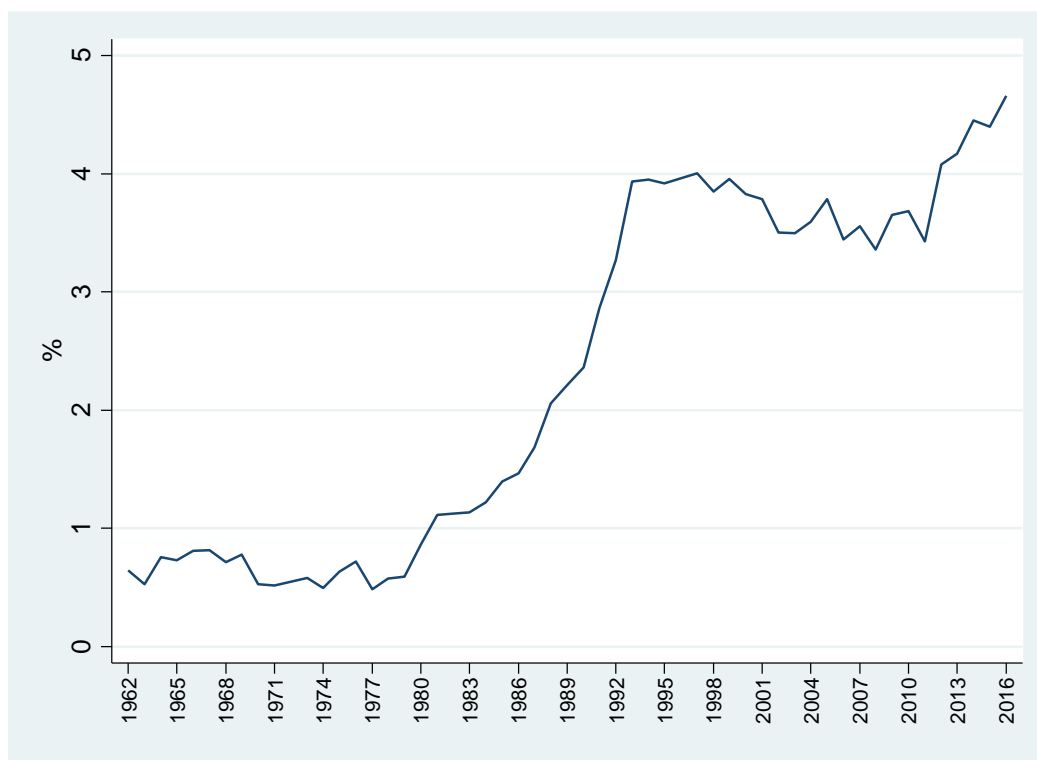
As stated in Table 2.1, the third reason for trade data discrepancy is the geographical definition of a trading partner. To this end, we drop any entity or country, which is not considered by its partners as a partner. For example, several large countries do not report any trade with tiny pacific islands, such as Palau and Nauru. To ensure comparability, we drop any observation relating to exporting or importing from them. Appendix B lists the dropped entities/countries. The share of exports reported by the dropped entities in global self-reported exports along the whole period of the study is 1.9%. This leaves us with 1,299,665 bilateral trade observations.

The causes of discrepancy related to timing and conversion to a 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. The UN Statistics Division has clear recommendations on conversion to foreign currency.

The other issue to be addressed is that most countries, as recommended by the UN Statistics Division, report their exports on FOB and their imports on CIF. This recommendation deters us from directly comparing corresponding bilateral trade claims since the importer includes the costs of insurance and transportation in the value of imports and the exporter does not include them in the value of exports. To enhance comparability, we convert all imports from CIF to FOB before constructing TDQI and reconciling trade data. Therefore, all the figures of trade reported and reconciled in this paper are valued on FOB basis. We use Miao and Fortanier (2017) bilateral relation and time-specific CIF-FOB conversion factors.[13] Miao and Fortanier (2017) use a combination of actual shipping and insurance data and gravity models to estimate the conversion factors.

The causes of discrepancy related to corruption and erroneous reporting are discussed in section 2.6.

For a specific year and trade flow, let N denote the number of trading partners involved in the calculation of TDQI for country i in a given year. For partners $j = 1, ..., N$:

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

---

[13] As Miao and Fortanier (2017) bilateral CIF-FOB conversion factors are not available before 1995, we use the average of those conversion factors over the period 1995-2016 to replace the missing values before 1995 for every bilateral relation. This approach is taken as the weighted average of global CIF-FOB conversion factors does not have any discernible time trend. In other words, if CIF-FOB factors between, say, the US and Japan average to 4% over the period 1995-2016, we assume they are also 4.5% for the period 1962-1994.

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 two corresponding trade values for country i's exports is:

$$S_{ij}^X = 1 - \frac{|X_{ij} - M_{ji}|}{X_{ij} + M_{ji}} \qquad [2.1]$$

The similarity between two corresponding trade values for country i's imports:

$$S_{ij}^M = 1 - \frac{|M_{ij} - X_{ji}|}{M_{ij} + X_{ji}} \qquad [2.2]$$

For all years in our data, we have 453,067 bilateral trade flows for which there are values reported by both partners (906,134 out of 1,299,665 observations) to calculate the similarity measures. [14] The remaining bilateral 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. The higher the measure, the higher the similarity. Unlike the method applied by the GTAP, the similarity measure is symmetric (refer to section 2.4.1 for more discussion).

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

---

[14] 453,067 X 2=906,134

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

$$Sh_{ij}^X = \frac{X_{ij}}{X_i} \qquad [2.3]$$

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

$$Sh_{ij}^M = \frac{M_{ij}}{M_i} \qquad [2.4]$$

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:

$$\bar{S}_i^M = \sum_{j=1}^N Sh_{ij}^M * S_{ij}^M \quad [2.5]$$

Average similarities for exports of country i:

$$\bar{S}_i^X = \sum_{j=1}^N Sh_{ij}^X * S_{ij}^X \quad [2.6]$$

Average similarities for imports and average similarities for exports for each country and year range from greater than 0 to 1. The higher $\bar{S}_i$, the higher the average similarity between country i's available data and the data reported by its partners. The rationale behind the measure of average similarities as a proxy for data quality is straightforward: the more a country's reports on bilateral trade differ from the corresponding reports of its partners, the more likely that country is a low-quality reporter.

Each single similarity measure between two corresponding trade values is included in the calculation of two average similarities: the average similarity for the imports of one country and the average similarity for the exports of the other. Therefore, a country with higher data quality gets its average similarity lowered due to trading with a country with lower quality. As j grows to its maximum N in Equations 2.5 and 2.6, all countries are assumed to be equally likely to trade with different

combinations of partners with the same average quality. Section 2.7 argues that this assumption is reasonable.

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. This would cause the weighted average similarities measure to be incorrect due to the bias of the trade shares used in constructing it: a country might appear to have higher quality because it reported a few values that happened to be in line with partners' reports. Other studies on reconciling international trade data do not take data availability into account when measuring reporter quality (see section 2.4). To account for the impact of data availability, we first develop a Trade Data Availability Index (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:

$$TDAI_i^M = \left(1 - \left(\frac{m_i^M}{n_i^M}\right)\right) * 100 \qquad [2.7]$$

Exports TDAI for i:

$$TDAI_i^X = \left(1 - \left(\frac{m_i^X}{n_i^X}\right)\right) * 100 \qquad [2.8]$$

TDAI ranges from 0 to 100. The higher the TDAI, the more available the bilateral trade data.

SITC recommends that countries report any bilateral trade flow exceeding $1000 US dollars. However, not reporting bilateral trade with a certain partner or partners might be the right practice from country i because such trade did not actually happen. Assuming all countries are equally susceptible to this scenario, TDAI remains comparatively correct, i.e. biased downwards for all countries equally. We argue that this bias in the measures of data availability, if exists, is quite small as several countries have a TDAI of 100 (see Appendix A).

We use TDAI to penalize a country for not reporting data at all:

$$TDQI_i^X = \bar{S}_i^X * TDAI_i^X \qquad [2.9]$$

$$TDQI_i^M = \bar{S}_i^M * TDAI_i^M \qquad [2.10]$$

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: a country that does not report data at all, has a quality of zero. TDQI is a function of available and unavailable data. The higher the index, the higher the quality of a country's trade data.

TDQI is built to simply tell us which partner in any bilateral trade flow is more likely to be more accurate, which serves as the basis of data reconciliation. In any bilateral trade flow, we reconcile the discrepancy in claims through picking the data from the partner with higher TDQI. Following Gehlhar (1996), we do not use the average of claims weighted by quality as a basis of reconciliation. Section 2.4.3 brings several reasons for not using the quality-weighted average of reports as a reconciliation method.

The most serious issue with 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 the UN Statistics Division recommendations of partner attribution. Otherwise, it is not possible to determine who traded what. The comparatively high TDQI for imports and exports 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; perhaps because the US properly informs them as the three countries reconcile their trilateral data before reporting to the SITC.

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 the goods actually originate from China. This might happen due to the change in goods ownership once in Hong Kong and before re-exportation (Ferrantino et al., 2012). Since this information might be only available to Hong Kong, it is hard for its partners to identify the correct country of origin.

We believe Hong Kong's exports reports are more accurate than the reports of the importers from it, even when these importers have a higher TDQI for their imports. This position is supported by two facts. First, Hong Kong's TDQI for exports started declining only after it started engaging actively in in re-exportation. Second, Hong Kong has very low corruption, as indicated by the Corruption Perceptions Index from Transparency International. Therefore, mis-invoicing, smuggling, and intentional partner misattribution are not common.

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 the TDQI scores. 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 the reconciled data. TDQI cannot address the role of re-exports given the available information 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 re-exports to the SITC, such as the Netherlands and Singapore.

2.6 Results and discussions

More countries are reporting their trade over time

The number of countries reporting bilateral trade increased over time: 77 countries reported bilateral exports to at least one partner in 1962 compared with 134 in 2016. For imports, the number rose from 76 to 133 (Figure 2.4).

FIGURE 2.4 : NUMBER OF COUNTRIES REPORTING TRADE TO AT LEAST ONE PARTNER

Note: Some of these countries cease to exist; others became countries after 1962.
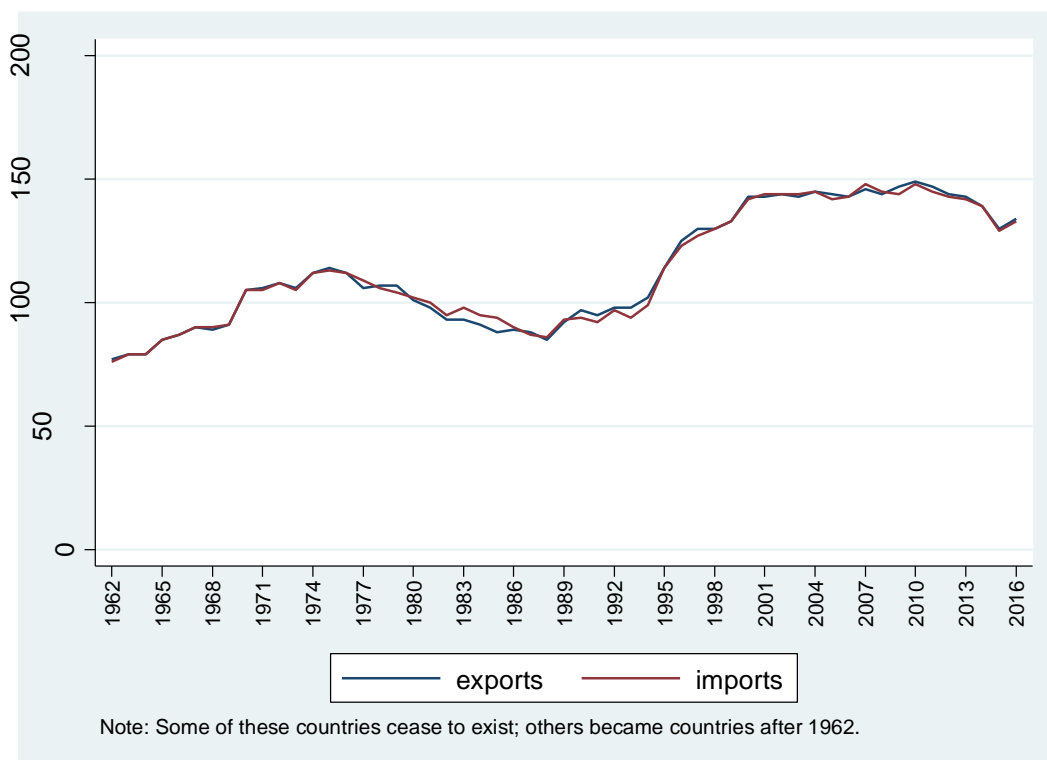
Trade data are more available for imports than for exports

Figure 2.5 calculates the unweighted average of the data availability index (TDAI) for all countries in a given year. TDAI has been systematically higher for imports than for exports. There are two possible factors why trade data are more available for imports: First, customs care more about determining the origin of imports than the destination of exports as import tariffs tend to be higher. Second, it is easier to identify the origin of imports than the final destination of exports at the time registration.

FIGURE 2.5 : ANNUAL AVERAGE OF TDAI FOR ALL COUNTRIES
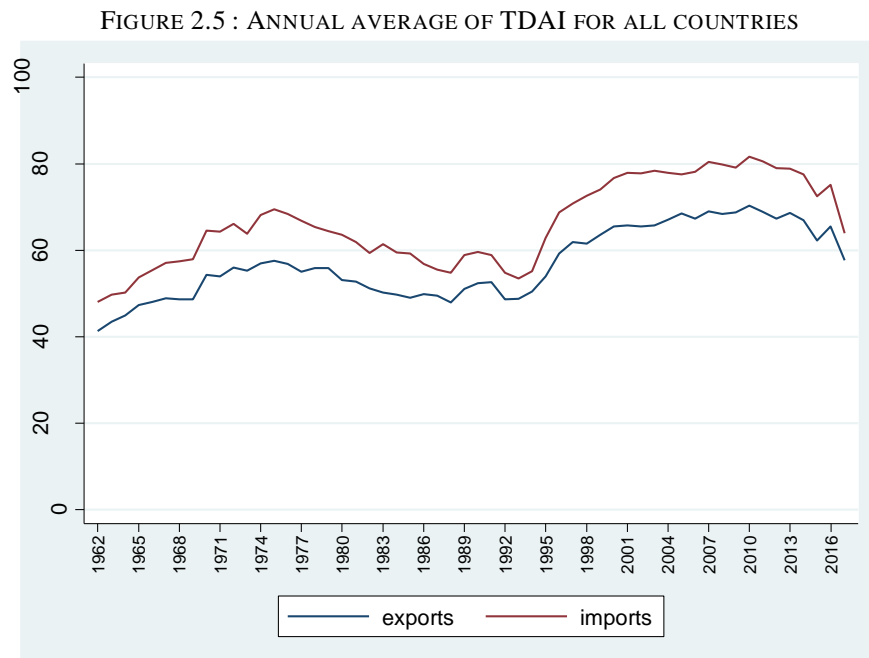
Importers' data are slightly more reliable

Figure 2.6 plots the distributions of the data quality index (TDQI) for imports and exports for all countries and years.

