Policy Research Working Paper 10112

Trade Fraud and Non-Tariff Measures

Hiau Looi Kee
Alessandro Nicita

WORLD BANK GROUP
Development Economics
Development Research Group
June 2022

Policy Research Working Paper 10112

Abstract

Similar to tariffs, non-tariff measures may induce trade fraud when they are restrictive. This paper examines whether discrepancies observed in the official trade statistics of importing and exporting countries are partly due to trade fraud from evading border non-tariff measures. To capture the restrictiveness of non-tariff measures, the paper estimates the ad valorem equivalent with importer-exporter-product variations. It presents a theoretical model and empirical evidence showing that discrepancies increase with ad valorem equivalents, consistent with the trade fraud due to traders intentionally mis-declaring countries of origin or misclassifying products in order to evade border non-tariff measures. The results are driven by homogeneous products and the trade between developed and developing countries.

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at hlkee@worldbank.org.

Trade Fraud and Non-Tariff Measures*

Hiau Looi Kee† Alessandro Nicita‡

World Bank United Nations

Key Words: trade discrepancies, non-tariff measures, ad valorem equivalent of NTMs, tariff evasion, trade fraud

JEL Classification Numbers: F10, F14

*We thank Pol Antras, Erhan Artuc, Chad Bown, Dave Donaldson, Davin Chor, Ana Fernandes, Michael Ferrantino, Penny Goldberg, Russell Hillberry, Aaditya Mattoo, Fernando Perro, Bob Rijkers, and the seminar participants of the IMF/WB/WTO Joint Trade Workshop, National University of Singapore, SAIS and George Washington University for their comments. Research for this paper has in part been supported by the World Bank's Multidonor Trust Fund for Trade and Development. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. The results and opinions presented in this paper do not represent the views of our institutions, the Executive Directors, or the countries they represent.

†Development Research Group, The World Bank, Washington, DC 20433, USA; Tel.: (202) 473-4155; Fax: (202)522-1159; e-mail: hlkee@worldbank.org.

‡United Nations Conference on Trade and Development. Email: Alessandro.Nicita@un.org.

"...buyers do not allow any 'Made in Bangladesh' label on the [garment] product ...they [requested] label these very products as 'Made in India' ... "

– Prabir De, "Non-Tariff Measure Study: Bangladesh, India and Nepal," 2016

"Five individuals and two domestic honey-processing companies have been charged with federal crimes... The charges assert that the Chinese-origin honey was misdeclared as other commodities upon importation into the United States and transshipped through other countries to evade anti-dumping duties."

– World Customs Organization, *Illicit Trade Report,* 2012

"Professional Fraud Facilitators... provide a complete fraud package to EU importers and Chinese exporters, with delivery of goods to EU port, change of container in transit country and documented new origin for goods (new B/L, invoices, origin certificates)."

- European Anti-Fraud Office, *Customs Origin Fraud and Professional Fraud Facilitators,* 2014

1 Introduction

In international trade, importing and exporting countries often record starkly different customs statistics for the same shipments, mostly with imports significantly larger than exports. For example, in 2018, the US recorded US$563 billion in overall imports from China, while China recorded US$480 billion in exports to the US.[1] The result is a trade discrepancy of nearly US$83 billion, or 15 percent. Not only is the presence of such trade discrepancies puzzling, it also distorts the true picture and may lead to public misperception and ill-informed policy decisions. At a detailed product level, the discrepancies in customs statistics are even

[1] These figures are from national statistics of the US and China.

more staggering and widespread across countries and products, both in terms of values and quantities. Figure 1 presents the distribution of trade discrepancies for a sample of more than 40 importing countries.[2] At the sample mean, the trade discrepancy is about 85 percent – suggesting that most import records are much larger than the export records, for both quantity and value.

What could cause such discrepancies in the official trade statistics? There are several contributing factors. The discrepancies could be at least in part due to the different valuation methods applied by customs agencies, which normally include freight and insurance costs in import values but not in export values.[3] Another contributing factor is that the transshipment and re-export of products going through third countries cause a mismatch in the country of origin.[4] The discrepancies could also result from trade fraud in the form of undervaluation in imports in order to evade import tariffs,[5] or the undervaluation in exports in order to evade export taxes or VAT.[6] The analysis of trade discrepancies has also been increasingly used to detect tariff evasion, especially by the customs authorities of developing countries.[7] However, even taken together, these contributing factors cannot fully explain the large discrepancies in the data, which tend to vary across products and trading partners, particularly with the recorded import quantity larger than the export quantity.

The objective of this paper is to assess whether part of the large discrepancies in the official statistics is actually trade fraud due to the avoidance of costs associated with the non-tariff measures (NTMs) of the importing countries.[8] Such trade fraud may not only

[2]In this paper, trade discrepancies are defined according to Equation (8) , which measures the percentage difference between the importer's and exporter's statistics. Figure 1 plots the distributions of the value discrepancies (gapvsh) and quantity discrepancies (gapsh) at the HS 6-digit product level across countries. Please refer to Table 1 for the list of importing countries.

[3]Hummels and Lugovskyy (2006).

[4]Ferrantino and Wang (2008).

[5]Bhagwati (1964), Fisman and Wei, (2004), and Javorcik and Narciso (2008).

[6]Ferrantino, Liu and Wang (2012).

[7]Cantens (2015); Grigoriou, Kalizinje and Raballand (2019); Grigoriou(2019); Chalendard, Fernandes, Raballand and Rijkers (2020).

[8]This paper focuses only on large discrepancies at the HS 6-digit level, in which the difference between

reduce government revenues but may also undermine the trade regulatory framework of the importing country, and ultimately alter the competitive environment by resulting in an unfair advantage gained by the fraudsters over their law-abiding competitors. By avoiding NTMs, trade fraud may jeopardize public health or safety, and weaken environmental protections. The three examples cited above in italics highlight such fraud, whereby traders either misreport country of origin or product codes to evade NTMs. We focus on the NTMs applied at the border, hereafter referred to as Border NTMs.[9] Examples of Border NTMs include customs controls, quota licensing, pre-shipment inspections, and additional fees paid at customs, among many others. Some of these NTMs are bilateral in nature, targeting specific products from specific partner countries. Compliance with these requirements is often financially costly and/or time-consuming,[10] therefore providing incentives for traders to commit trade fraud by misrepresenting the country of origin or misclassifying product codes on their customs documents (Li and Lin, 2022).

To study the role of Border NTMs in causing trade fraud, this paper first provides a simple theoretical model built on Swenson (2001) and Ferrentino, Liu and Wang (2012) to show that the restrictive NTMs cause importers and exporters to under-report the quantity traded on their respective customs declaration forms. However, given that the penalty and detection rates of the importing countries are generally higher than those of the exporting countries, importers will under-report less than exporters hence giving rise to the observed trade discrepancies. The theoretical model also shows that trade discrepancies increase with the ad valorem equivalent (AVE) of NTMs. Furthermore, the model shows that the effect of AVE on trade discrepancies is larger when products are homogeneous and when the exporting countries are developing countries. In addition, the model rationalizes origin and

the import value or volume and the export value or volume is more than 40 percent, in order to side step variations in the valuations and small statistical errors.

[9] For more details see Ederington and Ruta (2016) and UNCTAD (2015).

[10] See UNCTAD, World Bank (2018).

product classification fraud whereby exporters misdeclare the country or misclassify products in their customs declaration forms. Finally, the model shows that the relationship between trade discrepancies and tariffs depends on the effect of tariffs on the detection rate of the importing countries.

The empirical section of the paper heeds the message of Goldberg and Pavcnik (2016): "*Measure* (trade barriers) *before you estimate* (trade impacts)*!*" We first led the NTM data collection efforts of the UNCTAD and the World Bank which allow us to generate an NTM database of 50 importing countries at the detailed 6-digit Harmonized System (HS) level with bilateral variation across 96 trading partners.[11] We then rely on product level quantity-based gravity regressions to estimate the bilateral AVE of NTMs at the HS 6-digit level. Our estimated AVEs are reasonable and in line with previous studies, with a sample mean of 4.7 percent when NTMs are present. Using the estimated AVEs, this paper then assesses whether the presence of Border NTMs contributes to large trade discrepancies as shown in the theoretical model. The results show that exporting countries that face higher AVEs have larger trade discrepancies, consistent with the hypothesis of country origin fraud in order to gain more favorable market access conditions. Similarly, the regression results also demonstrate that the products that have higher AVEs are those that have larger trade discrepancies, consistent with the hypothesis of product classification fraud with the purpose of reducing the costs of compliance with Border NTMs. These results are consistent with the theoretical model and are robust to simultaneity and measurement errors in AVEs, which we address with the use of instrumental variables. Overall, this paper finds that a 10 percentage point increase in the AVEs could lead to a 15 percent increase in trade discrepancies, while a comparable increase in tariffs leads to only an 8 percent increase in trade discrepancies. These findings suggest that a significant part of trade discrepancies is due to trade fraud associated

[11] This is done by working with local consultants in identifying and codifying countries' laws and regulations that may affect international trade. Some of these laws and regulations target specific products or countries.

with NTM avoidance schemes. Robustness checks further reveal that NTM-induced trade fraud is more common for homogeneous products, as well as for trade between developed and developing countries, as predicted by the theoretical model.