FIGURE 2.6 : DISTRIBUTION OF TDQI FOR ALL COUNTRIES AND YEARS

Note: the Figure excludes observations with zero TDQI to improve presentation.
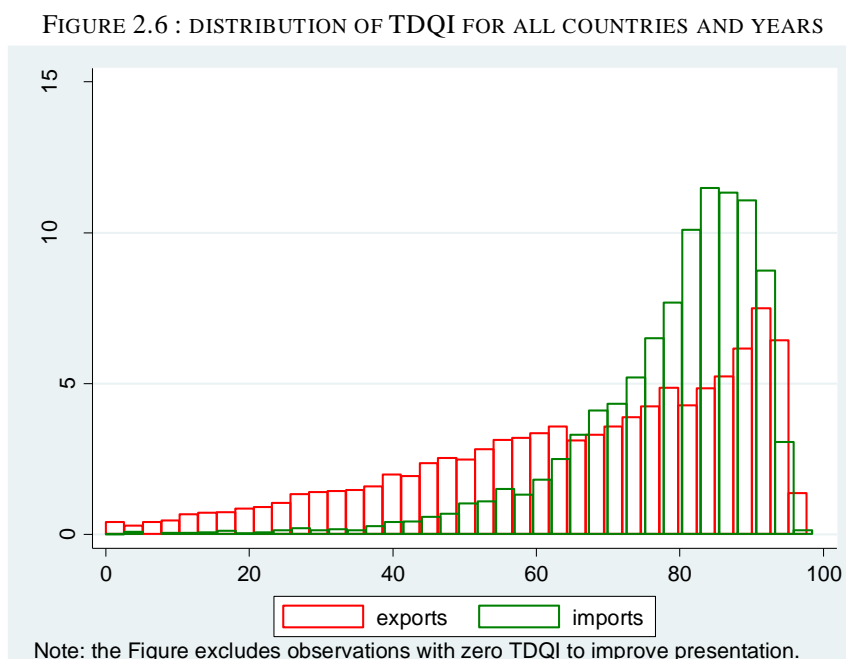
The average of TDQI is 66 for exports and 78 for imports. Average TDQI is higher for imports because both components of TDQI (the average similarities measure and the TDAI) tend to be higher for imports. As a result, it is more likely we pick the claim reported by the importer when reconciling a bilateral discrepancy. Several studies in the literature believe the data reported by importers are more accurate: Feenstra et. al (2005) rely solely on importers' claims in the reconciliation (refer to section 2.4.4). TDQI confirms that import data are relatively more accurate, however, the index also highlights the numerous cases where the opposite is the case. In fact, throughout the period of the study, exporters' data are used 43% of the time for reconciling bilateral discrepancies, while importers' data are used for reconciling the remaining 57% of the cases.

Countries with low data quality have increased their share in global trade

Figure 2.7 shows that average TDQI for imports and exports are roughly the same over time if we account for the trade share of each country in global trade. Global trade data quality using weighted average TDQI is largely in line with the results obtained using the method of lower-higher possibilities ratio explained in section 2.1 (Figure 2.2); the similarity serves as a robustness check for TDQI. We conclude that global trade data quality has been mostly deteriorating starting from the early 1990s. The decline in quality coincides with the increase in global trade and re-exportation.

FIGURE 2.7 : ANNUAL AVERAGE OF TDQI FOR ALL COUNTRIES WEIGHTED BY TRADE SHARES

Note: The trade shares used in calculating the weighted averages are using self-reported data.
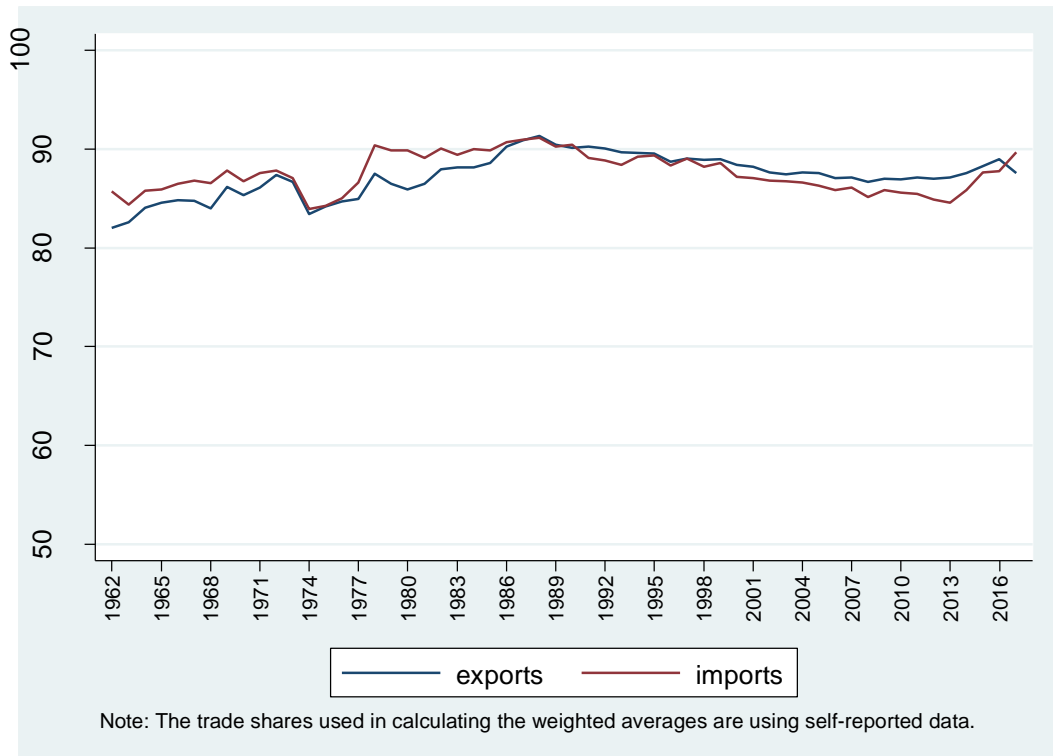
The deterioration in global trade data quality does not hold using the unweighted average of TDQI. We calculate the difference between TDQI in the last and the first years a country had trading partners reporting trade with it. For the case of exports, 115 countries had positive differences while 42 had negative differences; 100 compared with 56 for imports. Therefore, the quality of trade data improved in more

countries than worsened over time. Combining these numbers with the conclusion from Figure 2.7 means that the countries which have worsened in terms of TDQI have increased their share in global trade over time.

The quality and availability of data are strongly correlated

Appendix A lists TDQI and TDAI for imports and exports for the last year covered in this study (2016). The table is sorted by TDQI for exports from highest to lowest. Table 2.3 reports the correlations between all four indices. Correlations between TDAI and TDQI for imports and exports are all above 80%. The elevated correlations between TDQI and TDAI are expected as countries with higher data quality are more likely to report their trade. Similarly, high correlations between import and export measures of data quality and availability indicate that countries which report reliable export data are likely to do the same for their imports, and vice versa.

TABLE 2.3: CORRELATIONS BETWEEN TDQI AND TDAI FOR THE YEAR 2016

	TDAI exports	TDAI imports	TDQI exports	TDQI imports
TDAI exports	100%			
TDAI imports	94%	100%		
TDQI exports	94%	83%	100%	
TDQI imports	96%	97%	89%	100%

All correlations are statistically significant at 1% (t-test).

Countries on top of the list in Appendix A are mostly developed and transparent. On the other hand, countries with low data quality are, in general, those where corruption is endemic and border controls are largely non-existent. The findings echo those in Yeats (1990) paper: 'On the accuracy of economic observations: Do sub-Saharan trade statistics mean anything?' The answer to Yeats's question is: no. We add that this is not peculiar to Sub-Saharan countries.

Several factors can explain trade data quality

As discussed in the literature (section 2.3), several factors might be able to explain why the data reported by certain countries are more reliable than the data reported by others. To investigate this, we regress domestic TDQI on three different possible explanatory variables. First, the size of the economy, proxied by real GDP. Second, by controlling for population, we effectively examine the impact of the level of development as proxied in GDP per capita. Third, public corruption proxied in the Corruption Perceptions Index released annually by Transparency International. The higher the Transparency, the less corrupt the country. As stated in Table 2.1, in the context of trade data discrepancy, corruption can refer to smuggling, partner misattribution, and mis-invoicing.

TABLE 2.4: POSSIBLE DETERMINANTS OF TRADE DATA QUALITY

Specification	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: TDQI for exports			Dependent variable: TDQI for imports		
Independent variables						
Ln(GDP)	0.10***	0.11***	0.06***	0.07***	0.08***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ln(population)		-0.06***	0.03***		-0.06***	0.02**
		(0.01)	(0.01)		(0.01)	(0.01)
Ln(transparency)			0.21***			0.19***
			(0.01)			(0.02)
Constant	-1.79***	-2.01***	-1.76***	-0.97***	-1.20***	-0.98***
	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Observations	2,078	2,078	2,078	2,078	2,078	2,078
R-squared	0.41	0.42	0.48	0.24	0.26	0.31

The regressions control for country-specific and year effects.
Period: 2000–2015.
Standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Specifications 1 and 4 suggest that larger economies tend to have higher data quality. Adding the regressor of population in regressions 2 and 5 suggests that after controlling for the size of the economy, countries with smaller population, tend to have higher data quality. Put differently, countries with higher GDP per capita tend to have higher trade data quality.

Specifications 3 and 6 confirm several earlier studies in the literature, such as Fisman and Wei (2009), showing that lower corruption, i.e. higher public transparency, is associated with higher data quality. As shall be seen later, the fact that low-quality reporters tend to under-report their imports and exports, suggesting smuggling or tariff evasion, supports a causal relation between corruption and quality. However, more proper identification might be needed to confirm this relationship. Countries in the lowest quartile of average export and import data quality have an average of corruption perceptions index of 25.5 while countries in the highest quartile of data quality have an average of corruption perceptions index of 57.5.

We also note that countries which are more active in trade tend to have higher data quality, as Figure 2.8 shows. Regression analysis show that trading with one more partner is associated 0.7 rise in TDQI for the case of exports.[15]

---

[15] The regressions control for year effects and extend for the period 1962-2016. Similar results apply for imports.

FIGURE 2.8 : TDQI FOR IMPORTS AND EXPORTS VS NUMBER OF TRADING PARTNERS

Global trade is under-reported

Following the methodology explained in section 2.5, we use TDQI to reconcile international trade data. Figure 2.9 plots global trade using self-reported and reconciled values.

FIGURE 2.9 : GLOBAL TRADE USING SELF-REPORTED AND RECONCILED DATA
($US TRILLIONS)

Note:
Self-reported exports and imports are summed up to calculate global trade whether the partner reported trade or not. All data are on FOB basis.
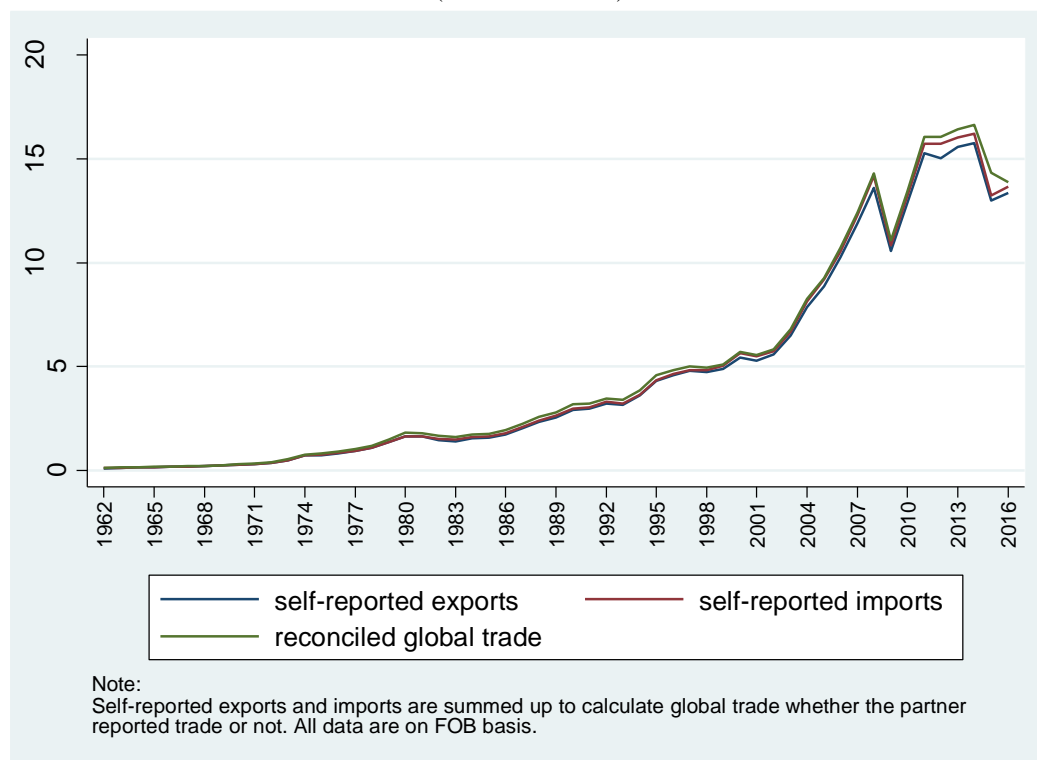
Almost along the whole period, global trade using reconciled data is higher than that using self-reported data: global aggregate trade is under-reported. Over the last five years combined, global trade using self-reported exports was 6.3% lower than that using reconciled data; the difference for imports was 3.3% (4.8% difference on average between self-reported and reconciled data). As shown next, the differences are because low-quality reporters under-report their trade.

Low-quality reporters under-report their imports and exports

To capture how reconciled and self-reported data differ on country-level, we calculate the ratio of the two for every country for the period 1962-2016. The ratio is calculated for a country's trade with the rest of the world combined as follows:

Reconciled-unreconciled ratio = (reconciled-unreconciled)/(reconciled + unreconciled)*100 [2.11]

The ratio is symmetric to differences in reconciled and unreconciled values and is confined between 0 and 100. Figures 2.10 and 2.11 report the distribution of the ratios for reporters with high, medium, and low quality (tertiles of countries by data quality).

FIGURE 2.10 : THE DISTRIBUTION OF RECONCILED-UNRECONCILED RATIO FOR EXPORTERS
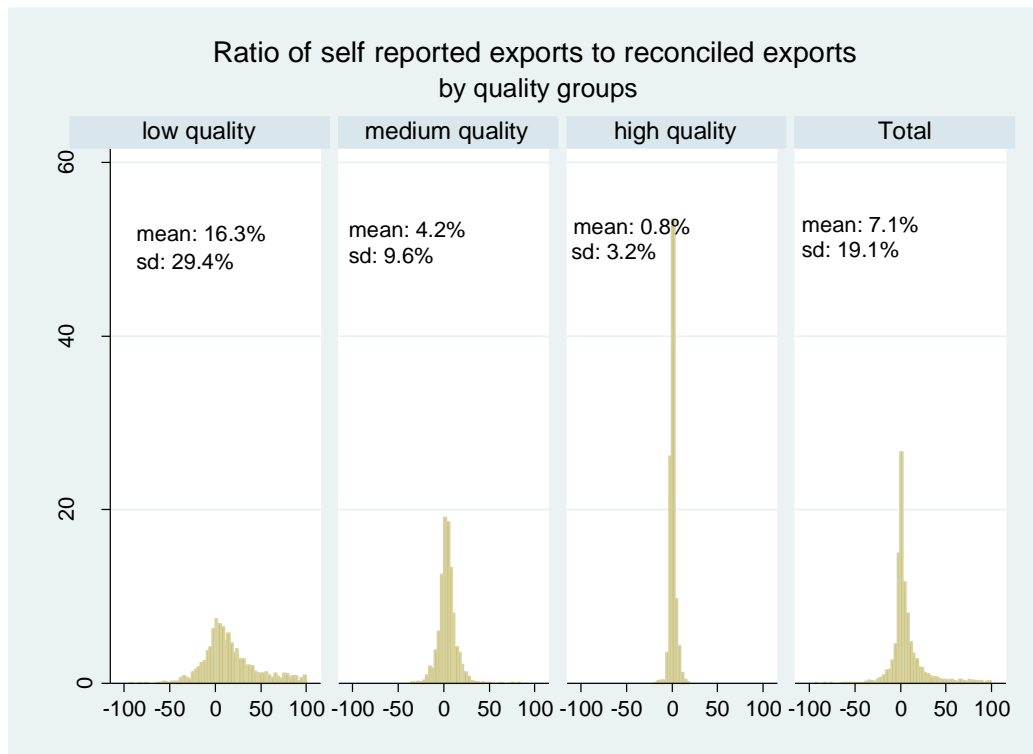
FIGURE 2.11 : THE DISTRIBUTION OF RECONCILED-UNRECONCILED RATIO FOR IMPORTERS
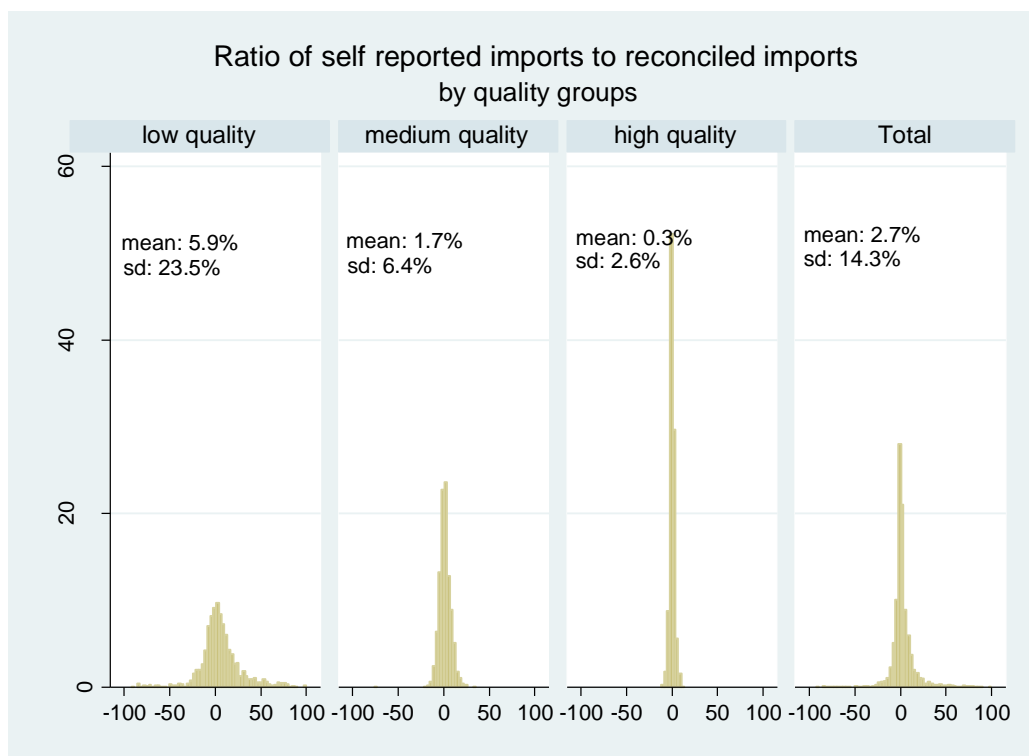
The histograms show that the deviations of reconciled from unreconciled data for high-quality reporters are relatively small for two reasons. First, in trade with low-quality reporters, the data are picked from the country with higher quality: self-reported figures are used for reconciliation. Second, when reconciling the data between high-quality reporters, the differences in reports tends to be smaller than the differences between countries with low-quality data: high-quality reporters tend to report similar values of trade with each other. The low standard deviations indicate homogeneity among high-quality reporters in terms of the similarity between reconciled and unreconciled data.

Low-quality reporters tend to under-report their trade as the positive means of the distributions suggest, which adds more evidence that corruption can explain a share of the discrepancies as this indicates smuggling or under-invoicing. The large standard deviations, on the other hand, indicate that there is a significant idiosyncratic component in the discrepancies, which implies erroneous reporting.

China tends to under-report exports and over-report imports

According to SITC data in 2016, the largest two traders in the world were China followed by the US; this holds using self-reported or reconciled data. We now take a look at these two countries.

China's exports are generally under-reported although the gap has been closing recently (Figure 2.12). This is not related to the fact that a share of China's exports goes through Hong Kong since the plotted data are for China's trade with the rest of the world combined. Ferrantino, Liu and Wang (2012) show strong statistical support for Chinese Value Added Tax evasion for the commodities traded directly between the US and China.

FIGURE 2.12: CHINA'S TRADE WITH THE REST OF THE WORLD ($US BILLIONS)

Note: China started reporting data to SITC in 1984; reconciled data starts from 1962.
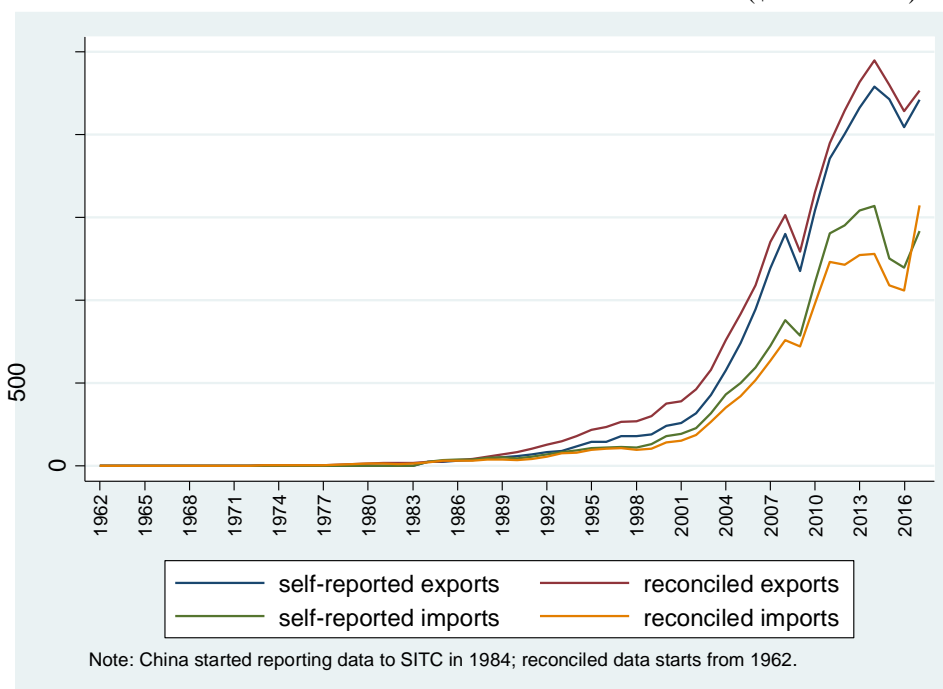
The less investigated issue in the literature is China's data discrepancy on the side of imports. China's reconciled imports are consistently smaller than self-reported throughout most of the period of the study. China tends to over-report its imports

from many high-quality reporters. However, this seems to have changed in 2016. Table 2.5 shows the difference in claims on China's imports for the year 2014.

TABLE 2.5: CHINA'S CLAIMS ON IMPORTS VS ITS PARTNERS' CLAIMS ON EXPORTS TO IT IN BILLIONS OF DOLLARS IN 2014

Partner	partner's TDQI for exports	China's TDQI for imports	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	46.1	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	39.5	China	0.95

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

US self-reported and reconciled trade are similar

US trade data remains relatively unchanged after reconciliation since US TDQI is comparatively high: US data are used for reconciliation in most cases. In the cases where US partners' data are used for reconciliation, the differences in claims are rather small (Figure 2.13).

FIGURE 2.13 : US TRADE WITH THE REST OF THE WORLD ($US BILLIONS)
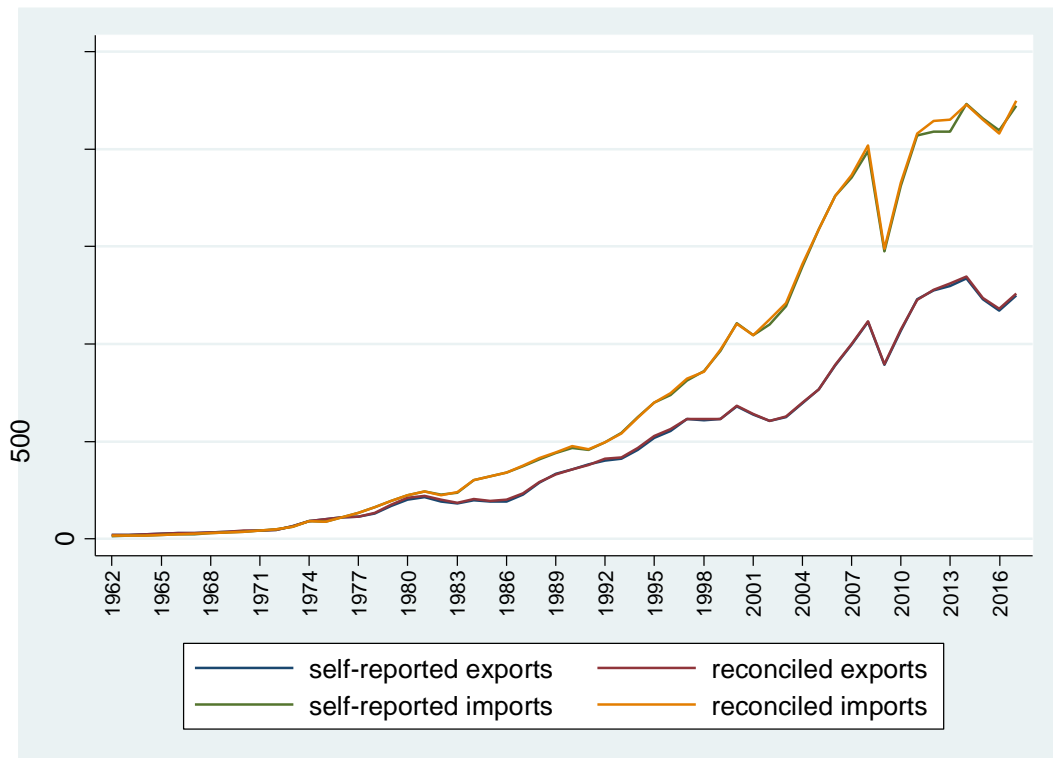
Using reconciled trade is important for understanding international trade

Bahmani-Oskooee et al. (2013) show how substantial data discrepancies can mean a study that uses data reported by a certain partner would reach a different conclusion if it used the data reported by the other (discussed in section 2.1). Using reconciled data in a measure that depends on trade in its construct makes a non-trivial difference as well. For illustration, we calculate trade openness for every

country in 2016, once using reconciled data and once using self-reported data.[16] Table 2.6 presents the findings, where high, medium and low quality refer to the tertiles of countries in terms of average TDQI for their imports and exports. The total sample size is 133 countries: the subsamples are 45 or 44 each.

TABLE 2 6: TRADE OPENNESS CALCULATED USING RECONCILED AND UNRECONCILED DATA

reporter quality	data source	average trade openness
high-quality reporters	unreconciled	59%
	reconciled	59%
medium-quality reporters	unreconciled	67%
	reconciled	71%
low-quality reporters	unreconciled	43%
	reconciled	57%

The less reliable third of countries are on average 14% more open to trade than thought in the literature, which largely depended on self-reported figures.[17] Trade openness for countries with higher data quality does not change materially using reconciled or unreconciled data.

2.7 Robustness checks

A key assumption for TDQI is that all countries are equally likely to trade with different combinations of partners with the same average of quality. However, what if a country trades extensively with low-quality reporters? It would appear as a low-quality reporter itself (low TDQI). From the Melitz (2003) model of trade, extended to a gravity framework by Helpman, Melitz and Rubinstein (2008), firms have to

---

[16] Trade openness is defined as (exports + imports)/GDP. GDP data are in real USD and are retrieved from the World Bank. We break the reporters into three groups (high, medium, and low-quality) based on the average quality of their imports and exports.
[17] From Table 2.6: 57% − 43%=14%

pay a fixed cost to export. Since this fixed cost differs across trading partners and is often higher in more 'difficult' destination countries, it is reasonable to expect that firms in a country like Iceland may not be able to incur the fixed cost of exporting to a country like Mali, while firms in the US might be able to. This means that we would observe the US trading with Mali, while Iceland would not. The TDQI for the US would then be lower than Iceland's if Mali was a low-quality reporter.

So how good is our assumption that all countries are equally likely to trade with different combinations of partners with the same average of quality? To answer this question, we look at domestic TDQI for each trade flow for each country in the year 2016 and compare it with the weighted average of TDQI for its trading partners. Figure 2.14 plots domestic TDQI for importers against the weighted average of their partners' TDQI for exports, and vice versa. The plot shows that the weighted averages of partners' quality for every country are clustered in a certain region (clustered horizontally). That is, whether countries are high-quality or low-quality reporters, the weighted average of their partners' TDQI is relatively the same, suggesting our assumption is reasonable.

But how about the few countries with low partner TDQIs? As the plot shows, the countries with low partner data quality are mostly countries with very small shares in global trade. Most importantly, what are the implications of having some countries that have lower weighted averages of their partners' data quality? Back to our example of the US and Iceland and their trade with Mali. This means this study is more likely to pick Iceland-US trade data from Iceland, instead of the US. As shown in Figures 2.10 and 2.11, as both countries are high-quality reporters, the discrepancies between their claims are small, which explains why US total trade with the rest of the world is roughly the same using reconciled or unreconciled data (Figure 2.13).

FIGURE 2.14 : REPORTER DATA QUALITY AND THE CORRESPONDING WEIGHTED AVERAGE OF
PARTNERS' QUALITY (2016)

Note:
countries with 0 TDQI are countries which did not report data in 2016.
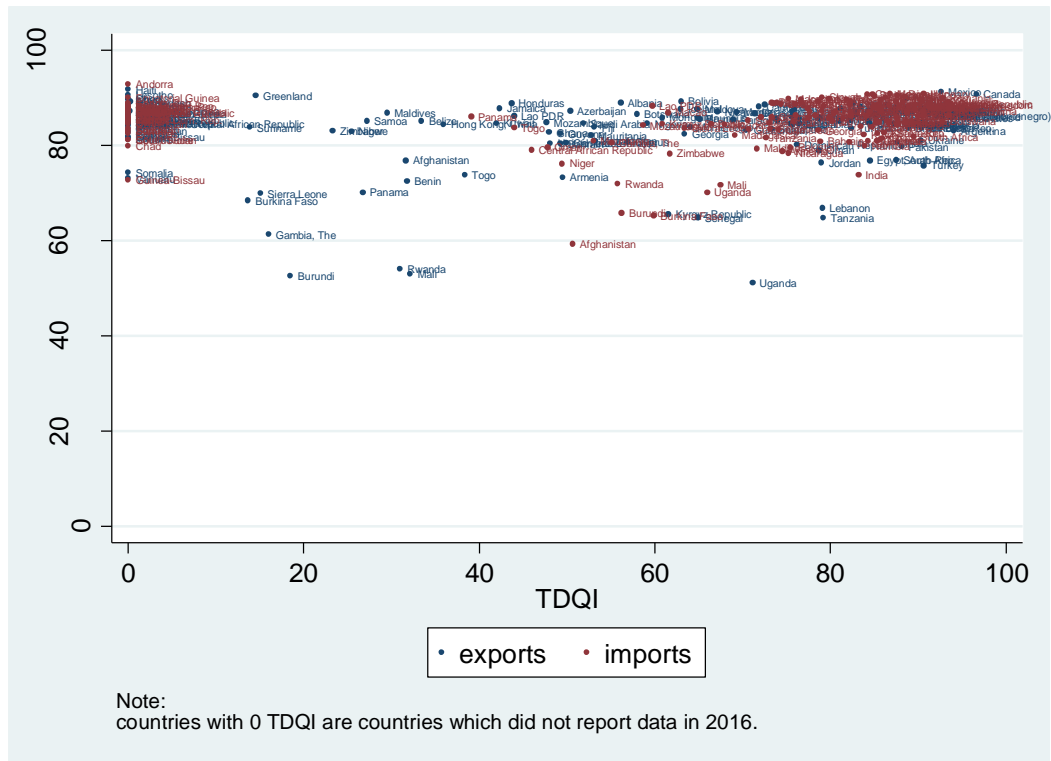
The same applies to years other than 2016. Figure 2.15 plots the distributions of the
weighted average of partners' data quality for different brackets of exporter quality.
The similarity in all three distributions suggests that all exporters, regardless of their
export data quality, have similar weighted averages of partners' data quality. Figure
2.16 confirms the same for the case of importers.