The rationale of this paper to investigate trade fraud relates to the existing literature on trade discrepancies and tariff avoidance. In particular, this paper follows on the "missing trade" approach initially proposed by Fisman and Wei (2004) to explain discrepancies in the statistics between Hong Kong SAR, China, and China. This approach was also used by Javorcik and Narciso (2008) whose examination of the exports statistics of Germany find that discrepancies with the statistics of the importing country tend to be greater when tariffs are higher. Similar results are found for India by Mishra, Topalova and Subramanian (2008). In addition, focusing on Chinese exports to the United States, Ferrantino, Liu and Wang (2012) show that Chinese exporters may under-report export values in order to avoid paying China's VAT, thereby helping to explain why reported export values are lower than the corresponding reported import values. Most recently, Demir and Javorcik (2020) also show import duty evasion as a margin through which firms adjust to changes in trade policy in Turkey, while Liu, Sheng and Wang (2021) show that trade data over-reporting is positively correlated with the exchange rate spread in China. This paper adds to the literature by looking at NTMs as another source of costs at the borders, which may result in similar schemes employing misdeclaration and misclassification with the purpose to reduce costs. In addition, unlike previous papers that mostly focused on one country's imports or exports, this paper uses detailed trade data from a wide range of importing and exporting countries, which allow us to detect customs origin fraud and product fraud.

Importantly, this paper also contributes to the literature on assessing the effect of trade policies, particularly non-tariff measures, on trade. In particular, this paper builds on the work of Kee, Nicita and Olarreaga (2009) estimating the AVE of non-tariff measures by providing a novel method to account for possible heterogeneous effects of NTMs across trading partners while employing econometric techniques in the estimation of AVEs that relate to the literature on the estimation of gravity models. Unlike previous papers, the AVE estimates of this paper are highly detailed and disaggregated at the HS 6-digit level and vary by importer and exporter. Such detailed variations could be useful for studies evaluating the impact of specific trade policy changes.

The paper proceeds as follows. Section 2 provides theoretical motivations to define trade discrepancies. Section 3 provides a data description for the paper. Section 4 presents the empirical model to estimate the AVEs of NTMs. Section 5 presents the results on the estimation of AVEs. Section 6 shows the empirical evidence for NTM-induced trade fraud, while robustness checks are presented in Section 7. Section 8 concludes.

2 Theoretical Motivations

This section provides a simple theoretical model built on Ferrantino, Liu and Wang (2012), which followed Swenson (2001). Those two studies emphasize the importance of transfer pricing, export taxes such as VAT and capital controls in affecting the under-reporting behavior of importers and exporters from China and the US. Different from the analysis of those two papers, we focus on a wider range of importing and exporting countries with different levels of development and institutions, where transfer pricing, VAT and capital controls may not be important. Instead, we emphasize the roles of tariffs and NTMs in inducing importers and exporters to under-report the quantity and value of trade differently, which leads to the observed trade discrepancies in the official statistics. Similarly to the previous models, traders weigh the benefits of under-reporting, which are the tariffs saved, against the costs of under-reporting, which are the potential penalties, to choose the optimal level of trade fraud in terms of under-reporting. In the model of this paper, in addition to the potential tariff evaded, which depends on trade value, under-reporting also allows traders to evade restrictive NTMs, such as inspections and testing requirements, which generally depend on the quantity traded. While the NTMs are only imposed by the importing countries, they may affect both the importing and exporting firms. Nevertheless, differences in the penalty rates and detection rates of the importing and exporting countries give rise to the observed trade discrepancies.

The following model focuses on under-reporting in quantity traded. Let index $k = \{X, M\}$ indicate exporting or importing firms. Exporting and importing firms submit respective customs declaration forms for the same shipment, but may declare different quantities, Q^X and Q^M, in order to minimize their trade costs. For the exporting firms, their exporting costs include the importing countries' NTM related transaction costs, such as pre-shipment inspection, testing requirements, conformity certificates and other obligations necessary to comply with the requirements of the importing country. These costs depend on the quantity and are captured parsimoniously by the AVE of NTM, multiplied by the size of the shipment, PQ^X. While these costs provide an incentive for exporting firms to under-report export quantities, exporting firms may also face penalties if they are detected to have under-reported in their export custom forms. The overall penalty cost depends on the penalty rate, α^X, the detection rate, s^X, the squares of the size of deviation, $\left(Q - Q^X\right)/Q$, and the true value of the shipment, PQ.[12] Each exporting firm will therefore choose to declare Q^X to minimize the total exporting costs, C^X:

$$\min_{Q^X} C^X = \underbrace{AVE \cdot PQ^X}_{\text{NTM-induced Cost}} + \underbrace{\frac{\alpha^X s^X}{2} \left(\frac{Q - Q^X}{Q}\right)^2 PQ}_{\text{Penalty Cost}}. \tag{1}$$

[12] Here we follow Ferrantino, Liu and Wang (2012) to model the cost of exporting depending on the squares of the size of deviation.

Under-reporting, such that $Q - Q^X > 0$, implies a lower NTM-induced cost, and also a higher penalty cost. At the equilibrium, the following first-order condition must hold:

$$\begin{aligned}\frac{\partial C^X}{\partial Q^X} &= AVE \cdot P - \alpha^X s^X \left(\frac{Q - Q^X}{Q}\right) P = 0 \implies \\ \left(\frac{Q - Q^X}{Q}\right) &= \frac{AVE}{\alpha^X s^X} > 0 \text{ if } \alpha^X s^X > 0 \text{ or} \qquad (2) \\ \frac{Q^X}{Q} &= 1 - \frac{AVE}{\alpha^X s^X}. \qquad (3)\end{aligned}$$

Thus, if the penalty and the detection rates are positive, it is always optimal for the exporting firm to under-report, such that $Q - Q^X > 0$. According to equation (2), the size of under-reporting increases with the cost of satisfying the importing country's NTMs, and decreases with the penalty and detection rates of the exporting country. If the exporting country does not impose penalties for under-reporting, such that $\alpha^X s^X = 0$, which is generally the case empirically as custom administrations may not have incentives or resources to accurately monitor exports, then

$$\begin{aligned}\frac{\partial C^X}{\partial Q^X} &= AVE \cdot P > 0 \implies \qquad (4) \\ Q^X &= 0.\end{aligned}$$

Equation (4) suggests that, in the absence of accurate controls of export shipments or penalty charges, C^X increases with Q^X. Hence, to minimize C^X, the optimal quantity to declare in the export customs form is 0. This is the case when the exporter commits either origin fraud by mis-declaring the country code by trans-shipping through a third country facing lower border costs associated with NTMs, or product classification fraud by misclassifying the product code as to face lower compliance costs with border NTMs. In both cases, the quantity of the true origin or true product code will be 0.

On the importers side, the per unit NTM-induced costs may include import licensing requirement and inspection, simply represented by the AVE of NTM. In addition, importing firms may also face a tariff, t, on the total value of the shipment. The importing firm will therefore choose to declare Q^M to minimize the total importing costs, C^M:

$$\min_{Q^M} C^M = \underbrace{t \cdot PQ^M}_{\text{Tariff cost}} + \underbrace{AVE \cdot PQ^M}_{\text{NTM-induced cost}} + \underbrace{\alpha^M s^M \left(\frac{Q - Q^M}{Q}\right)^2 PQ}_{\text{Penalty cost}}. \tag{5}$$

The first-order condition implies that

$$\frac{\partial C^M}{\partial Q^M} = t \cdot P + AVE \cdot P - \alpha^M s^M \left(\frac{Q - Q^M}{Q}\right) P = 0 \implies$$

$$\left(\frac{Q - Q^M}{Q}\right) = \frac{t + AVE}{\alpha^M s^M} > 0 \text{ if } \alpha^M s^M > 0 \text{ or} \tag{6}$$

$$\frac{Q^M}{Q} = 1 - \frac{t + AVE}{\alpha^M s^M} \tag{7}$$

Thus, equation (6) shows that the optimal size of under-reporting for the importers increases with the level of tariff, or the cost of complying with $NTMs$, but decreases with the penalty rate or detection rate of the importing country. In other words, the declared quantity Q^M will deviate from its true value Q depending on the level of tariff and the cost of compliance with NTMs, subject to the penalty and detection rates. Given that importing country governments need to collect tariff revenue from imports and assess whether the imported products meet their NTM requirements, $\alpha^M s^M$ are expected to be positive. Realistically, penalties are larger and checks are more careful for the importing countries than for exporting countries so we assume that:

$$\alpha^M s^M > \alpha^X s^X.$$

A trade discrepancy, $Disc$, is defined as the recorded import quantity minus the recorded export quantity, in percentage of the actual quantity:

$$Disc \equiv \frac{Q^M - Q^X}{Q} \tag{8}$$

$$= \frac{AVE}{\alpha^X s^X} - \frac{t + AVE}{\alpha^M s^M}$$

$$= \frac{AVE\left(\alpha^M s^M - \alpha^X s^X\right) - t\alpha^X s^X}{\alpha^X s^X \alpha^M s^M}. \tag{9}$$