FIGURE 2.15 : WEIGHTED AVERAGE OF PARTNERS' DATA QUALITY CONDITIONAL ON
EXPORTERS' DATA QUALITY (1962-2016)

Distribution of the weighted average of partner TDQIs by exporter quality brackets
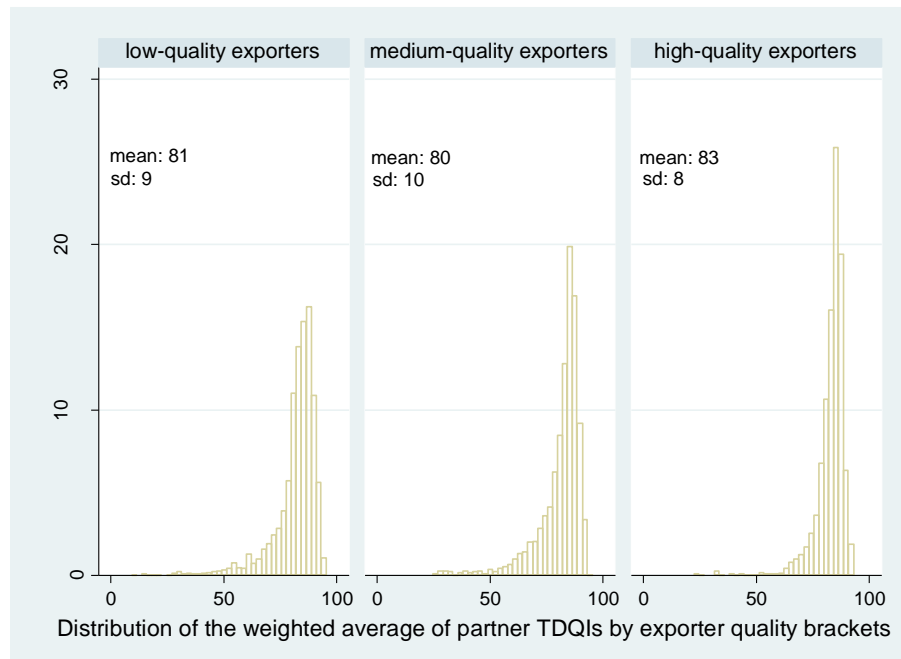
FIGURE 2.16 : WEIGHTED AVERAGE OF PARTNERS' DATA QUALITY CONDITIONAL ON
IMPORTERS' DATA QUALITY (1962-2016)

Distribution of the weighted average of partner TDQIs by importer quality brackets
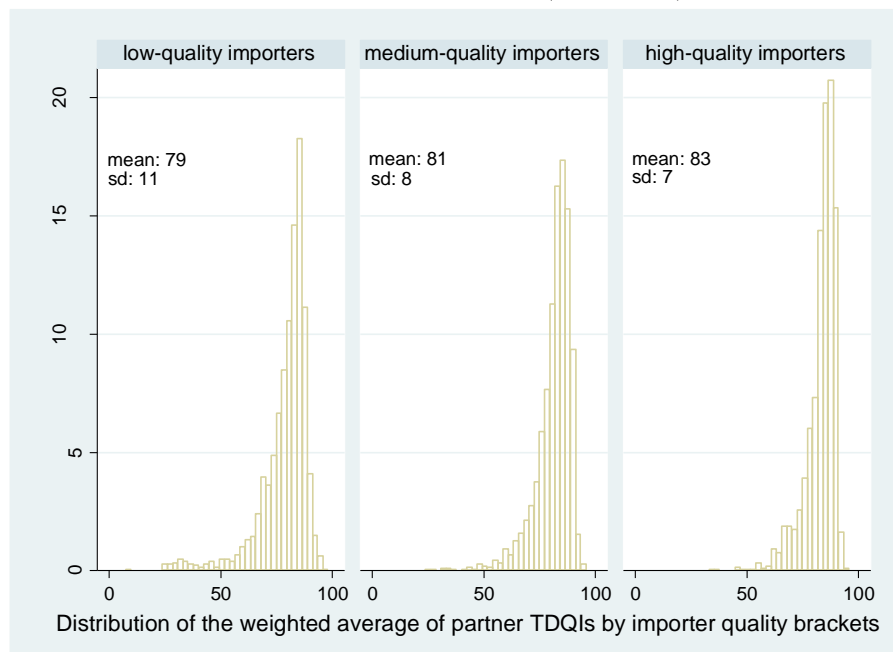
2.8 Conclusion

When two friends, John and Tony tell different stories about what happened between them, it is hard to quickly find out which claim is more reliable. If a third friend, Sarah, tells me she usually agrees with Tony but not with John, evidence starts to mount against John as now two people independently claim he is not telling the truth. If we follow the same logic with many more friends, we should be able to rank everybody in terms of their reliability. When we want to write the history of what happened between each pair of these friends, we should rely on the stories told by the more reliable half of them.

We use this logic to rank countries in terms of their trade data reporting quality and later reconcile trade data through picking the value reported by the more reliable partner in every bilateral relationship.

We confirm the wide belief that importers' data tend to be more reliable. We further show that low-quality reporters tend to under-report their imports and exports, giving evidence of under-invoicing and smuggling. While China tends to under-report its exports and over-report its imports, the US self-reported and reconciled data are roughly the same.

Using reconciled trade is important for understanding international trade. To illustrate, we calculate trade openness for every country in 2016, once using the data reconciled here and once using self-reported data. Low-quality reporters are, on average, 14% more open to trade using reconciled data.

The results are made available for public access.

APPENDIX A

Table A1: TDAI and TDQI for imports and exports for the year 2016

Reporter	TDAI exports	TDAI imports	TDQI exports	TDQI imports
Canada	100	100	97	88
Poland	100	100	95	82
Austria	100	100	94	85
Italy	100	100	94	93
Argentina	99	97	94	89
New Zealand	98	100	94	85
Germany	100	100	94	90
Spain	100	100	94	91
United States	100	100	93	93
Bulgaria	98	97	93	90
France	100	100	93	85
Mexico	98	100	93	87
Japan	100	100	92	85
Sweden	100	100	92	91
Finland	99	100	92	87
Korea, Rep.	100	100	92	86
Romania	100	96	91	93
Indonesia	100	99	91	87
Czech Republic	100	100	91	84
Turkey	100	99	91	89
Switzerland	100	99	91	85
Portugal	100	99	91	91
Slovenia	99	100	90	86
United Kingdom	100	100	90	94
Sri Lanka	99	100	90	85
Hungary	99	98	90	91
Ukraine	99	99	90	86
Thailand	100	100	90	88

Greece	99	97	89	86
Denmark	99	99	89	94
Australia	99	100	89	93
Brazil	100	99	89	91
Ireland	99	100	89	77
Slovak Republic	99	100	89	79
Chile	96	97	89	90
China	100	100	88	86
Pakistan	98	98	88	75
India	100	100	88	83
Serbia	97	99	88	79
South Africa	99	100	88	89
Netherlands	100	100	87	86
Morocco	98	98	87	87
Luxembourg	99	100	86	86
Peru	92	100	85	92
Russian Federation	100	100	85	89
Egypt, Arab Rep.	95	95	84	80
Qatar	96	87	84	71
Estonia	98	100	84	85
Malaysia	98	98	84	85
Norway	100	100	83	80
Belgium	100	100	83	92
Philippines	100	98	83	81
Israel	99	99	83	79
Latvia	100	98	83	85
Tunisia	89	96	82	86
Croatia	93	100	82	94
Macedonia, FYR	88	98	81	75
Belarus	96	100	81	89
Tanzania	90	99	79	73
Lebanon	94	99	79	87
Jordan	90	96	79	82
Myanmar	94	96	79	78

Oman	89	92	79	48
Lithuania	99	100	76	88
Colombia	94	100	76	93
Dominican Republic	81	100	76	91
Bosnia and Herzegovina	83	100	76	74
Nepal	84	98	76	82
Vietnam	86	87	76	77
Uruguay	91	99	76	82
Namibia	82	94	75	84
Malta	94	98	74	73
Costa Rica	86	100	74	90
Singapore	100	100	73	86
Algeria	83	96	73	88
Cambodia	89	97	72	78
Ecuador	88	97	71	88
Uganda	82	97	71	66
Guatemala	78	97	70	84
Madagascar	80	97	69	69
Paraguay	90	96	69	73
Nicaragua	76	95	67	75
Brunei	72	93	65	76
Mauritius	74	100	65	88
Senegal	84	94	65	73
Moldova	86	100	65	77
Montenegro	77	99	64	79
Georgia	78	90	63	79
Bolivia	74	99	63	76
Iceland	80	97	63	74
Kyrgyz Republic	78	99	62	61
Mongolia	66	99	61	90
Nigeria	70	97	59	78
Botswana	88	99	58	93
Cyprus	95	100	57	82
Albania	78	100	56	88

Kazakhstan	88	97	53	83
Mauritania	64	91	53	53
Saudi Arabia	89	94	52	84
Azerbaijan	85	92	50	76
Cameroon	70	96	50	74
Armenia	73	99	49	74
Bahrain	78	99	49	79
Guyana	69	96	49	85
United Arab Emirates	98	99	48	82
Ghana	75	96	48	73
Mozambique	78	100	48	59
Lao PDR	56	76	44	60
Honduras	65	92	44	68
Jamaica	54	83	42	72
Kuwait	88	99	42	84
Togo	53	94	38	44
Hong Kong, China	98	100	36	91
Mali	56	92	32	67
Benin	70	94	32	66
Afghanistan	35	69	32	51
Rwanda	79	97	31	56
Maldives	41	94	30	72
Panama	72	73	27	39
Niger	60	91	25	49
Zimbabwe	60	95	23	62
Burundi	46	89	18	56
Gambia, The	44	93	16	55
Sierra Leone	29	90	15	63
Suriname	65	91	14	80
Burkina Faso	64	98	14	60
Central African Republic	24	82	6	46
Iraq	11	0	3	0
El Salvador	79	100	0	85
Kenya	0	0	0	0

South Sudan	0	0	0	0
Haiti	0	0	0	0
Guinea	0	0	0	0
Bhutan	0	0	0	0
Turkmenistan	0	0	0	0
Yemen	0	0	0	0
Liberia	0	0	0	0
Djibouti	0	0	0	0
Somalia	0	0	0	0
Syrian Arab Republic	0	0	0	0
Libya	0	0	0	0
Korea, Dem. Rep.	0	0	0	0
Bangladesh	0	0	0	0
Comoros	0	0	0	0
Malawi	0	0	0	0
Chad	0	0	0	0
Congo, Rep.	0	0	0	0
Cuba	0	0	0	0
Tajikistan	0	0	0	0
Uzbekistan	0	0	0	0
Cote d'Ivoire	0	0	0	0
Zambia	0	0	0	0
Sudan	0	0	0	0
Congo, Dem. Rep.	0	0	0	0
Angola	0	0	0	0
Venezuela	0	0	0	0
Guinea-Bissau	0	0	0	0
Papua New Guinea	0	0	0	0
Gabon	0	0	0	0
Iran, Islamic Rep.	0	0	0	0
Lesotho	0	0	0	0

Note: Some countries have 0 TDQI because they have not reported data in 2016.

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-Aruba, Saint Kitts and Nevis, Saint Lucia, Montserrat, Martinique, Mayotte, New Caledonia, Other Asia not elsewhere specified, Palau, Occupied Palestinian Territory, French Polynesia, Reunion, Solomon Islands, Saint Pierre and Miquelon, Sao Tome and Principe, Seychelles, Turks and Caicos Islands, Tonga, Trinidad and Tobago, Tuvalu, Saint Vincent and the Grenadines, Virgin Islands (U.S.), Wallis and Futura Islands, Eritrea, 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, Saint Kitts and Nevis, Saint Lucia, Montserrat, Martinique, Other Asia not elsewhere specified, Sao Tome and Principe, Turks and Caicos Islands, Trinidad and Tobago, Unspecified, St. Vincent and the Grenadines, British Virgin Islands, Marshall Islands, Special Categories, Saint Pierre and Miquelon, Holy See, Aruba, American Samoa, Antarctica, French Southern Territories, Bouvet Island, Cocos (Keeling) Islands, Cook Islands, Cape Verde, Christmas Island, Western Sahara, Falkland Island, Faeroe Islands, Federated State of Micronesia, Guam, Heard Island and McDonald Island, British Indian Ocean Territories, Kiribati, Northern Mariana Islands, Mayotte, New Caledonia, Norfolk Island, Niue, Nauru, Pitcairn, Palau, Occupied Palestinian Territory, French Polynesia, Reunion, South Georgia and the South Sandwich Islands, Saint Helena,

Solomon Islands, San Marino, Seychelles, Tokelau, Tonga, Tuvalu, United States Minor Outlying Islands, Virgin Islands (U.S.), Wallis and Futuna Islands, Neutral Zone, US Miscellaneous Pacific Islands, Saint Kitts-Nevis-Anguilla-Aruba, Former Panama Canal Zone, Ryukyu Island, Pacific Islands, Saint Barthélemy, Bonaire, Saint Maarten (Dutch part), British Antarctic Territory, Sikkim, Eritrea, and Ethiopia (excludes Eritrea). Note that we include Ethiopia (including Eritrea).

CHAPTER 3:
Why You Should Use High Frequency Data to Test the Impact of Exchange Rate on Trade

3.1 Introduction

Among a variety of theories attempting to explain the impact of exchange rate movements on trade balance, J-Curve Theory gained most of the attention. The theory suggests that, as a result of currency depreciation, a country's trade balance follows a J-shaped pattern as it deteriorates in the short run and recovers to a higher level in the long run (Figure 3.1). The initial deterioration is due to the presumed stickiness of the traded quantities for reasons such as previously signed contracts on the supply side, and habit formation on the demand side. Actually, if the pattern of short-run deterioration and long-run improvement of the trade balance as a result of currency depreciation occurs (the J-Curve Phenomenon holds) it indicates that the Marshall-Lerner Condition (MLC) is met, too (Clarke & Kulkarni, 2010).

In simplified terms, the MLC states that if the central bank devaluates the currency aiming to improve trade balance, the demand for the nation's imports and exports should be elastic enough to cause a favourable volume effect. The condition under the simplest circumstances states that the sum of the absolute value of the two demand elasticities for imports and exports must exceed one (Brown & Hogendorn, 2000). Although this paper tests only the J-Curve Theory, its rationale is still applicable to other theories attempting to test the impact of exchange rate on trade.