Under the assumption that penalty and detection rates are higher in the importing countries, Equation (9) suggests the discrepancy increases with the costs of complying with the NTMs of the importing countries:

$$\frac{\partial Disc}{\partial AVE} = \frac{\alpha^M s^M - \alpha^X s^X}{\alpha^X s^X \alpha^M s^M} > 0. \tag{10}$$

Intuitively, more restrictive NTM requirements of the importing country, which are parsimoniously represented by the larger AVE, increase the trade costs for both the exporting firms (pre-shipment inspection, testing and certification requirements, etc.) and the importing firms (import licensing requirement, cargo inspection, etc.). Therefore, both the importers and the exporters have incentives to under-report their quantity on their customs declaration forms in order to reduce such costs. However, the penalty and detection rates in the importing countries are normally higher, due to more accurate customs procedures driven by revenue collections and regulatory concerns, unlike the exporting customs that generally have fewer reasons to precisely assess shipments against the export declarations. This results in larger under-reporting by the exporters than by the importers, which implies larger trade discrepancies between the official statistics of the importing and exporting countries.

In addition, the rate of detection, s, may depend on a few factors. First, for the importing country, products that have higher tariffs may be subjected to more careful inspection in order to verify import duties. Thus, s^M increases with t :

$$s^{M\prime}(t) > 0.$$

Second, it is generally easier to detect fraud related to misclassification of products or misdeclaration of country of origin for differentiated products than for homogeneous products. For example, customs officers can easily tell a German car from a Japanese car, but may have a hard time identifying Chinese steel rods from Brazilian steel rods. Thus, s could be higher for differentiated products than homogeneous products. This could lead to the effect of NTMs on discrepancies to be smaller for differentiated products than for homogeneous products if s^X increases more than s^M.

Third, developed countries may have better customs administrations and resources to inspect products, relative to developing countries. Thus, the differences between the detection rates of the importing and exporting countries will be larger when the exporting country is a developing country rather than a developed country. This implies that the effect of $NTMs$ on discrepancies will be larger when the developed countries import from the developing countries.

Finally, the effects of tariffs on trade discrepancies depends on the effect of tariffs on the detection rate:

$$\frac{\partial Disc}{\partial t} = -\frac{1}{\alpha^M s^M}\left[1 - \frac{AVE + t}{s^M} s^{M\prime}(t)\right] \lesseqgtr 0, \text{ given } s^{M\prime}(t) > 0. \tag{11}$$

Equation (11) shows that, if the detection rate is very sensitive to tariffs, then an increase in tariffs could lead to a rise in trade discrepancies. This is because while the importing firms may be increasing under-reporting when facing a higher tariff, they risk being detected and face a much larger penalty. So in equilibrium, the size of under-reporting may actually decrease with a higher tariff which implies a larger trade discrepancy, given a fixed level of export quantity.

In summary, this simple model delivers the following testable hypothesis:

1. Trade discrepancies increases with the restrictiveness of NTMs imposed by the importing countries
2. Origin fraud and product classification fraud exist, particularly for trade between high-income importing countries and developing exporting countries
3. The effects of importing country NTMs on trade discrepancies are larger for trade between high-income importing countries and developing exporting countries
4. The effects of importing country NTMs on trade discrepancies are larger for homogeneous products than for differentiated products
5. The effects of tariffs on trade discrepancies could be positive

3 Data

The tariff and NTM data used in this paper are from the UNCTAD TRAINS database. Trade data originates from the UNSD COMTRADE database. The NTM data is collected only periodically, and the data utilized in this exercise was collected between 2015 and 2019. As NTM data generally varies little across time, the data on NTM is assumed to reflect the existing NTMs as of 2018. Consequently, trade and tariff data are matched to the same year.

For the purpose of this paper, we focus on Border NTMs, while retaining other types of NTMs as control variables.[13] Border NTMs include costs incurred by imported goods such

[13]NTMs are distinguished between border and non-border variables on the basis of the international classification of non-tariff measures (UNCTAD, 2019) and defined as custom measures by Ederington and Ruta (2016)

as customs controls, quota licensing, pre-shipment inspections, and additional fees paid at customs, among many others.[14]

Table 1 presents the list of importing countries included in our sample, together with their coverage ratios of border NTMs. There are a total of 49 importing countries with 96 exporting countries. The sample is largely determined by the availability of NTM data in the UNCTAD TRAINS database. There is a wide range of variations in terms of the coverage ratios across these different importing countries. We exploit this wide range of variations across countries to estimate the AVEs at the HS 6-digit level.

Trade discrepancies are constructed by matching import data with export data at the HS 6-digit level according to Equation (8) for both quantity and value, based on UNSD Comtrade HS 6-digit level import and export data. Since we do not observe the true quantity, Q, we replace it with the average value of import and export quantity and value. The quantity data were thoroughly examined, and we include only observations in which the importer and the exporter use the same quantity unit in their reporting, to make sure that the discrepancies are not due to differences in the units of measurement in quantity. In addition, the quantity units remain the same within any given HS 6-digit product.[15]

4 Estimation of the Ad Valorem Equivalent of NTMs

In the theoretical model, the impacts of NTMs on trade discrepancies work through the ad valorem equivalent (AVE) of NTMs. To empirically test the hypothesis, we will need estimates of these $AVEs$, which vary by product and importer-exporter pair. There is no such detailed estimates available in official published statistics. We will need to estimate these $AVEs$. The closest estimates are provided by Kee, Nicita and Olarreaga (2009), at

[14] Please refer to Ederington and Ruta (2016) for details.

[15] For the regression analysis, we always control for product fixed effects so that we do not rely on differences in the quantity units across products to estimate the coefficients of interest.

product and importer level, without variations across exporters.

In this paper, to estimate the AVE of the NTMs, we run quantity-based gravity regressions for each of the nearly 5,000 HS 6 digit-level products for 2018. The dependent variable of the gravity regressions is the quantity traded of an HS 6-digit product between a country pair, and the right-hand variables include a dummy variable indicating the presence of Border NTMs and the bilateral applied tariff on the product specific to the country pair. In these gravity regressions, the coefficients of tariffs can be interpreted as the trade elasticities, which are used to re-scale the coefficients of the NTM dummy variables to obtain the AVE of NTMs. The dependent variable is treated as a discrete, non-continuous variable, and we use a count data regression technique such as Poisson or negative binomial models for our estimations.[16]

4.1 Econometric Issues

There are some econometric issues that need to be addressed before we present our estimation model. First is the distribution of the dependent variable. For each HS 6-digit product, there is an excessive number of bilateral country pairs that do not trade. The share of quantity imported that is positive is only 6.8 percent in the sample. This leads to the large presence of zero observations in the bilateral data set as shown in Figure 2. Moreover, for the country pairs that do trade, the volume of their trade has a very large range that leads to an extreme over-dispersion as shown in Figure 3. Overall, the ratio of the variance of the positive quantity and its mean is 4.83e+08 in our data set. The large presence of zeroes together

[16] This is necessary because nearly 25 percent of the observations have "Number of items" as the unit of measurement for quantity imported, which clearly is discrete in nature. The rest of the 75 percent of the observations have "Weight in kilograms" as the unit of measurement, which is often recorded as a discrete variable. Thus, nearly 100 percent of the observations recorded quantity imported as discrete integers. Therefore, it is appropriate to use a count data regression technique, such as Poisson or negative binomial models for our estimation. Furthermore, Santos Silva and Tenreyro (2006, 2011) showed that the Poisson pseudo-maximum likelihood (PPML) estimator usually used for count data works well even when they used the value of aggregate bilateral imports, which itself is not a count variable, as the dependent variable.

with over-dispersion led us to use the zero-inflated negative binomial (ZINB) model for the main estimation specification. Nevertheless, for each HS 6-digit product, in addition to the ZINB, we also run the negative binomial (NB), the zero-inflated Poisson (ZIP), the ordinary least squares (OLS) in log and the Poisson models, and use the Vuong test, which is a model specification test, to pick the best-fitting regressions for the AVE estimates.

For ZINB, we assume that the observed trade data is the realization of the mixture of two distinct distributions: one distribution governs the probability of zero trade (participation or extensive margin), and the second distribution governs the realization of positive trade (volume or intensive margin). Let Q_{nij} be the quantity of product n imported by country i from country j, and let $h\left(Q_{nij}, \theta|\mathbf{X}\right)$ denote the negative binomial density with mean $e\left(\mathbf{X}\beta\right)$, dispersion parameter α, while θ includes α and β. Here, $\mathbf{X}$ is a vector of variables to explain the positive import quantity, which are standard gravity variables, $\mathbf{Z}_{ij}$, as well as tariffs and NTMs:

$$
\begin{aligned}
Q_{nij} &\sim \begin{cases} 0\text{, with probability } p_{nij} \\ \tilde{Q}_{nij}\text{, with probability } 1-p_{nij} \end{cases}, \\
\tilde{Q}_{nij} &= 0,1,2,... \sim h\left(Q_{nij}, \theta|\mathbf{X}\right).
\end{aligned}
$$

Therefore, the import quantity density distribution can be modeled as

$$
pr\left(k\right) = \begin{cases} p_{nij} + \left(1-p_{nij}\right) h\left(Q_{nij}=k, \theta|\mathbf{X}\right)\text{, if } k=0 \\ \left(1-p_{nij}\right) h\left(Q_{nij}=k, \theta|\mathbf{X}\right)\text{, if } k=1,2,... \end{cases} \tag{12}
$$

The parameter, p_{nij},which is used to increase the presence of zeros in the data set, could depend on covariate $\mathbf{W}$ and is commonly modeled to follow a logit function to ensure its range:

$$p_{nij} = \frac{\exp\left(\mathbf{W}\gamma\right)}{1+\exp\left(\mathbf{W}\gamma\right)} \in (0,1)\,. \tag{13}$$

In general, the same set of variables could be included in both $\mathbf{W}$ and $\mathbf{X}$ and we do not need to identify any additional variables for the participation equation (see Cameron and Trivedi, 1998). Nevertheless, we still include a common religion indicator in $\mathbf{W}$, motivated by Helpman, Melitz and Rubinstein (2008) to estimate Equation (13).

The second modeling issue is related to the estimation of AVE with bilateral variations. There are two main reasons why product level trade elasticities and AVEs may have bilateral variation across exporting countries for each importing country. First, the import demand of the products from different source countries are likely to have different price elasticities empirically.[17] By incorporating bilateral variation into these trade elasticities, we will be better equipped to analyze the impacts of trade negotiations or agreements with specific partner countries. For our current context, the bilateral differences in import demand further imply that the resulting AVEs will have bilateral variations, given that these trade elasticities are used to convert the trade impacts of NTMs into AVEs. Second, some countries in our data set have bilateral variations in their NTMs to target specific partner countries, which naturally will lead to bilateral variations in AVEs.[18] Furthermore, even if the NTMs of the importing countries are not country or product specific, the compliance costs of NTMs are likely to vary across exporting countries and products, which will give rise to bilateral variations in the AVEs of an importing country at a product level.[19] Rather than assuming and imposing that all trading partners face the same restrictiveness due to an NTM, we allow

[17]For example, when faced with a 10 percent tariff increase, the responsiveness of the US import demand for cars from Germany could be very different from cars from the Republic of Korea.

[18]For example, Kazakhstan, the Russian Federation and Belarus created the Eurasian Customs Union in 2010 so their NTMs do not apply to member countries.

[19]For example, Sri Lanka has a certification requirement (A830) for more than 1,600 imported products, such as Mackerel (HS 160415), which comes from China, Thailand and Peru. These trading partners have different stages of development and sophistication when conforming with such an NTM requirement.

the data to speak for themselves, by adopting an econometric model that allows bilateral variations in the coefficient for the NTMs.[20]

Our empirical strategy relies on using the market powers of the exporter and importer to estimate the size of the AVEs and the trade elasticities. This is motivated by the terms of trade models or trade theories (Bagwell and Staiger, 2011, for tariffs, and Staiger and Sykes, 2011, for NTMs). In these theories, importing countries with market power (i.e. countries which can influence international prices through their trade policies) will have terms of trade gains to impose tariffs and NTMs. This is because they can "pass-though" some of the costs to the trading partners given that their trading partners have price sensitive supply curves. If the importing country has no market power, imposing NTMs will cause imports to decrease and the domestic price of the imports to increase to the full extent of the compliance costs. On the contrary, if the exporting country has no market power, NTMs of the importing country will have no effect on imports or domestic prices. The resulting AVE depends on the market powers of the importing country and the exporting country.

To this end, we interact the tariff and NTM of an HS6 digit product with the following two variables: the share of the exporter in the world trade of this product, and the share of the importer in the world trade of this product. Both variables capture the market powers of the exporter and the importer in the world market for each HS 6-digit product. We expect the trade impact to be lower if the exporter has a larger share in the world trade of the product, which could result in a smaller trade impact due to the presence of the importing country's NTMs. However, it is also possible that the exporter could easily divert its exports to other markets when faced with the burdensome NTMs of a specific importer, and the trade impact will be larger. Likewise, if the importing country is an important market with a large

[20] Overall, we consider the estimation of bilateral variations in trade elasticities and AVEs to be crucial and necessary for improving the empirical applications of these estimates, particularly for researchers and policy makers to evaluate or renegotiate trade agreements.

market share of a product, it could be more difficult to comply with the country's NTMs. Then the compliance cost for the exporting countries could be higher and that would lead to a larger trade impact. However, it is also possible that if the market share of the importer is large, then it is more difficult for exporting countries to divert their products to other markets, so the trade impact of the NTMs could be lower. Overall the trade effects of the market shares of the importer and the exporter in the world market due to the presence of NTMs depends on the specific products and the country-pair. Therefore we let the data reveal the dominant force. A similar argument can be made for tariffs and the resulting bilateral import elasticities.

The third econometric issue to be addressed is related to the endogeneity of the right-hand side variables. Specifically, tariffs and NTMs are trade policies set by the importers and could be endogenous to the trading volume. Including tariffs and NTMs in the regressions will lead to inconsistent estimates. To address this endogeneity issue, we use the simple averages of tariffs and NTMs of the three closest neighboring countries in our sample as the instrumental variables (IVs) respectively. A recent work by Jiao and Wei (2020) provides a theoretical justification based on the median voter theorem regarding why trade policies of other countries are exogenous, and hence satisfy the exclusion restriction, but may affect trade policies of a country which make them reasonable IVs. In addition, the trade policies of neighboring countries are likely to be correlated with the trade policy of the importing country due to a regional trading agreement or a common cultural/history background, but the bilateral imports of the importing country from a particular trading partner should not have any direct impact on the bilateral trade policies of neighboring countries on the same trading partner and thus satisfy the exclusion restrictions.[21] Please note that this could

[21] For example, bilateral imports of Sri Lanka from China could be lower due to the presence of Sri Lanka's tariffs and NTMs on Chinese products. However, we worry that these tariffs and NTMs could depend on Sri Lanka's imports of these Chinese products due to domestic political economy and local industry lobbying considerations. To address this endogeneity issue, we use the simple average of the bilateral tariffs of India,

be an issue if these countries are all part of a regional trade bloc with common external tariffs and NTMs. However, we believe this is not an issue in our data set as we have a wide range of countries in the sample and not all countries are part of any regional trade bloc with common external trade barriers (for this reason, the European Union is treated as one country). Moreover, the domestic lobbying activities of neighboring countries are not expected to be correlated among neighboring countries.[22] We will check the first-stage regression results to make sure that these neighboring countries' bilateral trade policies have the right signs and acceptable F-statistics to rule out the issue of weak instruments.

In the second stage of the IV regressions, we replace the bilateral tariffs with their fitted values from the first-stage regressions. For the bilateral border NTMs, which is a discrete dummy variable, we run a probit (selection) regression for the bilateral NTM variable in the first stage, and construct the inverse Mills ratio according to the first-stage regression to be included along side the bilateral NTMs in the second-stage regressions (see Heckman, 1979).

4.2 Econometric Models

Taking into account the above econometric issues, first we run the first-stage IV regressions, and then we run the product-level quantity-based gravity regressions based on cross-section data using five models: the zero-inflated negative binomial model, the negative binomial model, the zero-inflated Poisson model, the Poisson model and the log-OLS model.

Bangladesh and Pakistan on Chinese products as the instrument for the bilateral tariff of Sri Lanka on the same Chinese products. Similarly, we use the simple average of the bilateral NTMs of India, Bangladesh and Pakistan on Chinese products as the instrument for the bilateral NTMs of Sri Lanka on the same Chinese products. The reasoning is that, while the trade policies of India, Bangladesh and Pakistan could be correlated with the trade policies of Sri Lanka due to regional and cultural proximity, the bilateral imports of Sri Lanka from China of a particular product should NOT a priori directly affect the trade policies of India, Bangladesh and Pakistan on these same Chinese products. In other words, we do not expect domestic political economy and local industry lobbying of Sri Lanka on Chinese products to have any direct impacts on the determinations of bilateral trade policies of India, Bangladesh and Pakistan on these Chinese products. Hence these instruments satisfy the exclusion restriction.

[22] For example, the industrial structure of Sri Lanka is very different from the industrial structures of India, Bangladesh and Pakistan in terms of the products they produce and trade.

For the first-stage regression for tariffs, we use the average tariff of the three closest countries, $\bar{t}_{nij}$, as an instrument for the tariff, t_{nij}, to get the fitted tariff, $\hat{t}_{nij}$:

$$\hat{t}_{nij} = \mathbf{Z}_{ij}\hat{\nu}+\hat{\nu}^t\bar{t}_{nij}, \tag{14}$$

where $\hat{\nu}$ and $\hat{\nu}^t$ are the least squares estimates of the coefficients of the second-stage control variables, Z_{ij}, and the average tariff, $\bar{t}_{nij}$.