FIGURE 3.1: THE J-CURVE PHENOMENON
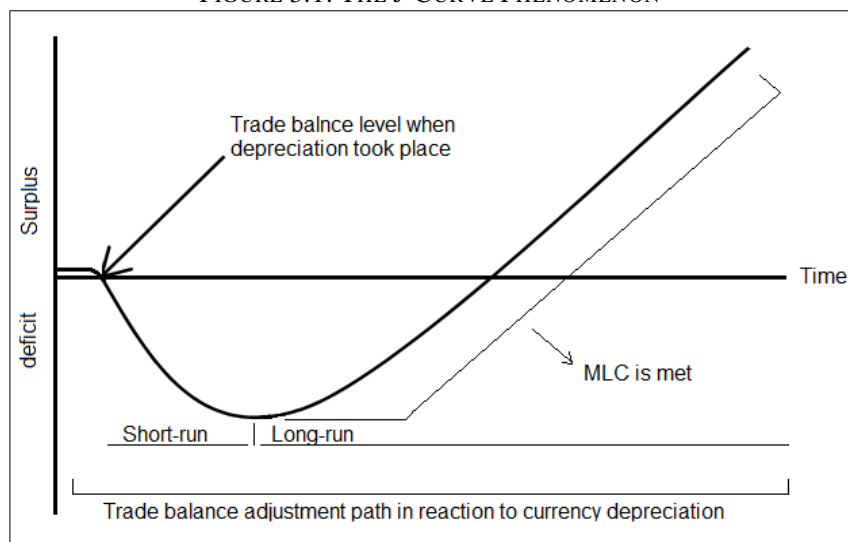
Source: Author's depiction

Bahmani-Oskooee and Hegerty (2010) review the literature on the J-Curve Theory and show how studies in the literature vary primarily in terms of the data used in investigating the Theory. The first group of studies used aggregate trade of a country's trade with the rest of the world combined. The second group investigated the Theory bilaterally, stating that studies in the first group suffered from 'aggregation bias' as aggregating trade might obscure the dynamics of individual bilateral relation. The third group reduced the aggregation bias further by using commodity level data. Most studies, to varying degrees, find support of the J-Curve phenomenon (Bahmani-Oskooee and Hegerty 2010).

Among the hundreds of studies which tested the J-Curve Theory, the most common empirical methodology is a combination of the Autoregressive Distributed Lag (ARDL) approach to cointegration and the Error Correction Model (ECM) (Pesaran et al., 2001). The desirability of this methodology in the literature is largely because it does not require a prior determination of which regression variables are I(0) and which are I(1), as well as the direct and easy inference of short and long run impacts of exchange rate on trade balance.

However, the literature suffers from a key problem: the assumption that data frequency has no role in empirically investigating the relationship between exchange rate and trade. In this paper, we advocate the use of high frequency data when testing the J-Curve Theory – or other theories of similar nature – for four reasons.

First, while it is true that cointegration is a long-run notion, the short-run impact of exchange rate and error correction on trade is a vital component of the ECM. Low frequency data, such as annual or quarterly, might obscure much of the evidence needed for understanding the short-run dynamics of the relationship as high and low values might average out over the course of one period in low frequency data.

Second, regarding cointegration itself, it is thought by many researchers in the field that using low frequency data is as good as using high frequency data for testing long-run association. As stated by Bahmani-Oskooee (1996), '…using annual data over the 1960-90 period in this paper is as good as using quarterly or monthly data over the same period'. However, cointegration tests with low frequency data have lower power. As shown in a Monte Carlo experiment by Zhou (2001), the loss of power applies especially for small samples of less than 50 years (as is the case in most macroeconometric applications) and for models which involve autoregressive terms. In other words, low frequency data might suppress the evidence of true cointegration.

Third, quantities and prices might change after exchange rate movements in a period as short as instantly. Using quarterly or annual data implicitly assumes that short run changes in trade values and quantities take at least three months or a year to take place. Fourth, using annual data for ARDL cointegration increases the likelihood of including I(2) variables owing to the smoothing imposed on the data by the aggregation over time. This may lead to spurious cointegration conclusions.

To address these issues, our paper tests the J-Curve Theory between the US and Canada using the ARDL-ECM approach for two identical periods using monthly, quarterly, and annual data.

The findings confirm that (a) low frequency data might suppress the evidence of cointegration, (b) higher frequency allows better understanding for the short-run dynamics of the relationship, and (c) lower frequency might cause some series to become I(2), which invalidates the ARDL cointegration approach.

3.2 Methodology

We choose US-Canada bilateral trade relation as an example to investigate the research questions since it continues to be one of the most significant global trade relations in terms of total merchandise.[18] Additionally, the data quality is perceived to be high (low trade data discrepancy) and are available for a long period.

Since the J-Curve Theory is being tested using real exchange rate (REX), we use the whole available sample of data starting from 1962, rather than since the breakdown of the Gold Standard in early 1970s. As shown in Figure 3.2, especially during a financial crisis like that around 2008, annual exchange rate obscures some of the information available in monthly data. Another problem of similar nature might arise from end of period or mid-period sampling of low frequency series.

---

[18] SITC data, described in section 2.2, suggests Canada-US trade relation was the third largest in the world in 2016 in terms of the sum of exports and imports.

FIGURE 3.2 : ANNUAL VERSUS MONTHLY CAD/USD REAL EXCHANGE RATE 1962-2015
(PERIOD AVERAGE)
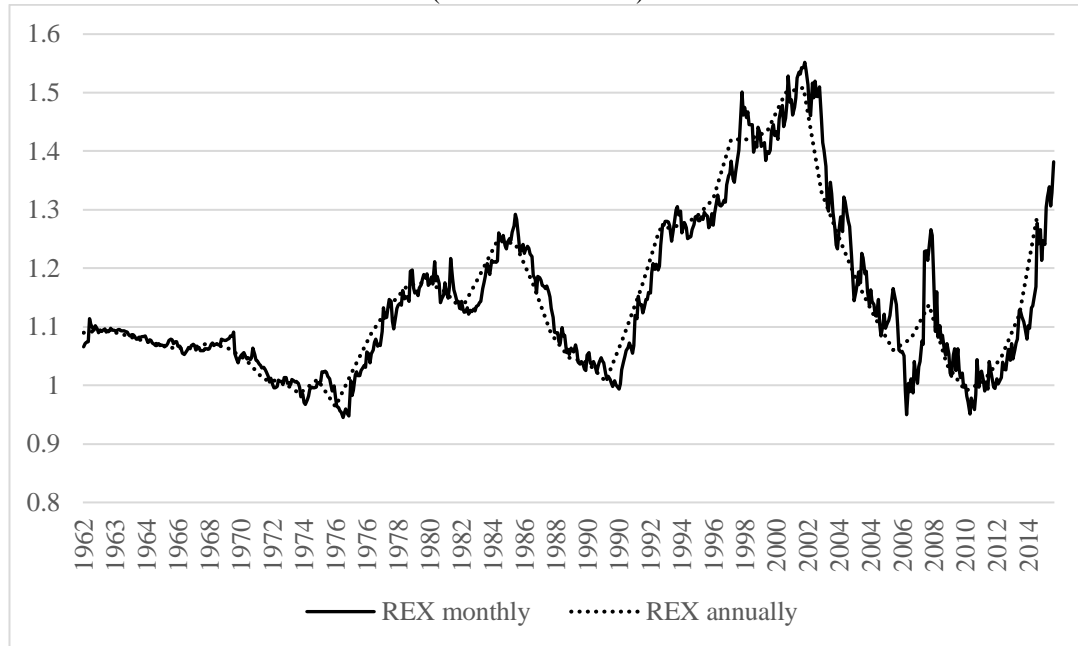
The model used for investigating the J-Curve Theory is identical to that used in many studies in the field, which is a combination of ARDL cointegration as developed by Pesaran et al. (2001), and the ECM. We select two time periods for the empirical analysis. The first extends along the whole available period of 1962–2015. The second follows Bahmani-Oskooee and Wang (2008) on the same two countries and spans from 1962 to 2004. We use monthly, quarterly, and annual data for the same variables. That is, we investigate six variations of the following model. Two time periods, using three different data frequencies for each.

$$\Delta LnTB_t = \alpha + \sum_{k=1}^{n_1} \emptyset_k \Delta lnTB_{t-k} + \sum_{k=0}^{n_2} \beta_k \Delta lnUSGDP_{t-k} +$$
$$\sum_{k=0}^{n_3} \gamma_k \Delta lnCGDP_{t-k} + \sum_{k=0}^{n_4} \varphi_k \Delta lnREX_{t-k} + \delta_1 lnTB_{t-1} + \delta_2 lnUSGDP_{t-1} +$$
$$\delta_3 lnCGDP_{t-1} + \delta_4 lnREX_{t-1} + u_t \qquad [3. 1]$$

where TB is trade balance defined as US exports to Canada divided by US imports from Canada. REX is real exchange rate defined as the number of Canadian Dollars per USD adjusted by the price levels in each country as measured by consumer

price index. CGDP and USGDP are the real GDPs of Canada and US. These two variables are not available on monthly frequency, so we apply repeated quarterly data on these in the monthly regressions since they are used only as controls.[19] All the data are retrieved from IMF-IFS Database. The data are available here.

Besides the first differences, the model includes a linear combination of lagged level variables as a proxy for lagged error-correction term ECt-1. The term reflects the speed of adjustment of the dependent variable (return to equilibrium) after a deviation has occurred in the independent variables. To justify retention of the lagged level variables in the model, Pesaran et al. (2001) proposed the use of standard F-test to determine the joint significance of lagged level variables with a new set of critical values. They tabulated the critical values for lower and upper bounds depending on the order of integration of the included variables. If the computed F statistic is larger than the upper-bound tabulated value, the null hypothesis of no cointegration can be rejected.

After establishing cointegration, the short-run impacts can be inferred from the coefficients of the differenced variables. The short-run coefficients are reported once the test passes all the diagnostic tests of LM autocorrelation residual-based test, White Test of heteroscedasticity, and Ramsey misspecification RESET Test. Long-run impacts can be calculated from the same specification in 2.1. However, this study is more interested in short term dynamics as they are sufficient to test whether the J-Curve Theory holds or not.

---

[19] Using repeated quarterly observations in a monthly time series is not ideal. However, we have tested the sensitivity of the findings to the inclusion of domestic and partner GDPs altogether and the results did not change materially.

3.3 Findings

Before running our six regressions, we first ensured that all our variables are either I(1) or I(0) using the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. The variables are all found to be I(1), regardless of the data frequency, using either test. The optimal lag structure for each regression is chosen through Schwarz Information Criterion. The following table summarizes the estimates provided by each regression. For brevity, we report the short-run estimates only for the variable of interest, LnREX (Table 3.1).

TABLE 3. 1: REGRESSION OUTPUTS OF DEPENDENT VARIABLE $\Delta LnTB_t$

Period	1962-2015			1962-2004		
Frequency	Monthly	Quarterly	Annually	Monthly	Quarterly	Annually
ARDL (F)	3.95	4.61	4.19	4.2	2.07	3.58
Cointegrated?	Yes	Yes	Yes	Yes	No	No
$\Delta LnRex_t$			0.31(1.49)		-0.19(-1.03)	
$\Delta LnRex_{t-1}$		0.29 (2.37)**				0.24(0.96)
$\Delta LnRex_{t-2}$	0.29 (2.31)**			0.36(1.86)*		
$\Delta Ln REX_{t-3}$		0.28(2.22)**				
LnREXt	-0.44(-6.12)***	-0.27(-2.72)***	-0.42(-3.07)***	-0.55(-4.78)***	-0.54(-3.44)***	-0.27(-1.14)
LnUSGDP	0.99 (3.84)***	0.18(0.42)	1.17(4.70)***	0.43(0.99)***	0.59(0.91)	-1.27(-4.6)***
LnCGDP	-1.03(-3.9)***	-0.98(-3)***	-1.19(-4.68)***	-1.26(-5.25)***	-1.3(-3.79)***	1.19(4.11)***
Intercept	-0.01(-0.31)	1.06(3.54)***	-0.01(-0.17)	1.07(2.54)***	0.98(1.75)*	0.08(0.66)
Adj. R^2	41%	33%	21%	41%	34%	35%
LM Test	1.8	0.69	1.49	2.79*	1.85	2.32

*,**,*** indicate significance at 10%, 5%, and 1% respectively. Numbers inside parenthesis are t-ratios. Critical ARDL F values are 3.77 and 4.35 at 10% and 5% levels of significance, respectively. LM serial correlation test reported values are F stats of Breusch-Godfrey Test. HAC robust standard errors are applied for hypothesis testing when autocorrelation is present.

The results indicate the presence of the J-Curve phenomenon for all four cointegrated regressions as REX appears to have positive effect in the short run and negative effect in the long run. That is, a depreciation of USD is expected to worsen the US trade balance with Canada in the short-run and improve it in the long-run.

The findings confirm that aggregation over time suppresses the evidence of the long-run effect of exchange rate on trade. This is because no cointegration could be established only using annual or quarterly data for the period 1962-2004. In line with Zhou (2001), the results suggest that lower frequency data might capture the long-run impact of exchange rate on trade only when the period is long enough. The annual sample of 1962-2004 was also used by Bahmani-Oskooee and Wang (2008) for the same trade relation but using sectoral trade data. Thus, it is also likely for them to have not found enough support for cointegration due to aggregation over time. They found support of cointegration for 80 sectors out of 152.

Additionally, for the regressions of 1962-2015, we see that monthly data can better explain the dynamicity of the relationship. It shows that the impact of REX on TB appears after two months, while quarterly and annual results are less accurate and show that the impact kicks in any time within the first quarter, or year, respectively. As reflected in adjusted R2, we also note that monthly data – by using more information – provides a better fit of the same variables over the same period.

Finally, using low frequency data might also cause a rarely discussed problem, which impacts the application of ARDL cointegration. That is the inclusion of I(2) series. This is particularly problematic for studies in this field as there has been a tendency not to test for unit root at all since ARDL allows for the inclusion of I(0) and I(1) series without prespecifying the order for the test. For instance, Bahmani-Oskooee and Ratha (2004) state that: 'The new critical values tabulated by Pesaran et al. (2001) do take into consideration the stationarity properties of the series. For this reason, there is no need for pre-testing of unit roots'.

Table 3.2 selects a particular sample (1984-2009) from our study as an example of how aggregation over time might make some series I(2). The results of unit root tests are compared with the other time periods tested in the regressions presented in Table 3.1.

TABLE 3. 2: UNIT ROOT TEST FOR THE VARIABLES FOR DIFFERENT PERIODS

	1962-2015 (54 years)			1962-2004 (43 years)			1984-2009 (26 years)		
Series\frequency	M	Q	A	M	Q	A	M	Q	A
LnRex	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(2)
LnTB	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)
LnUSGDP	NA	I(1)	I(1)	NA	I(1)	I(1)	NA	I(1)	I(2)
LnCGDP	NA	I(1)	I(1)	NA	I(1)	I(1)	NA	I(1)	I(2)

Results are confirmed by ADF as well as PP. Frequencies: M is monthly, Q is quarterly, A is annual

It should be noted that many studies tested the impact of exchange rate on trade using annual data for 26 years or shorter.

3.4 Conclusion

This paper argued that the impact of exchange rate on trade should be investigated using monthly data and showed that using low frequency data might suppress the evidence of a true relationship. It was further shown that using high frequency data allows better understanding of the dynamics of the relationship and reduces the likelihood of using I(2) series.

A caveat of the approach taken in this paper is the use of aggregate bilateral trade data instead of using commodity-level data. Implicit in our use of aggregate trade is the assumption that exported goods are priced in the currency of the exporter, which is not always the case. Gopinath et al. (2010) show that the currency used in pricing the traded goods affects the relationship between the exchange rate and trade volumes.

CHAPTER 4:

Housing Leverage and Consumption Expenditure: Evidence from New Zealand Microdata[20]

4.1 Introduction

Understanding household consumption spending is crucial for modelling business cycles and designing macroeconomic policy. Traditional theories of consumption suggest that income and wealth are important determinants of household spending (see e.g. Fisher, 1930; Friedman, 1957). In New Zealand, housing wealth and housing mortgage debt represent a substantial proportion of household assets and liabilities, respectively. As a result, large swings in house prices have a significant impact on household balance sheets and may also have a material impact on consumption spending decisions. In this paper, we investigate the impact of housing wealth and housing debt on household consumption spending.