For the first-stage regression for the NTMs, because NTM is a dummy variable that equals one when at least one NTM is present, the first-stage regression is a probit regression with a density function, $f(.)$, in which we use the average presence of NTMs in the three closest countries, $\overline{NTM}_{nij}$, as an instrument for NTM_{nij} :

$$\begin{aligned} NTM_{nij} &\in \{0,1\}, \\ f\left(NTM_{nij}|\mathbf{X},\overline{NTM}_{nij}\right) &= \Phi\left(\mathbf{Z}_{ij}\delta+\delta^{NTM}\overline{NTM}_{nij}\right)^{NTM_{nij}}\left[1-\Phi\left(\mathbf{Z}_{ij}\delta+\delta^{NTM}\overline{NTM}_{nij}\right)\right]^{1-NTM_{nij}}, \\ invMill_{nij} &= \frac{\phi\left(\mathbf{Z}_{ij}\hat{\delta}+\hat{\delta}^{NTM}\overline{NTM}_{nij}\right)}{\Phi\left(\mathbf{Z}_{ij}\hat{\delta}+\hat{\delta}^{NTM}\overline{NTM}_{nij}\right)}, \end{aligned}$$

where we retrieve the inverse Mills ratio, $invMill_{nij}$, constructed based on $\hat{\delta}$ and $\hat{\delta}^{NTM}$, the coefficients of the second-stage control variables, Z_{ij}, and the average NTM presence, $\overline{NTM}_{nij}$.

Finally, we also run the first-stage regressions for the four interaction terms that involved tariffs and NTMs, using the interaction terms with $\bar{t}_{nij}$ and $\overline{NTM}_{nij}$ as the respective instruments:

$$\begin{aligned} share_{ni}\hat{t}_{nij} &= \mathbf{Z}_{ij}\hat{\omega}_1+\hat{\omega}_1^t share_{ni}\bar{t}_{nij}, \\ share_{nj}\hat{t}_{nij} &= \mathbf{Z}_{ij}\hat{\omega}_2+\hat{\omega}_2^t share_{nj}\bar{t}_{nij}, \\ share_{ni}NTM_{nij} &= \mathbf{Z}_{ij}\hat{\omega}_3+\hat{\omega}_3^t share_{ni}\overline{NTM}_{nij}, \\ share_{nj}NTM_{nij} &= \mathbf{Z}_{ij}\hat{\omega}_4+\hat{\omega}_4^t share_{nj}\overline{NTM}_{nij}. \end{aligned} \tag{15}$$

The main specification of second-stage regression for the ZINB, the NB, the ZIP and the Poisson models is

$$\begin{aligned} \ln E\left(Q_{nij}|\mathbf{X}\right) &= \beta_n + \beta_{nij}^t \hat{t}_{nij} + \beta_{nij}^{NTM} NTM_{nij} + \beta^M invMill_{nij} + \mathbf{Z}_{ij}\beta + \varepsilon_{nij}, \\ \beta_{nij}^t &= \beta_n^t + \beta_1^t share_{ni} + \beta_2^t share_{nj}, \\ \beta_{nij}^{NTM} &= \beta_n^{NTM} + \beta_1^{NTM} share_{ni} + \beta_2^{NTM} share_{nj}. \end{aligned} \tag{16}$$

The control variables included in Z_{ij} are the standard gravity variables: the log of the gross domestic product (GDP) of the importer and the exporter, the bilateral distance between the importer and the exporter, the landlocked indicators for the importer and the exporter, the common boarder indicator, and also a dummy controlling for other types of non-border NTMs measures. The bilateral coefficients of tariffs and NTMs, β_{nij}^t and β_{nij}^{NTM}, are obtained by using the interaction terms based on the importer's share in the world market, $share_{ni}$ and the exporter's share in the world market, $share_{nj}$. For log-OLS, the second-stage regression is similar to equation (16) but with the log of the quantity imported, $\ln\left(Q_{nij}\right)$, as the dependent variable.

For the zero-inflated specifications, such as the ZINB and the ZIP models, we also include a logit regression to explain the presence of excessive zeros in the bilateral trade. Although it is not necessary to have an extra variable to explain zeros (see Cameron and Trivedi, 1998), we still include a common religion indicator in the zero regression, motivated by Helpman, Melitz and Rubinstein (2008), in addition to all the second-stage control variables, Z_{ij}, to form W in equation (13).

Once all five versions of equation (16) are estimated, we use Vuong tests to select the best fit models to retrieve $\hat{\beta}_{nij}^{t}$ and $\hat{\beta}_{nij}^{NTM}$ for the construction of the AVE. For all five models, $\hat{\beta}_{nij}^{t}$ has the interpretation of the (instantaneous) semi-elasticity of trade with respect to a one percentage point increase in the tariff:

$$\hat{\beta}_{nij}^{t} = \frac{\partial \ln E\left(Q_{nij}|\mathbf{X}\right)}{\partial t_{nij}}. \tag{17}$$

To obtain the instantaneous elasticity of trade with respect to tariffs, we multiply $\hat{\beta}_{nij}^{t}$ with t_{nij}:

$$\varepsilon_{nij} \equiv \frac{\partial \ln Q_{nij}}{\partial \ln t_{nij}} = \hat{\beta}_{nij}^{t} t_{nij}.$$

Note that unlike the trade elasticity in the literature (see Simonovska and Waugh, 2014 for a review), these trade elasticities are estimated at the HS 6-digit level with bilateral variations. In addition, these trade elasticities are jointly estimated with the AVE of the NTMs, which reduces the potential downward bias due to the negative correlations among tariffs, NTMs and trade flow.

To construct the AVE of the NTMs, we need to first construct the proportionate change in quantity imported (or the expected quantity imported for the count models) due to the presence of NTMs:

$$\begin{aligned}
\ln E\left(Q_{nij}|\mathbf{X},NTM_{nij}=1\right)-\ln E\left(Q_{nij}|\mathbf{X},NTM_{nij}=0\right) &= \hat{\beta}_{nij}^{NTM} ==> \\
\ln\left(\frac{E\left(Q_{nij}|\mathbf{X},NTM_{nij}=1\right)}{E\left(Q_{nij}|\mathbf{X},NTM_{nij}=0\right)}\right) &= \hat{\beta}_{nij}^{NTM} \\
\frac{E\left(Q_{nij}|\mathbf{X},NTM_{nij}=1\right)}{E\left(Q_{nij}|\mathbf{X},NTM_{nij}=0\right)} &= \exp\left(\hat{\beta}_{nij}^{NTM}\right) \\
\frac{E\left(Q_{nij}|\mathbf{X},NTM_{nij}=1\right)-E\left(Q_{nij}|\mathbf{X},NTM_{nij}=0\right)}{E\left(Q_{nij}|\mathbf{X},NTM_{nij}=0\right)} &= \exp\left(\hat{\beta}_{nij}^{NTM}\right)-1.
\end{aligned} \tag{18}$$

We then use the semi-elasticity of trade with respect to a one percentage point increase in the tariff, $\hat{\beta}_{nij}^{t}$, to convert the proportionate change in quantity imported due to NTMs to the ad valorem equivalent tariff:

Definition 1 *The ad valorem equivalent of NTM (AVE_{nij}) of product n, in importing country i, from exporting country j, measures the ad valorem tariff that induces the same proportionate change in the quantity imported as the presence of NTM_{nij}, or*[23]

$$AVE_{nij}=\frac{\exp\left(\hat{\beta}_{nij}^{NTM}\right)-1}{\hat{\beta}_{nij}^{t}}. \tag{19}$$

Equation (19) above makes it clear that AVE by construction will have product-importer-exporter variation. Variation across products is obvious, because all the regressions are estimated at the HS-6 product level, so all the coefficients will have product variations. The reasons we have variation across importers and exporters are twofold. First, even if NTMs

[23] There are other ways to define the AVE, such as the equivalent tariff that induces the same change in the quantity imported, or the equivalent tariff that induces the same rate ratio change in the quantity imported. In those cases, the formula for the AVE could be slightly different. Kee, Nicita and Olarreaga (2009) use the import demand elasticity, ε, to obtain the following:

$$AVE=\frac{\exp\left(\hat{\beta}^{NTM}\right)-1}{\varepsilon}.$$

are multilateral with no partner country specific measures (which is not true for some countries in our sample), the bilateral variation in import demand elasticities at the denominator will ensure that AVEs have variation across importers and exporters. Furthermore, compliance costs may vary across exporting countries even if they face the same NTMs of an importing country, which are captured by the numerator of Equation (19) . Overall, the size and magnitude of the AVE of a product of an importing country from an exporting country depends on both the compliance costs and market power of both the importing country and the exporting country.

Finally, we bootstrap this procedure 50 times to obtain the bootstrap standard errors of AVE_{nij}.

5 Results for the AVE Estimations

The first-stage regressions fit very well. The average first-stage F-statistics for tariffs is 142.47, with a median of 104.01, 1st percentile of 8.46, and 99th percentile of 633.56. Similarly, the average first-stage Chi-squared statistics for NTMs is 442.19, with a median of 346.07, 1st percentile of 12.25, and 99th percentile of 1532.15. These results indicate that tariffs and NTMs of neighboring countries are not weak instruments for own tariffs and NTMs.

For the second stage regressions, the NB model works best for 71 percent of the products, the ZINB for 10 percent of the products, Poisson for 2 percent, and ZIP for 1 percent. For the remaining 16 percent of products, non-linear models did not converge and therefore OLS was used. Based on these estimates, we constructed the AVE of NTMs according to equation (19). At the sample mean, when the border measures are present, the average AVE is 4.7 percent, whereas the mean tariff is 6.9 percent.

What is the relationship between NTMs and the tariffs in our sample? Column (1) of Table (2) presents the regression result when we regress the tariffs on the AVEs, controlling for importer, exporter and product fixed effects. The negative and significant coefficient suggests that the restrictiveness of NTMs and tariffs are negatively correlated. Thus, at the sample mean, NTMs and tariffs are policy substitutes. Column (2) focuses on the variation of tariffs and NTMs when we compare among all exporting countries controlling for import-product and exporter fixed effects. The coefficient on AVEs remains negative, suggesting that NTMs are more restrictive for those exporting countries that face lower tariffs for a given importing country and product. This result is consistent with the argument that more restrictive NTMs are in place for those countries that receive tariff preferences.

Column (3) focuses on the variation of tariffs and NTMs when we compare among products controlling for importer-exporter fixed effects. The coefficient on AVEs remains negative, which is consistent with the hypothesis that NTMs are more restrictive for those products that have lower tariffs for a given pair of importing and exporting countries. In other words, when the high income countries grant tariff preference to their trading partners on certain products, they may impose more restrictive NTMs on other products from the same exporting countries.

Despite the complicated estimation procedure with more than 700,000 estimates, our elasticity estimates are quite reasonable. First, at the sample mean, the average trade elasticity for HS 6-digit products is about -4.7, which is a bit higher than Simonovska and Waugh's (2014) structural estimates. Given that our estimates are at a more disaggregate level, the higher average elasticity is to be expected. In addition, we controlled for NTMs in the estimation of these trade elasticities based on tariffs. Given that tariffs and NTMs are negatively correlated in our sample and are both negatively correlated with trade flows, including NTMs in the trade flows regressions to estimate trade elasticity would minimize the bias of the coefficient on tariffs and lead to higher elasticity estimates.

In addition, as a robustness check, we split the HS 6-digit products into homogeneous versus differentiated products according to Rauch's classification. As expected, the homogeneous products are significantly more elastic than the differentiated products. At the sample mean, the trade elasticity for the homogeneous products is -6.2, while the trade elasticity for the differentiated products is -4.2, and the difference is statistically significant even after we controlled for the importer-exporter pair fixed effects.[24]

Finally, instead of using the market shares of importing and exporting countries to measure market powers in Equation 16, we use the logs of the GDPs of the importing and exporting countries as a robustness check. This is because larger countries may have more market power to influence world prices according to the terms of trade theory. One problem with using the log of GDP is that it does not vary by products. In other words, we are assuming that larger countries have more market power for all products, which may not be the case empirically and may introduce bias in the estimations. This is why we prefer our original market share variables which measure the market power of each country for each product. Nevertheless, results based on interacting the log of GDP of the importing and exporting countries are quite consistent with our original results. The correlation between the AVE estimated using market shares and the AVE estimated using GDPs is 0.36, while the correlation between the import demand elasticity estimated using market shares and the import demand elasticity estimated using GDPs is 0.58. Regressing the AVE estimated using GDPs on the AVE estimated using market shares results in a R-square of 0.74. So both sets of estimates are positively correlated. Overall we are confident that the results based on our original market share estimations are robust and consistent.

[24] Regression results are available upon request.

6 Results on NTM-Induced Trade Fraud

6.1 Least Squares Downward Bias and Instrumental Variable Estimations

With the estimated AVEs we verify the five testable hypothesis of the theoretical model. Table 3 presents the least squares results when we regress trade discrepancies on the AVE of Border NTMs and tariffs. The dependent variables for Columns (1) to (3) are the quantity discrepancies, while the dependent variables for Columns (4) to (6) are the value discrepancies. Both sets of discrepancies are constructed according to equation (8) . We also tried other ways of measuring the dependent variables, but the results are very similar. For all the regressions in this table, we control for importer-, exporter- and product-fixed effects, and we cluster the standard error by importer-product, which is the level of aggregation of most NTM variables. Clustering is particularly necessary, as the AVEs are estimated with errors.

Columns (1) and (4) regress trade discrepancies on the estimated AVEs of border NTMs. Columns (2) and (5) include bilateral tariffs in the regressions. Columns (3) and (6) sum up the AVEs and tariffs together as one variable and include the presence of the other NTMs as control variables in the regressions. In all specifications, the least squares estimated coefficients on the AVEs are positive and are statistically significant at the 1 percent level.

However, there are reasons to believe that these least squares estimates for the AVEs have downward bias. First, a primary concern is that corruption at the border could cause trade discrepancies and at the same time be paired with a lack of enforcement on NTM regulations. This may result in both larger discrepancies and less restrictive AVEs, therefore biasing the LS estimation towards zero. Moreover, a large part of the trade discrepancies could be due to idiosyncratic and random recording errors and not NTM-induced trade fraud. In this scenario, it will be harder to find a systematic relationship between the trade discrepancies and the NTMs. This will further push the least squares estimates toward zero. Finally, there may be a concern that the AVEs are contaminated with estimation errors that could lead to systematic downward bias in the least squares results presented in Table 3. We address all these concerns with the use of instrumental variables, which are correlated with the AVEs but not necessarily correlated with trade. For importing country i, product n and exporting country j, we use the average AVE of exporting country j of the product n in the non-i markets, $\overline{AVE}_{n\text{\textasciitilde}ij}$, as the instrument for AVE_{nij}. This is a valid instrument, because it captures the compliance cost of exporting country j for product n, facing similar NTM measures in the other markets. Thus, potentially it could be positively correlated with AVE_{nij}, and yet it should not affect the trade discrepancies between country i and j on product n.[25]

Table 4 presents the first-stage regression results for the IV estimations. Columns (1), (2), (4) and (5) are for AVE_{nij}, corresponding to Columns (1), (2), (4) and (5) of Table 3. Likewise, Columns (3) and (6) are for $AVE_{nij} + Tariff$, corresponding to Columns (3) and (6) of Table 3. All standard errors are clustered by importer-product. Across all specifications, the coefficients on the average AVE of exporting country, $\overline{AVE}_{n\text{\textasciitilde}ij}$, are positive and statistically significant. Together with the high first-stage F-statistics, they confirm that $\overline{AVE}_{n\text{\textasciitilde}ij}$ is a valid instrument.

Table 5 presents the second-stage regression results. Similar to Table 3, the dependent variables for Columns (1) to (3) are the quantity discrepancies, while the dependent variables for Columns (4) to (6) are the value discrepancies, and we cluster the standard errors by importer-product. Compared to Table 3, the results in Table 5 are much larger, confirming the downward bias in the previous least squares estimates. The coefficients for the AVEs are

[25] We also tried using the trade weighted average AVE of exporting country j of product n in non-i markets as the instrument. The results are very similar.

consistently higher and statistically significant across all columns. Overall, at the sample mean, a 10 percentage point increase in the AVEs leads to a 15 percent increase in trade discrepancies, while a comparable increase in tariffs leads to only an 8 percent increase in trade discrepancies. This suggests that a significant part of the trade discrepancies results from NTM avoidance schemes, as predicted by the theoretical model. These results are consistent with the testable hypothesis 1, which is that trade discrepancies increases with the restrictiveness of NTMs. Moreover, the positive relationship between trade discrepancies and tariffs also provides support for testable hypothesis 5 of the model.

6.2 Sources of Trade Fraud: Origin Fraud and Product Classification Fraud

What can explain the positive relationship between trade discrepancies and AVEs? It could be that firms mis-declared the country of origin of their products in order to avoid the burdensome NTMs. This would be consistent with origin fraud. If this is the case, we expect that, given a specific importing country and product, exporting countries that have higher AVEs are also those that have larger trade discrepancies with the importing country. Tables 6 and 7 test the origin fraud hypothesis with least squares and IV regressions, respectively. First-stage regressions are presented in Table 8. Instead of controlling for a full set of fixed effects, we control for importer-product fixed effects and exporter fixed effects. Thus, we rely on the variation of the AVEs of the exporters in the same importing market with the same product to identify the coefficient of the AVEs. Similar to the previous tables, the dependent variables for Columns (1) to (3) are the quantity discrepancies, while the dependent variables for Columns (4) to (6) are the value discrepancies, and we cluster the standard errors by importer-product. Both least squares and IV results show that, within the importer-product pairs, the exporters that face higher AVEs are those that have larger trade discrepancies with the importers. This is consistent with origin fraud, whereby the firms exporting products coming from countries subject to high AVEs tend to mis-declare the country of origin to benefit from a lower AVE.