FIGURE 4. 1 : NEW ZEALAND HOUSE PRICE INFLATION

Source: Reserve Bank of New Zealand

New Zealand house prices rose rapidly in the early 2000s. The country as a whole experienced double-digit house price inflation in the six years leading up to the Global Financial Crisis (GFC). After a brief, but significant, downward adjustment during the GFC, house price inflation increased again, especially in Auckland, the most populous city in the country. The increase in house prices corresponded with an increase in household debt over the same period. As shown in Figure 4.2, the mean debt-to-income ratio (DTI) has increased in most income quantiles over the last decade. The increase in DTIs has been particularly apparent for the quantile of households with the lowest income. Finally, using aggregate data for the period 1982-2016, Wong (2017) shows that the impact of housing wealth on household consumption is weakening over time. Motivated by these observations, we investigate whether household leverage can help to explain the weakening housing wealth effect on consumption documented by Wong (2017).

FIGURE 4. 2 : MEAN OF DTI BY INCOME GROUP
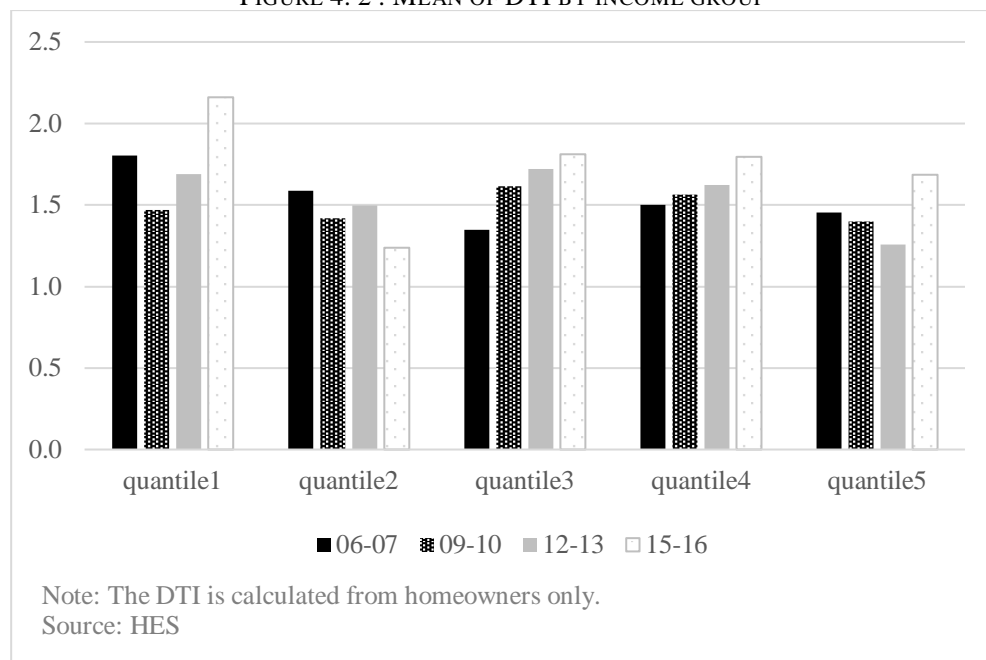
Note: The DTI is calculated from homeowners only.
Source: HES

We use microdata from Statistics New Zealand's 'Household Economic Survey' (HES) to investigate how household leverage affects the marginal propensity to consume (MPC) out of housing wealth. HES data provide detailed information on household spending, income and loans. We combine HES data with Real Estate Institute of New Zealand (REINZ) micro house price data to address the endogeneity issues that arise from using household-level cross-sectional data.

Empirically, estimating the effect of housing wealth changes on household expenditure faces two types of endogeneity issues. First, any evidence of an association between housing wealth variations and consumption changes could be driven by unobservable confounding factors such as future income expectations or household preferences. Second, naive regressions with total household spending can suffer from reverse causality, in which high housing-related spending leads to higher property values.

This paper seeks to address both endogeneity challenges and illustrate how the data choices relating to the variable of housing have econometric implications, which in

turn affects our conclusions regarding the impact of housing wealth on consumption. To deal with the endogeneity arising from confounding factors, we use local average house prices as an instrument for individual house prices. The instrumental variable (IV) estimation improves the estimated elasticity of housing wealth to expenditures through two channels. First, the instrument exploits cross-locality variation in house prices to extract the exogenous component of house price differences at the individual level. Cross-locality differences in house prices are mainly driven by local amenity differences. Those factors are arguably exogenous to household characteristics, such as risk aversion, housing preference, and future income expectations. Second, the instrument also helps to reduce the estimation bias caused by measurement errors in the house prices reported in the HES as HES house prices are estimated by local governments for the purpose of levying rates and therefore do not accurately reflect the values of houses.

The Household Economic Survey covers approximately 3000 households each quarter. In contrast, the REINZ housing sales data cover about 95% of all housing transactions in New Zealand and the house price data are highly accurate. To deal with the second cause of endogeneity, we take advantage of the detailed reporting on household expenditures in the HES and separate household spending from those related to housing improvements, maintenance and mortgage-related costs. These types of expenditures drive house values higher and lead to reverse causality. We conduct our regressions on non-housing consumption expenditures.

In the empirical analysis, we first assess the validity of average local house prices as an instrument for individual house prices. The first stage regression suggests that the instrument can explain up to 22% of the variation in individual house prices reported in HES. We then run a benchmark regression of total household expenditure excluding housing-related spending on housing wealth. The IV estimation suggests that using household-level prices leads to downward bias, which is the result of various causes of endogeneity issues discussed above. The

average MPC out of a one-dollar increase in exogenous housing wealth is around 2.2 cents. All regressions control for income, household characteristics, and regional and time fixed effects. We also split non-housing expenditure into durables and non-durables. In line with other studies in the literature, such as Mian, Rao and Sufi (2013), we find that durable consumption is more sensitive to changes in housing wealth than non-durables.

We then focus on the role of household leverage in determining the MPC out of housing wealth. In this analysis, we study how leverage measures, such as the loan-to-house-value ratio (LVR) and the DTI, affect the estimated MPC out of housing wealth. Overall, we find that household leverage weakens the MPC associated with housing. To examine the robustness of these findings, we investigate whether household spending responds differently depending on the age and type of home ownership. The findings confirm that the consumption of mortgagors is less sensitive to housing wealth as compared to outright homeowners. The regression with an age-housing wealth interaction also shows that the response of younger households to changes in their housing wealth is weaker than the response of older households, which tend to be less leveraged.

Mian, Rao and Sufi (2013) (MRS hereafter) use county and zip code-level data from the United States (US) to estimate the MPC out of housing equity shocks. They obtain MPC estimates in the range of 5 to 7 cents for every dollar change in housing net worth. This estimate is higher than our result of 2.2 cents. In addition, in contrast to our finding, MRS show that consumption responses to wealth shocks are stronger in highly indebted counties.

There are several possible reasons why our results differ from those of MRS. First, MRS construct a pseudo panel of data by aggregating geographic areas into county-level data units. Conversely, while they estimate the aggregate MPC at the county level, our estimate is the MPC at the individual household level. As shown in Yao, Fagereng, and Natvik (2015), using detailed Norwegian household data,

aggregating household level data into municipality or county levels magnifies estimated MPCs relative to their micro-level counterparts.

Second, MRS study the period 2006-2009, during which the US housing market experienced a severe downturn, while in our New Zealand sample, house price inflation was mostly positive. The difference in our estimated interaction coefficients could suggest that the role of leverage on the housing wealth effect is asymmetric, depending on whether house prices are increasing or decreasing.

More broadly, our paper is closely related to a growing literature on the housing wealth effect. A large housing wealth effect relative to non-housing wealth is documented by Case, Quigley and Shiller (2005) and Carroll, Otsuka and Slacalek (2011). Using aggregate data, the estimated MPC out of housing wealth in these studies range from 3 to 5 cents on a dollar gain in housing wealth. Studies that use microdata reveal a more detailed picture about the housing wealth effect. For example, Campbell and Cocco (2007) use the UK Family Expenditure Survey and show that the elasticity of consumption to house prices depends on age and tenure types. In line with our study, MPCs are larger for older homeowners, perhaps because lifetime horizon effects are at play. For the case of New Zealand, Smith (2010) also used the Household Economic Survey to investigate the direct impact of housing wealth on consumption (for the period 1984-2007). Unlike Smith (2010), this paper focuses on the moderating role of leverage in testing the impact of housing wealth on consumption.

Since the GFC, a new strand of papers has emerged focusing on the role of household leverage in shaping the housing wealth effect on consumption. Dynan (2012) studies the direct impact of debt on consumption using household-level panel data and shows that leverage lowers household consumption. Her study, however, does not investigate the moderating role of leverage in the impact of housing wealth on consumption.

Motivated by this US literature, international evidence has been fast accumulating, revealing additional insights. Yao, Fagereng, and Natvik (2015) use detailed Norwegian household data and find that housing leverage, measured by the loan-to-value ratio, amplifies the housing wealth effect by about 19-30 cents. Hviid and Kuchler (2017) use a large Danish household panel dataset and document asymmetric MPCs out of positive and negative house wealth shocks. Household consumption appears more sensitive in response to negative housing wealth shocks than to positive ones.

The asymmetric housing wealth effect was suggested earlier by Engelhardt (1996) and Skinner (1989, 1994) using US microdata. Our paper documents a different asymmetric housing wealth effect, when controlling for the household leverage ratio. As stated above, we find that New Zealand household spending is less sensitive to housing wealth when the leverage ratio is high. We provide micro evidence that rising household leverage has contributed to the declining MPC out of housing wealth that Wong (2017) observes in the aggregate data.

4.2 The Household Economic Survey in New Zealand

The HES collects comprehensive data on household residents living in permanent dwellings. The HES surveys household members aged 15 and above. The HES covers multiple aspects of household economics including highly disaggregate household expenditures, income, and loans. The survey also covers demographics, home ownership status, and house values. The data are stratified by different population benchmarks including age, sex, population per region, two-adult and non-two-adult households, and people of Maori ethnicity. This stratification guarantees proper weighting of households and a high degree of comparability across time since the data are cross-sectional rather than longitudinal. Data are collected in one-year waves extending from July to June.

Our sample of households excludes the households that have more than one family. In line with the literature, our study is primarily interested in non-housing expenditures. The focus on non-housing expenditure is to break the reverse causality between housing expenditures and house prices. The excluded housing expenditures are expenses on house maintenance, improvements, and mortgage repayment. Smith (2010) shows that the level of 'ex-housing expenditure' from HES is tightly correlated with household consumption in the System of National Accounts.

The HES expenditure data are only disaggregated into housing and non-housing components triennially. We, therefore, focus our analysis on the four waves with detailed expenditure data only. These waves are 2006-2007, 2009-2010, 2012-2013, and 2015-2016. We also use the disaggregate expenditure data to break our non-housing expenditures into durables, and non-durable components. Appendix C lists the items of expenditures that fall into each of these categories.

The address of each household in the survey is reported at different levels of aggregation. The population of New Zealand is broken down into 47,062 meshblocks (MB), 2,020 area units (AU), and 65 territorial authorities (TA). In the empirical analysis, we control for regional fixed effects by using TA dummies and use the average house selling prices at the AU level.[21]

The HES does not report complete wealth-related data in the triennial waves which contain detailed expenditures. For the particular purpose of this paper, however, the HES reports the rateable value of the primary dwelling of the household and the year it was valued. The primary dwelling is the dwelling occupied by the

---

[21] For more information about geographic boundaries, visit:
http://www.stats.govt.nz/browse_for_stats/Maps_and_geography/Geographic-areas/digital-boundary-files.aspx.

respondents at the time of the interview. The rateable value of the dwelling is estimated by the territorial authority for levying rates. For the dwellings rated in years prior to the survey date, we use REINZ data on house price inflation to ensure all house values are up to date. We inflate house prices at the TA level.

Households fall into one of three housing tenures: renters; owners with a mortgage; and owners without a mortgage. Since this study is interested in homeowners only, we drop the renters from the study. As the house values are available for the primary property only, we exclude households with multiple properties from our analysis. This is done through two stages. First, we exclude any household that owns a house and receives rental income on another property. Second, we exclude the households with LVRs above 0.8. These exclusions lead to a 13% reduction of the total sample size but make the LVR figure more sensible.[22] This is because it is most likely that only households with multiple properties can have LVRs higher than 0.8, especially after the Reserve Bank of New Zealand introduced LVR restrictions in 2013.[23] The overall sample size used for the empirical analysis includes 4644 households.

To capture the actual income of each household, we use the gross disposable annual income data reported by the New Zealand Treasury, which is based on HES raw data. Finally, inflation-adjusted house prices, disposable income, and debt data are used to construct two different measures of household leverage. First, DTIs are constructed by using total household debt and disposable income. Second, LVRs are computed as the ratio of total household debt over the primary house value at the time of survey. Both measures are based on outstanding debt, rather than at the

---

[22] Before excluding those observations, the average LVR was 31.3, which means there are some extreme values in the sample.

[23] The choice of this threshold was subjective and based on what we believed is sensible. The results were not very sensitive to the choice of LVR exclusions ranging from 0.8 to 0.95, i.e. 5% deposit.

time of loan origination, to capture the actual level of leverage at the time of interview.

4.2.1 Descriptive statistics

Table 4.1 reports the mean of the main variables in our regression. We first show them for the full sample and then present them for each wave. As our later analysis will focus mainly on households who own their homes, we only report the statistics with respect to homeowners in Table 4.1. In general, the age and the size of households are stable across waves.[24] Over time, expenditures have increased substantially, along with disposable income. However, during the same period, the increase in house prices resulted in a significant rise in debt as mortgages constitute the main share of household debt. Income growth has lagged behind the increase in debt over time, as illustrated by the upward trend in the DTI ratio. The average LVRs reported in Table 4.1 are lower than those that are typical when loans are first originated because the averages here reflect borrowers who have paid down some of their debt, and also include households that are entirely debt free.

---

[24] The main respondent to the survey is deemed to be the household head.

TABLE 4. 1: DESCRIPTIVE STATISTICS FOR HOMEOWNERS OVER TIME

	all waves	2006/07	2009/10	2012/13	2015/16
Sample size	4,644	1,201	1,180	1,196	1,067
Age of household head	57.2	56.2	57.0	58.0	57.7
Number of persons	2.3	2.4	2.3	2.3	2.4
Non-housing expenditure	41,841	37,152	38,546	44,060	49,128
Disposable income	66,024	51,402	62,062	71,856	82,605
Housing wealth (using the HES data)	418,784	357,868	352,661	405,461	592,396
Total debt	135,382	102,634	131,681	136,539	184,977
DTI	1.9	1.8	2.0	2.0	2.1
LVR	0.32	0.29	0.33	0.34	0.33

Notes: All nominal values are in New Zealand dollars. Homeowners include mortgagors and outright owners. Throughout the study, the number of reported observations is rounded up or down randomly to a multiple of three in compliance with Statistics New Zealand rules. Owners with multiple properties are dropped from the analysis.

Table 4.2 breaks down homeowners into mortgagors and outright owners. Owners with mortgages tend to be younger than outright owners; the latter tend to be closer to the age of retirement. Mortgagors also have more people in the household (their children), earn higher income, and have higher non-housing spending than outright owners. In per capita terms, however, outright owners still have higher income and expenditure. The two types of owners have similar (gross) housing wealth, but their debt levels are of course markedly different. Mortgage debt is the largest component of household debt, and mortgagors have debt levels that are typically 3 times higher than those of outright owners. As a result, their leverage ratios are also much higher.

TABLE 4. 2: DESCRIPTIVE STATISTICS BY TENURE TYPE

	Mortgagors	Outright homeowners
Sample size	1,670	2,974
Age of household head	44.4	63.3
Number of persons	2.9	2.0
Non-housing expenditure	45,847	39,447
Disposable income	77,129	59,388
Housing wealth	419,614	418,289
Total debt	144,350	44,759
DTI	2.1	0.6
LVR	0.34	0.11

Notes: All nominal values are in New Zealand dollars.

Table 4.3 presents the descriptive statistics for households in different age cohorts. From this table, we observe that most households fall in the oldest age cohort. Households with a head older than 60-year-old are more than three times as common as households with a head younger than 40. This age composition potentially impacts on our quantitative results. Older households typically have lower income, consume less, and live in lower-value housing compared to households in young and prime cohorts. Most importantly for this study, the LVRs and DTIs tend to decline with age as older households tend to have paid down most of their debts.