The other type of trade fraud is product classification fraud, whereby products are misclassified as other products in order to evade or reduce the costs associated with NTMs compliance. This type of fraud results in, given a specific importing-exporting countries pair, products that have higher AVEs are also those that have larger trade discrepancies. Tables 9 and 10 test the product classification fraud hypothesis with least squares and IV regressions, respectively. First-stage regressions are presented in Table 11. Instead of controlling for a full set of fixed effects, we now control for importer-exporter fixed effects and product fixed effects. In other words, we rely on the variation in the AVEs across products within the importer-exporter pairs to identify the coefficient on the AVEs. Similar to the previous tables, the dependent variables for Columns (1) to (3) are the quantity discrepancies, while the dependent variables for Columns (4) to (6) are the value discrepancies, and we cluster the standard errors by importer-product. Both tables show that, within the importer-exporter pairs, products that face higher AVEs are those that have larger trade discrepancies. This is consistent with product classification fraud, whereby firms mis-declared products to lower costs associated with NTMs.

Overall, evidence presented in Tables 3 to 10 are consistent with NTM-induced trade fraud, when firms either mis-declared their country of origin or their product codes in order to avoid the burdensome NTMs of the importing countries. This provides evidence supporting the second testable hypothesis of the theoretical model. In addition, comparing the point estimates presented in Tables 7 and 10 suggests that product classification fraud is more sensitive to restrictive NTMs, relative to country of origin fraud. In other words, facing restrictive NTMs, more traders may choose to misreport product classification codes than the country of origin. This is consistent with the anecdotal evidence presented at the start of the paper in italics.

7 Robustness Checks

The regression results presented above show that larger trade discrepancies are associated with more restrictive NTMs. To further validate this, we split the sample according to the product characteristics and the country characteristics.

7.1 Homogeneous versus Differentiated Products

Homogeneous products are easier for firms to either mis-declare the product codes or country of origin in order to evade restrictive NTMs. For example, it is easier to declare aluminum tubes (HS 760810) from China as aluminum tubes from Brazil, given that aluminum tubes are similar regardless of their origin. In contrast, it is harder to declare cars (HS 870390) from Japan as cars from Germany, given that cars are heterogeneous (Toyota vs Volkswagen). Likewise, it is easier to declare aluminum tubes as aluminum rods (HS 760410), than to declare cars as buses (HS 870210). Therefore we expect the effects of NTMs on trade discrepancies to be larger for homogeneous products.

Table 12 presents the second-stage IV regression results for origin and product classification fraud, with a sub-sample of homogeneous and differentiated products. We instrument the AVEs the same way as before while controlling for tariff, and the presence of other NTMs. The first-stage results are available upon request. Columns (1) to (6) are for origin fraud, while Columns (7)-(12) are for product classification fraud. Columns (1)-(3) and (7)-(9) consist of a sub-sample of differentiated products, while Columns (4)-(6) and (10)-(12) are the results for homogeneous products. The effects of NTMs on trade discrepancies remain positive and statistically significant for all these specifications, with the point estimates larger for homogeneous than differentiated products.[26] These results support testable hypothesis 4 of the model. This is also consistent with the findings of Javorcik and Narciso (2008), who claim that quantity measures are more relevant for the mis-declaration of homogeneous goods since homogenous goods are often sold/measured by weight.[27]

7.2 High-Income Countries versus Developing Countries

Countries with higher income often have better infrastructure and institutions to enforce rules of law, including NTMs, and at the same time have better customs enforcement to accurately record trading transactions. We expect the effects of NTMs on trade discrepancies to be smaller for trade between high-income importing and exporting countries, than between high-income importing country and developing exporting countries.

Table 13 presents the second-stage IV regression results for a sub-sample of high-income importing countries. Similar to Table 12, Columns (1) to (6) are for origin fraud, while Columns (7)-(12) are for product classification fraud. Columns (1)-(3) and (7)-(9) are the results when the exporting countries are developing countries. In these columns, all the AVE variables remain positive and highly significant. In contrast, Columns (4)-(6) and (10)-(12) are the results when the exporting countries are also high-income countries. None of the AVE variables are statistically significant. Together these results suggest that trade fraud mainly exist between high-income importing countries and developing exporting countries. These results support testable hypothesis 3 of the theoretical model.

Overall, these results are consistent with our hypothesis that NTM-induced trade fraud

[26] Results based on regressions for a combined sample of homogeneous and differentiated products cannot reject the hypothesis that the effects of NTMs on trade discrepancies are lower for differentiated products than for homogeneous products, for both product fraud and origin fraud, controlling for tariffs.

[27] When we restrict the sample of homogeneous products which are measured by weight for Table 12, the results are consistently larger and stronger, confirming the finding of Javorcik and Narciso (2008).

is particularly relevant for homogeneous products and for the trade between developed and developing countries. This is consistent with the theoretical model that it is easier for firms to mis-declare product codes or countries of origin to evade restrictive NTMs when products are homogeneous, and when customs administrations are less developed.

8 Concluding Remarks

This paper examines whether some of the discrepancies in the official statistics of importing and exporting countries result from trade fraud in which traders deliberately mis-declare the country of origin or product code in order to avoid costly or burdensome NTMs at the borders. Our approach utilizes the detailed product-level trade statistics of a wide range of importing and exporting countries. We first provide a theoretical model to link the restrictiveness of NTMs to trade fraud and to rationalize the presence of origin and product classification fraud. We then estimate the ad valorem equivalents of the NTMs of the importing countries, with variations at the product and exporting country levels and relate the estimated ad valorem equivalents to the existing trade discrepancies, in terms of the quantity and value discrepancies. Our findings indicate that trade flows where larger discrepancies are observed are those that face more restrictive NTMs, suggesting the misclassification of product codes or the misrepresentation of the country of origin in order to reduce or evade the compliance costs associated with Border NTMs. We find that NTM-induced trade fraud is more pronounced for homogeneous products and for trade between developed and developing countries.

One additional contribution of this paper is the quantification of the costs associated with Border NTMs, filling an important void in the trade policy literature which thus far has focused primarily on tariffs. Given that tariffs have been steadily declining for the past few decades, recent trade negotiations and agreements mostly focus on NTMs. From a methodological point of view, this paper presents a novel method for estimating the ad-valorem equivalent of NTMs. In addition, our finding that AVEs and tariffs are substitutes suggests that it is important to include NTMs in trade policy analysis.

Finally, the general results of this paper indicate that NTMs at the border add important costs to trade and that traders seek to avoid these costs by resorting to misclassification of products and/or misreporting quantities. Avoidance of NTMs at the border not only undermines the trade policy of the importing country, for example by misreporting country of origin, but also undermines regulations that seek to guarantee consumers' health and environmental protection.

References

[1] Bagwell, K. and Staiger, R. (2011). "What Do Trade Negotiators Negotiate About? Empirical Evidence from the World Trade Organization," *American Economic Review* 101: 1238–1273.

[2] Bhagwati, J. (1964). "On the Underinvoicing of Imports," *Bulletin of the Oxford University Institute of Economics & Statistics* 27(4): 389–397.

[3] Broda, C., and Weinstein, D. (2006). "Globalization and the Gains from Variety," *Quarterly Journal of Economics* 121(2): 541-585.

[4] Cameron, C. and Trivedi, P. (2013), *Regression Analysis of Count Data*, 2nd edition, Econometric Society Monograph No.53, Cambridge University Press.

[5] Cantens, T. (2015). "Mirror Analysis and Revenue Frauds," *World Customs Organization Research Paper* No. 35.

[6] Chalendard, C., Fernandes, A., Raballand, G., and Rijkers, B. (2020). "Technology (Ab)use and Corruption in Customs," World Bank mimeo.

[7] Demir, B., and Javorcik, B. (2020). "Trade policy changes, tax evasion and Benford's law," *Journal of Development Economics* 144, 102456.

[8] Ederington, J. and Ruta, M. (2016). "Non-Tariff Measures and the World Trading System". *World Bank Policy Research Working Paper* 7661.

[9] Ferrantino, M., and Wang, Z. (2008). "Accounting for Discrepancies in Bilateral Trade: The Case of China, Hong Kong and the United States," *China Economic Review* 19: 502-520.

[10] Ferrantino, M., Liu, X., and Wang, Z. (2012). "Evasion behaviors of exporters and importers: Evidence from the U.S.–China trade data discrepancy," *Journal of International Economics* 86: 141–157.

[11] Fisman, R. and Wei, S. (2004). "Tax Rates and Tax Evasion: Evidence from "Missing" Imports in China," *Journal of Political Economy*, vol. 112, no. 2.

[12] Goldberg, P., and Pavcnik, N. (2016). "The Effects of Trade Policy," *Handbook of Commercial Policy*, edited by Kyle Bagwell and Robert Staiger.