TABLE 4. 3: DESCRIPTIVE STATISTICS BY AGE GROUPS

	Young (20–40)	Prime (40–60)	Old (60–80)
Sample size	610	1,702	2,183
Number of persons	3.14	2.68	1.62
Non-housing expenditure	45,968	50,601	31,477
Disposable income	79,618	80,857	44,937
Housing wealth	460,415	493,081	441,887
Total debt	174,132	128,953	51,070
DTI	2.46	1.82	1.3
LVR	0.43	0.29	0.16

Notes: All nominal values are in New Zealand dollars.

4.3 Empirical approach

Because our data are a series of repeated cross-sectional waves, we set up the following baseline regression equation:

$$logC_i = \beta_0 + \beta_1 log\ HP_i + \beta_2 log\ Y_i + \sum_{k=1}^{K} \beta_k\ Z_i^k + \mu_i \quad [4.1]$$

Where C_i is non-housing expenditure, HP_i is housing wealth, and Y_i is disposable income. All regressions in this study include an N×K matrix of N observations and K control variables (i =1,…, N; k =1,…, K). The control variables are age, age squared, education dummies, number of people employed, household composition dummies (single-person, couple, couple with children), an ethnicity dummy, territorial authority dummies, and wave dummies.

If $log\ HP_i$ is truly exogenous, we can interpret the estimated coefficient β_1 as the MPC out of housing wealth. However, endogeneity issues are always a challenge to interpreting empirical results. The major concern is that confounding factors could be driving both the consumption expenditure and housing wealth of the household. With our cross-sectional data, we cannot control for all possible household characteristics. As a result, estimating β_1 directly using HES house prices would lead to biased results and the direction of the bias would depend on the correlation between the unobserved confounding factors and the variables in the regression. For this reason, we use house prices at the area unit (AU) level to identify the effect of house prices on the consumption of households in the HES.[25] We regard the cross-AU variations in average house prices as being exogenous relative to the characteristics of individual households. In particular, we use average house sale prices at the AU level as an instrument for household-level

---

[25] AUs are roughly comparable to zip-codes in the US. For more information about geographic definitions in New Zealand refer to section 4.2.

house prices from the HES. We derive the mean sales price of three-bedroom residential houses for each AU and each year between 2006 and 2016 using REINZ data which cover 95% of actual housing sales in New Zealand. Armstrong et al. (2017) match REINZ and HES house price data on household level and report a correlation of 62% for the period (1992-2015). We then use the address information and interview years in the HES to match each household to the corresponding mean house sale price in the AU. This, to a large extent, resolves the endogeneity problem due to unobservable household characteristics driving the empirical results. For example, in an AU, residences are typically mixed from different backgrounds, jobs and demographic characters. As a result, an individual household's preference or high-income expectation would be unlikely to be correlated with the average home selling prices in an AU, unless it was driven by region-specific economic factors, which should be captured by the regional dummies in the regression equation. In addition, the use of an IV approach also helps in overcoming the measurement error in the HES house price data, which are based on the capital values estimated for city councils periodically.

Regressing the log of HES house prices on the log of local average house prices yields an intercept coefficient of 0.78. The coefficient is statistically significant at the 1% level and the regression has a goodness of fit of 22%. We use this first stage regression to obtain the fitted values which represent the exogenous component of the variation in individual house prices.

4.4 Results and discussions

A typical lifecycle model postulates that any form of household consumption responds to changes in its permanent income, which is the annuity value of expected lifetime resources. These resources can be in the form of physical wealth (current housing and financial wealth) or human wealth (current labour income plus the present discounted value of expected future labour income). In such models, changes in wealth affect consumption through the 'wealth channel': any increase in

wealth increases consumption through its impact on expected lifetime resources. When the value of these resources increases, the household can shift its consumption schedule upward without violating its budget constraint. In the standard lifecycle model, the household increases spending in every remaining period of its lifetime by a constant equal to the increase in permanent income, regardless of which component of lifetime resources actually increased (De Veirman and Dunstan 2008).

There are alternative explanations for the impact of housing wealth on consumption. House prices may affect consumption by relaxing or tightening borrowing constraints as a house can be used as a loan collateral. For borrowing-constrained homeowners, an increase in house prices relaxes borrowing constraints, even if there is no wealth effect associated with the house price increase. In other words, an increase in house prices may lead to an increase in consumption not because of a wealth effect, but because it allows borrowing constrained homeowners to smooth consumption over the life cycle (Ortalo-Magné and Rady 2001). The impact of housing wealth on consumption, from this perspective is known as the 'collateral channel'.

4.4.1   Marginal propensity to consume out of housing wealth

Table 4.4 shows the estimation results for the baseline regression for three specifications. In specification i, housing wealth is measured at a household-level and it is sourced from the HES as described in Section 4.2. The estimated coefficient β_1 is 0.09 and it is statistically significant. In column ii, we report the estimate based on average house prices at the AU geographic level. In this case, AU house prices are used as a proxy for exogenous variations in individual housing wealth. Specification iii uses the instrumental variable approach described in the previous section.

TABLE 4. 4: BASELINE REGRESSION RESULTS USING DIFFERENT MEASURES OF HOUSING WEALTH

Dependent variable	Log non-housing expenditures		
	i	ii	iii
Log housing wealth (HES)	0.09*** (0.03)	-	-
Log housing wealth (AU average)	-	0.17*** (0.02)	-
Log housing wealth (IV)	-	-	0.22*** (0.03)
Log income	0.40*** (0.02)	0.40*** (0.02)	0.40*** (0.02)
Other controls	✓	✓	✓
Observations	4,644	4,644	4,644
Adjusted R^2	0.57	0.57	0.57

Note: The control variables are listed in section 4.3. The standard errors are robust and are reported in parentheses. *** indicates significance at 1%.

The estimated MPC using AU prices is almost twice as large as the estimate obtained from HES housing wealth. Wong (2017) who estimated the impact of housing wealth on household consumption using aggregate data for the period 1982-2016 showed a comparable elasticity of 0.22. As we discussed before, the regression based on HES housing value might suffer from multiple problems, which might lead to bias in different directions. In particular, the endogeneity due to unobservable confounding factors might cause upward bias if the unobservable household characteristics are positively correlated with house prices. For example, a higher income expectation could be driving both consumption and the housing value of a household. Similarly, if a household is more impatient, they will also spend more and buy an expensive house. On the other hand, measurement errors in HES house prices could lead to attenuation bias. As the estimated coefficients increase from regressions i to ii and iii, this suggests that the measurement error in the HES house prices has a stronger impact on the results than the endogeneity issue discussed above. Using average local house prices as a proxy variable in specification ii yields a smaller MPC as compared to the result reported in

specification iii. This is largely because the local average house price variable cannot fully explain the changes in the HES house prices unless used as an instrument. We rely on the IV estimation for the rest of the paper.

The estimated elasticity of consumption spending to housing wealth is about 0.22. This means a 1% increase in housing wealth is associated with a 0.22% increase in consumption expenditure. In dollar terms, the average MPC out of a one-dollar increase in housing wealth is around 2.2 cents.[26]

In Table 4.5, we separate non-housing expenditure into durable and nondurable spending. Columns iii in Table 4.4 and i in Table 4.5 are identical. We present the results again in Table 4.5 to give a sense of how MPC differs from durables, to non-durables, to total expenditures. In line with MRS, durable expenditures respond more strongly to changes in income and housing wealth compared to non-durable spending.

---

[26] As in Table 4.1, the average annual non-housing household expenditure (C) is $41,841 NZ dollars for homeowners. Using our house price sales data, the average house value (HP) at the AU level is $418,784 NZ dollars. The elasticity is $\Delta C/\Delta HP*HP/C = 0.22$, implying $\Delta C = 0.22*C/HP*\Delta HP$. Setting $\Delta HP=\$1$, $0.22*42,690/418,784 = \$0.022$ dollar, i.e.2.2 cents.

TABLE 4. 5: BASELINE REGRESSION RESULTS USING DIFFERENT DEFINITIONS OF NON-HOUSING EXPENDITURES

Dependent	Log non-housing Expenditures	Log non-durables	Log durables
	i	ii	iii
Log housing wealth (IV)	0.22***	0.17***	0.26***
	(0.03)	(0.03)	(0.1)
Log income	0.40***	0.37***	0.60***
	(0.02)	(0.02)	(0.05)
Other controls	✓	✓	✓
Observations	4,644	4,644	3,792
Adjusted R^2	0.57	0.56	0.14

Note: The control variables are listed in section 4.3. For the definition of durables and non-durables refer to Appendix C. The number of observations is not identical as we could not perfectly match the addresses reported in HES and REINZ. The last column of durables has fewer observations as not all households reported durable non-housing expenditures within the two-week period of the survey. When restricting the samples in regressions i and ii to the same sample of regression iii, the results remain unchanged to the second decimal place. The standard errors are robust and are reported in parentheses. *** indicates significance at 1%.

4.4.2 The effect of leverage

Table 4.6 summarises the regressions on homeowner spending, controlling for the effect of leverage using either LVRs or DTIs. For both measures, we report the results of three specifications. In the first specification, column i, we only include the leverage ratio as a new independent control variable. The coefficient is estimated to be negative and highly significant. It appears that, relative to outright owners, an increase in LVR by 1 percentage point reduces the consumption spending of borrowers by 0.25%. This result confirms the debt-overhang channel, highlighted by Dynan (2012): household debt is an independent driver of consumption spending.

TABLE 4. 6: THE ROLE OF LEVERAGE IN DETERMINING MPC OUT OF HOUSING WEALTH FOR HOMEOWNERS

Dependent variable	Log non-housing expenditures					
Leverage ratio	LVR			DTI		
	i	ii	iii	i	ii	iii
Log housing wealth (IV)	0.20***	0.21***	0.20***	0.22***	0.21***	0.22***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Leverage ratio	-0.25***	1.09	-	-0.03***	-0.15	-
	(0.04)	(1.18)		(0.01)	(0.21)	
Leverage $\times$ log housing wealth	-	-0.10	-0.02***	-	0.01	-0.003***
		(0.09)	(0.003)		(0.02)	(0.0005)
Log income	0.40***	0.40***	0.40***	0.39***	0.39***	0.39***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Other controls	✓	✓	✓	✓	✓	✓
Observations	4,644	4,644	4,644	4,644	4,644	4,644
Adjusted R^2	0.57	0.57	0.57	0.57	0.57	0.57

Note: The control variables are listed in section 4.3. The standard errors are robust and are reported in parentheses. *** indicates significance at 1%.

Debt overhang channel is likely to be more significant in New Zealand as compared to some other countries due to the institutional setup governing mortgages. In the event of a decline in income or an increase in interest rate, households with lower leverage have higher flexibility to restructure or refinance their mortgage. Households with lower leverage are also more likely to maintain positive equity in the event of a decline in house prices. However, the full recourse nature of mortgage lending in New Zealand means that the ability to continue servicing the loan is a more critical determinant of default for home occupiers.[27]

In specification ii, we allow the leverage ratio to interact with house prices. This specification provides another channel through which the leverage ratio can affect consumption, namely the housing wealth channel. As discussed in MRS, collateral constraints are part of the mechanism that translate movements in housing wealth into movements in consumption. During a severe housing market downturn, as in 2008/9 in the US, highly leveraged households face a binding borrowing constraint and their consumption falls more sharply in response to declines in housing wealth. Interestingly, in our regression ii, the estimated coefficient on the interaction term is negative, but the coefficient on the leverage ratio is positive. Both estimates are statistically significant.

---

[27] In some countries, mortgages are backed by the house itself only. These are called non-recourse mortgages, as banks cannot pursue a borrower directly for the debt. If a house is purchased for $100 with 90% LVR and the house price declines by more than 10%, the household might choose to strategically default on their mortgage and turn over the key to the bank. This is known as strategic default since the household chooses to default not because they cannot service the mortgage anymore but to minimise their wealth losses. The risk of strategic default declines over time as the household pays back the debt gradually, i.e. expands its equity buffer. Most mortgages in New Zealand are full recourse contracts where the creditor has the right to pursue the borrower for the debt, regardless of what happens to the asset tendered as a security (Harris and Meir 2015). The risk of strategic default is, therefore, less likely in New Zealand.

A closer look at the correlation between these two independent variables shows that the leverage ratio is highly correlated with the interaction term between leverage and housing wealth in the regression.[28] This collinearity problem causes the associated coefficients to be imprecisely estimated. This motivates us to run the third specification with only the interacted leverage ratio and housing wealth covariate. Column iii shows that once the collinearity issue is removed, the estimated coefficient becomes negative and statistically significant. Using a DTI as a measure of household indebtedness delivers similar results as seen in the right half of the table.

Taking all the regressions together, we conclude that household indebtedness affects consumption spending through two channels. First, the debt-overhang channel as highlighted by Dynan (2012). Second, household indebtedness, especially via mortgage debt, acts as a drag on consumption spending, not only through the level effect, but also by influencing the slope of the housing wealth effect. Because of the collinearity issue, our data cannot separately disentangle these two channels. However, our empirical results also confirm the existence of the house wealth channel through which household indebtedness affect consumption spending.

A particularly interesting aspect of our results is that the estimated interaction coefficient is negative, which suggests that households with high leverage are less sensitive to exogenous house price variations. This result contrasts strikingly with MRS, who show that the household spending response to housing wealth shocks is stronger in regions with highly indebted households. They argue that the finding

---

[28] See the cross-correlations table in Appendix D.

confirms the collateral constraint channel, which is studied by Iacoviello (2005) and Kiyotaki and Moore (1997).

In the light of our empirical result, we argue that the relationship between household indebtedness and consumption spending is more complex than implied by MRS. One possible explanation for the difference in the impact of household indebtedness is that US studies mainly focus on the period after the Great Recession, when the US housing market suffered from a severe downturn, while in the most of the sample period for our New Zealand dataset house prices were growing (See Figure 4.1). The effect of the leverage ratio might be asymmetric in terms of how the consumption spending of borrowers responds to increases or decreases in housing wealth. Therefore, more theoretical modelling of the interaction between household debt and consumption is desirable in future research.

4.4.3 Robustness tests

According to descriptive statistics in Tables 4.2 and 4.3, household leverage is correlated with tenure types and age. For example, mortgagors tend to be significantly more leveraged relative to outright owners, because the main type of household debt is mortgage debt. Table 4.3 shows that older households tend to be less leveraged as they have paid most of their mortgages down. To confirm the robustness of the findings reported in section 4.4.2, in this section, we use age and tenure type as a proxy for leverage. This exercise overcomes the potential measurement errors in the household leverage data.

In Table 4.7, specification i allows the age of the household head to interact with housing wealth. We find that age has a small but significant impact on the marginal propensity to consume out of housing wealth, suggesting that the expenditure of older households is slightly more responsive to changes in housing wealth. Given that older households have lower leverage, this result is consistent with our finding in the previous section. Campbell and Cocco (2007) also show similar results for

the UK. However, the results appear to be at odds with Attanasio et al. (2009). A major difference between the approach taken in this paper and Attanasio et. at. (2009) is that we control for the level of income in our regressions.

In column ii, we interact a mortgagor dummy with housing wealth. The dummy takes one for mortgagors and zero for outright owners. The result in column i in this table as well as those in Table 4.6, implies that mortgagors are less sensitive to changes in housing wealth relative to outright owners.

TABLE 4. 7: ESTIMATED MPC BY AGE AND TENURE TYPE FOR HOMEOWNERS

Dependent variable	Log non-housing expenditures			
	Benchmark	i	ii	iii
Log housing wealth (IV)	0.22***	0.14***	0.20***	0.22***
	(0.03)	(0.03)	(0.03)	(0.03)
Age×log housing wealth	-	0.001***	-	-
		(0.0002)		
Mortgagor dummy × log housing wealth	-	-	-0.01***	-
			(0.002)	
DTI× log housing wealth	-	-	-	-0.005***
				(0.01)
Age×DTI× log housing wealth	-	-	-	0.00006**
				(0.00)
Log income	0.40***	0.40***	0.39***	0.39***
	(0.02)	(0.02)	(0.02)	(0.02)
Other controls	✓	✓	✓	✓
Observations	4,644	4,644	4,644	4,644
Adjusted R^2	0.57	0.57	0.57	0.57

Notes: The control variables are listed in section 4.3. Standard errors are reported in parentheses. *** indicates significance at 1%.

In column iii, we re-estimate the regression equation as in Table 4.6, but add a three-way interaction term between age, DTI and housing wealth. The idea of this specification is to check if our estimate is driven by a life cycle pattern. Leverage is systematically high for younger compared to older households. If the housing MPC is also systematically correlated to the life cycle, our interaction term in Table 4.6 will pick up this life cycle effect as well. The empirical result shows a small but

positive estimate, suggesting that the spending of older households is more sensitive to changes in housing wealth, when they have higher leverage. More importantly, after controlling for this life cycle effect, the interaction between leverage and housing wealth is still negative with a similar magnitude.

For completeness, Appendix D reports the regression results when using AU average house prices as a proxy, instead of an instrument. All empirical findings do not change materially.

4.5 Conclusion

This paper conducts a microeconometric analysis exploring how changes in housing wealth impact household consumption in New Zealand once one controls for the role of leverage and other household characteristics. The paper highlights the impact of data choices regarding the variable of housing and how it affects the econometric estimation. We find that household indebtedness plays a significant role in determining the wealth effects of housing. In contrast with the literature, we find that highly indebted households spend less out of increases in housing wealth than do less indebted households. The microeconometric evidence helps to explain the empirical finding of a weakening MPC out of housing wealth over time, as documented by Wong (2017) using New Zealand aggregate data. Our results point to rising household leverage as a significant driver behind the change in the MPC out of housing wealth in New Zealand since the 2000s.

Although survey-based data are richer in information, one caveat is that households can sometimes mis-report their activity, resulting in data inaccuracies. Another caveat of the approach followed in this study is not controlling for whether households respond to expected or unexpected increases in housing wealth. The repeated cross-sectional nature of the data prevented us from achieving this goal. Permanent Income Hypothesis postulates that consumption should respond only to unpredictable changes in income or wealth.

APPENDIX C

	Type	Description	NZHEC code
Non-housing expenditures	Durables	Furniture, furnishings and floor coverings	5.1
		Household appliances	5.3
		Tools and equipment for house and garden	5.5
		Purchase of vehicles	7.1
		Audio-visual and computing equipment	9.1
		Major recreational and cultural equipment	9.2
		Jewellery and watches	11.3.01
	Non-durables	Fruit and vegetables	01.1
		Meat, poultry and fish	01.2
		Grocery food	01.3
		Non-alcoholic beverages	01.4
		Restaurant meals and ready-to-eat food	01.5
		Alcoholic beverages	02.1
		Cigarettes and tobacco	02.2
		Illicit drugs	02.3
		Clothing	03.1
		Footwear	03.2
		Actual rentals for housing	04.1
		Home ownership	04.2
		Household energy	04.5
		Household textiles	05.2
		Glassware, tableware and household utensils	05.4
		Other household supplies and services	05.6
		Medical products, appliances and equipment	06.1
		Out-patient services	06.2
		Hospital services	06.3
		Private transport supplies and services	07.2
		Passenger transport services	07.3
		Postal services	08.1

		Telecommunication equipment	08.2
		Telecommunication services	08.3
		Other recreational equipment and supplies	09.3
		Recreational and cultural services	09.4
		Newspapers, books and stationery	09.5
		Accommodation services	09.6
		Package holidays	09.7
		Miscellaneous domestic holiday costs	09.8
		Early childhood education	10.1
		Primary, intermediate and secondary education	10.2
		Tertiary and other post school education	10.3
		Other educational fees	10.4
		Personal care	11.1
		Prostitution	11.2
		Personal effects not elsewhere specified	11.3
		Insurance	11.4
		Expenditure incurred whilst overseas	13.5
		Sales of clothing and footwear	14.1
		Sales and trade-ins of property and materials for property improvement and maintenance	14.2
		Sales and trade-ins of household contents	14.3
		Sales, trade-ins and refunds for health (excluding insurance claims)	14.4
		Sales and trade-ins of vehicles, vehicle parts and accessories	14.5
		Sales and trade-ins for communication	14.6
		Sales, trade-ins and refunds of equipment for recreation and culture	14.7
		Refunds for education	14.8
		Sales, trade-ins and refunds of miscellaneous goods, cash receipts from insurance claims	14.9

Note: The complete New Zealand Household Economic Survey Classification (NZHEC) can be found at:
http://www.stats.govt.nz/browse_for_stats/people_and_communities/Households/household-economic-survey-classifications.aspx

APPENDIX D

TABLE D1: CROSS CORRELATIONS TABLE

	Total expenditure	Durable	Non-durable	Disposable income	House wealth	DTI	(DTI)X (income)	(DTI)X (housing)
Total expenditure	1							
- Durable	0.53	1						
- Non-durable	0.95	0.30	1					
Disposable income	0.64	0.30	0.63	1				
Housing wealth	0.43	0.13	0.44	0.50	1			
DTI	-0.12	-0.11	-0.10	-0.15	0.25	1		
(DTI)X(income)	-0.09	-0.09	-0.07	0.09	0.28	0.99	1	
(DTI)X(housing)	-0.10	-0.10	-0.08	-0.12	0.29	0.99	0.99	1

TABLE D2: THE ROLE OF LEVERAGE IN DETERMINING MPC OUT OF HOUSING WEALTH FOR HOMEOWNERS

Dependent variable	Log non-housing expenditures					
Leverage ratio	LVR			DTI		
	I	ii	iii	i	ii	iii
Log housing wealth	0.15***	0.16***	0.16***	0.17***	0.17***	0.15***
(AU average)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)
Leverage ratio	-0.25***	0.77	-	-0.03***	-0.13	-
	(0.04)	(0.9)		(0.006)	(0.16)	
(Leverage) × (log housing wealth)	-	-0.08	-0.02***	-	0.00	-0.003***
		(0.07)	(0.003)		(0.02)	(0.0005)
Log income	0.40***	0.40***	0.40***	0.39***	0.39***	0.39***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Other controls	✓	✓	✓	✓	✓	✓
Observations	4,644	4,644	4,644	4,644	4,644	4,644
Adjusted R^2	0.57	0.57	0.57	0.57	0.57	0.57

Note: These regressions are based on homeowners only, which includes both outright owners and mortgagors. The control variables are listed in section 4.3. The standard errors are robust and are reported in parentheses.

BIBLIOGRAPHY

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

Armstrong, J., Dunstan, A., & Irrcher, T. (2017). *Evaluating alternative monthly house price measures for New Zealand* (No. AN2017/02). Reserve Bank of New Zealand.

Attanasio, O. P., Blow, L., Hamilton, R., & Leicester, A. (2009). Booms and busts: Consumption, house prices and expectations. *Economica*, *76*(301), 20-50.

Bahmani-Oskooee, M. (1996). Decline of the Iranian Rial during the post-revolutionary period: A productivity approach. *The Journal of Developing Areas, 30*(4), 477-492.

Bahmani-Oskooee, M., & Ratha, A. (2004). The J-curve dynamics of US bilateral trade. *Journal of Economics and Finance, 28*(1), 32-38.

Bahmani-Oskooee, M., & Wang, Y. (2008). The J-curve: evidence from commodity trade between US and China. *Applied Economics, 40*(21), 2735-2747.

Bahmani-Oskooee, M., & Hegerty, S. W. (2010). The J-and S-curves: a survey of the recent literature. *Journal of Economic Studies*, *37*(6), 580-596.

Bahmani-Oskooee, M., Hegerty, S., & Harvey, H. (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*(8), 1183-1213.

Baker, S R (forthcoming) Debt and the response to household income shocks: Validation and application of linked financial account data, *Journal of Political Economy*.

Bostic, R, S Gabriel and G Painter (2009) Housing wealth, financial wealth, and consumption: New evidence from micro data, *Regional Science and Urban Economics*, 39(1), 79-89.

Brown, W. B., & Hogendorn, J. S. (2000). *International Economics: In the Age of Globalization*: University of Toronto Press.

Campbell, J Y, and J F Cocco (2007) How do house prices affect consumption? Evidence from micro data, *Journal of Monetary Economics*, 54(3), 591-621.

Carroll, C D (2001) A theory of the consumption function, with and without liquidity constraints, *The Journal of Economic Perspectives*, 15(3), 23-45.

Carroll, C D and M S Kimball (2006) Precautionary saving and precautionary wealth, *Center for Financial Studies Working Paper*, 2006/02.

Carroll, C D, M Otsuka and J Slacalek (2011) How large are housing and financial wealth effects? A new approach, *Journal of Money, Credit and Banking*, 43(1), 55-79.

Case, K E, J M Quigley and R J Shiller (2003) Home-buyers, housing and the macroeconomy, *Berkeley Program on Housing and Urban Policy*, W04-004.

Clarke, A., & Kulkarni, K. G. (2010). Testing the Application of Heckscher-Ohlin Theorem to Contemporary Trade Between Malaysia and Singapore. *Journal of Emerging Knowledge on Emerging Markets, 1*(1), 10.

De Saint Vaulry, A. (2008). *Base de données CHELEM-commerce international du CEPII*: Centre d'études prospectives et d'informations internationales.

De Veirman, E., & Dunstan, A. (2008). *How do housing wealth, financial wealth and consumption interact? Evidence from New Zealand* (No. DP2008/05). Reserve Bank of New Zealand.

Dynan, K (2012) Is a household debt overhang holding back consumption?, *Brookings Papers on Economic Activity*, 2012(1), 299-362.

Engelhardt, G V (1996) House prices and home owner saving behavior, *Regional Science and Urban Economics*, 26(3), 313-336.

Feenstra, R. C., Lipsey, R. E., Deng, H., Ma, A. C., & Mo, H. (2005). World trade flows: 1962-2000 (Vol. Working Paper 11040): National Bureau of Economic Research.

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.

Fisher, I (1930) *Theory of Interest,* Macmillan, New York.

Fisman, R., & Wei, S. J. (2004). Tax rates and tax evasion: evidence from 'missing imports' in China. *Journal of political Economy*, *112*(2), 471-496.

Fisman, R., & Wei, S. J. (2009). The smuggling of art, and the art of smuggling: Uncovering the illicit trade in cultural property and antiques. *American Economic Journal: Applied Economics*, *1*(3), 82-96.

Fortanier, F. (2016). Towards merchandise trade statistics without asymmetries. *OECD, The Statistics Newsletter* (Issue No. 64, March 2016).

Friedman, M. (1957) *A Theory of the Consumption Function*. Princeton University Press, Princeton.

Gaulier, G., & Zignago, S. (2010). BACI: international trade database at the product-level (the 1994-2007 version). *Centre d'Etudes Prospectives et d'Informations Internationales*.

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

Gopinath, G., Itskhoki, O., & Rigobon, R. (2010). Currency choice and exchange rate pass-through. *American Economic Review*, *100*(1), 304-36.

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

Guo, D. (2010). Mirror Statistics of International Trade in Manufacturing Goods: The Case of China. Research and statistics branch: working paper 19/2009. *Vienna, United Nations Industrial Development Organization*.

Harris, R., & Meir, A. (2015). Recourse structure of mortgages: a comparison between the US and Europe. *CESifo DICE Report*, *13*(4), 15-22.

Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The quarterly journal of economics*, *123*(2), 441-487.

Hviid, S J, A Kuchler & Nationalbank, D. (2017) Consumption and savings in a low interest-rate environment, *Danmarks National Bank Working Papers*, 116.

Iacoviello, M (2005) House prices, borrowing constraints, and monetary policy in the business cycle, *The American Economic Review*, 95(3), 739-764.

Juster, F. T., J P Lupton J P Smith, and F Stafford (2006) The decline in household saving and the wealth effect, *The Review of Economics and Statistics*, 88(1), 20-27.

Kaplan, G (2012). Inequality and the life cycle, *Quantitative Economics*, 3(3), 471-525.

Kaplan, G and G L Violante (2014) A model of the consumption response to fiscal stimulus payments, *Econometrica*, 82(4), 1199-1239.

Kaplan, G, K Mitman and G L Violante (2016) Non-durable consumption and housing net worth in the great recession: Evidence from easily accessible data, *National Bureau of Economic Research Working Paper*, 22232.

Kiyotaki, N and Moore, J (1997) Credit Cycles, *Journal of Political Economy*, 105(2), 211-248.

La Cava, G, H Hughson and G Kaplan (2016) The household cash flow channel of monetary policy, *Reserve Bank of Australia Research Discussion Paper*, 2016-12.

Lehnert, A (2004) Housing, consumption, and credit constraints, Board of Governors of the Federal Reserve System Finance and Economics Discussion Series, 2004-63.

Levin, L (1998) Are assets fungible? Testing the behavioral theory of life-cycle savings, *Journal of Economic Behavior & Organization*, 36(1), 59-83.

Makhoul, B., & Otterstrom, S. M. (1998). Exploring the accuracy of international trade statistics. *Applied Economics*, *30*(12), 1603-1616.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, *71*(6), 1695-1725.

Mian, A., K Rao and A Sufi (2013) Household balance sheets, consumption, and the economic slump, *The Quarterly Journal of Economics*, 128(4), 1687-1726.

Miao, G. and Fortanier, F. (2017) 'Estimating Transport and Insurance Costs of International Trade', *OECD Statistics Working Paper 2017/04*; http://dx.doi.org/10.1787/8267bb0f-en

Ortalo-Magné, F., & Rady, S. (2002). Tenure choice and the riskiness of non-housing consumption. *Journal of Housing Economics*, *11*(3), 266-279.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics, 16*(3), 289-326.

Shaar, K., & Khaled, M. (2017). Why you should use high frequency data to test the impact of exchange rate on trade. *Applied Economics Letters*, 1-4.

Simola, H. (2012). Russian import statistics in the mirror of world exports. BOFIT. *Institute for Economies in Transition*, (2).

Skinner, J (1989) Housing wealth and aggregate saving, *Regional Science and Urban Economics*, 19(2), 305-324.

Skinner, J S (1994) Housing and saving in the United States, in Y Noguchi and J M Poterba, eds, *Housing Markets in the United States and Japan,* 191-214. University of Chicago Press.

Smith, M. (2010). *Evaluating household expenditures and their relationship with house prices at the microeconomic level* (No. DP2010/01). Reserve Bank of New Zealand.

Vincent, J. R. (2004). *Detecting illegal trade practices by analyzing discrepancies in forest products trade statistics: An Application to Europe, with a focus on Romania*. The World Bank.

Wang, Z., Gehlhar, M., & Yao, S. (2010). A globally consistent framework for reliability-based trade statistics reconciliation in the presence of an entrepôt. *China Economic Review*, *21*(1), 161-189.

West, L. A. (1995). *Reconciling China's trade statistics*. International Programs Center, Population Division, US Bureau of the Census.

Wong, M (2017) Revisiting the wealth effect on consumption in New Zealand, *Reserve Bank of New Zealand Analytical Note*, AN2017/03.

Yao, J, A Fagereng and G Natvik. (2015) Housing, debt and the marginal propensity to consume, *Norges Bank Research Paper*, 1-38.

Yeats, A. J. (1990). On the accuracy of economic observations: Do sub-Saharan trade statistics mean anything? *The World Bank Economic Review, 4*(2), 135-156.

Zhou, S. (2001). The power of cointegration tests versus data frequency and time spans. *Southern Economic Journal*, 906-921.