[13] Grigoriou, C. (2019). "Mirror analysis as a support for risk management and valuation: a practical study." *World Customs Journal* 13(2): 91-113.

[14] Grigoriou C., Kalizinje, F., and Raballand, G. (2019). "How helpful are mirror statistics for Customs reform? Lessons from a decade of operational use." *World Customs Journal* 13(2): 105-114.

[15] Heckman, J. J. (1979). "Sample Selection as a Specification Error," *Econometrica* 47(1): 153–161.

[16] Helpman, E., Melitz, M., and Rubinstein, Y. (2008). "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics* 123(2): 441-487.

[17] Hillberry, R., and Hummels, D. (2013). "Trade elasticity parameters for a computable general equilibrium model." *Handbook of Computable General Equilibrium Modeling* 1: 1213-1269.

[18] Hummels, D., and Lugovskyy, V. (2006). "Usable Data? Matched Partner Trade Statistics as a Measure of Transportation Costs" *Review of International Economics* 14(1).

[19] Javorcik, B. and Narciso, G. (2008). "Differentiated Products and Evasion of Import Tariffs," *Journal of International Economics* 76: 208-222.

[20] Jiao, Y., and Wei, S. (2020). "Natural Barriers and Policy Barriers," NBER working paper 28171.

[21] Kee, H. L., and Nicita, A. (2021). "Discrete Trade Reforms with Heterogeneous Elasticities: Assessing the Impacts of a No-Deal Brexit," World Bank Policy Research Working Paper 8195.

[22] Kee, H. L., Nicita, A., and Olarreaga, M. (2008). "Import demand elasticities and trade distortions," *Review of Economics and Statistics* 90(4): 666–682.

[23] Kee, H. L., Nicita, A., and Olarreaga, M. (2009). "Estimating Trade Restrictiveness Indices," *Economic Journal* 119: 172–199.

[24] Kee, H. L., Neagu, C., and Nicita, A. (2013). "Is Protectionism on the Rise? Assessing national trade policies during the crisis of 2008," *Review of Economics and Statistics* 95(1): 342–346.

[25] Lee, J., and Swagel, P. (1997). "Trade barriers and trade flows across countries and industries," *Review of Economics and Statistics* 79 (3): 372-382.

[26] Li, Y., and Lin, F. (2022). "Beyond tariff evasion: bypass effect of FTAs to circumvent technical barriers," *Review of World Economics.*

[27] Liu, R., Sheng, L., and Wang, J. (2021). "Faking Trade for Capital Control Evasion: Evidence from Dual Exchange Rate Arbitrage in China," working paper.

[28] Martin, W., and Pham, C. (2020). "Estimating the gravity model when zero trade flows are frequent and economically determined," *Applied Economics* 52(26): 2766-2779.

[29] Mishra, P., Subramanian, A., and Topalova, P. (2008). "Tariffs, enforcement, and customs evasion: Evidence from India," *Journal of Public Economics* 92 (1011), 1907–1925.

[30] Santos Silva, J.M.C. and Tenreyro, S. (2006). "The Log of Gravity." *Review of Economics and Statistics* 88 (4): 641–658.

[31] Santos Silva, J.M.C. and Tenreyro , S. (2011). "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator." *Economics Letters* 112(2): 220–222.

[32] Sequeira, S. (2016). "Corruption, Trade Costs and Gains from Tariff Liberalization: Evidence from Southern Africa," *American Economic Review* 106(10): 3029-3063.

[33] Simonovska, I., and Waugh, M. (2014). "The elasticity of trade: Estimates and evidence," *Journal of International Economics* 92(1): 34-50.

[34] Staiger, R., and Sykes, A. (2011). "International Trade, National Treatment, and Domestic Regulation," *Journal of Legal Studies* 40: 149-203.

[35] Swenson, D. (2001). "Tax reforms and evidence of transfer pricing," *National Tax Journal* 54 (1), 7–25.

[36] UNCTAD (2015). *International Classification of Non-Tariff Measures*. United Nations publication.

[37] UNCTAD, and World Bank (2018). *The unseen impact of non-tariff measures: insights from a new database.*

[38] World Integrated Trade Solutions (WITS)-Users Manual (2011).

Figure 1: Sample Distribution of Trade Discrepancies

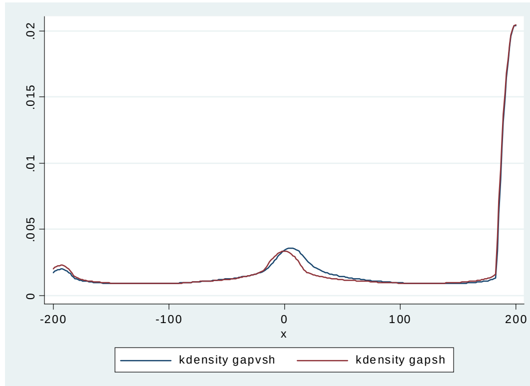

Figure 2: Proportion of Positive Quantity

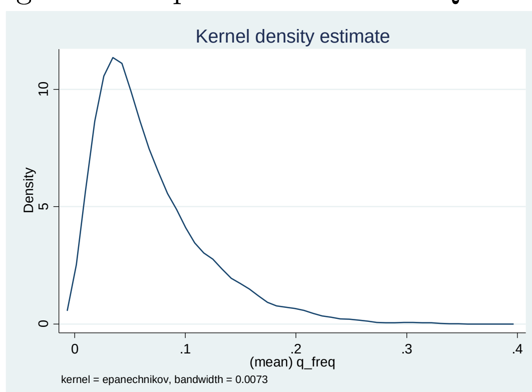

Figure 3: Over Dispersion: Ratio of Variance over Mean of Positive Quantity

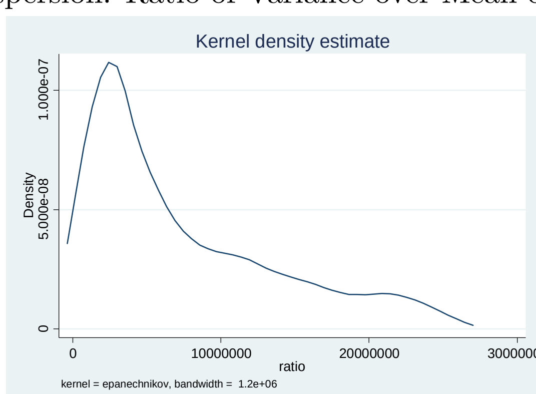

Table 1: Coverage Ratio of Non-Tariff Measures by Importing Countries

Importing Country	Percent	Importing Country	Percent
ARE	99.22	JPN	67.08
ARG	51.38	KAZ	62.80
AUS	79.32	KOR	94.39
BGD	98.20	LKA	89.87
BOL	90.11	MAR	37.30
BRA	88.27	MEX	49.14
CAN	56.54	MYS	72.48
CHE	60.62	NGA	99.00
CHL	57.43	NZL	86.21
CHN	71.86	PAK	56.47
CIV	94.78	PER	89.13
CMR	71.90	PHL	99.01
COL	97.25	PRY	82.37
CRI	59.45	RUS	69.53
CUB	98.03	SAU	38.77
DZA	64.74	SEN	98.36
ECU	75.31	SGP	84.98
ETH	97.87	THA	90.05
EUN	60.98	TUN	46.83
GHA	97.07	TUR	86.34
GTM	86.63	URY	73.84
HKG	84.33	USA	85.44
HND	96.19	VEN	92.22
IDN	78.95	VNM	84.25
IND	79.19		

Data Source: UNCTAD TRAINS. Coverage Ratio measures the percentage of HS 6 digit products that are subjected to at least one border NTMs in the importing country.

Table 2: Second Stage IV Regression, Dependent Variable: Tariffs

	(1)	(2)	(3)
AVEborder	-0.125***	-0.0740***	-0.0874***
	(0.0309)	(0.0137)	(0.0284)
Observations	253,965	248,215	253,818
Importer Fixed Effects	Yes		
Exporter Fixed Effects	Yes	Yes	
Product Fixed Effects	Yes		Yes
Importer-Product Fixed Effects		Yes	
Importer-Exporter Fixed Effects			Yes

Notes: Standard errors are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively;
AVE_{nij} is instrumented using the average AVE of exporting country j of the product n in the non-i markets.

Table 3: Downward Bias Least Squares Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVE of NTM-border	0.403***	0.383***		0.360***	0.348***	
	(0.0663)	(0.0688)		(0.0650)	(0.0675)	
Tariff		0.829***			0.707***	
		(0.0413)			(0.0421)	
AVE of NTM-border+Tariff			0.736***			0.636***
			(0.0346)			(0.0353)
SPS/TBT NTMs			-3.769***			-3.277***
			(0.991)			(0.965)
Other NTMs			-3.999***			-4.022***
			(0.961)			(0.955)
Constant	95.83***	89.31***	93.25***	98.70***	93.09***	96.72***
	(0.326)	(0.464)	(0.760)	(0.319)	(0.465)	(0.749)
Observations	256,711	250,432	250,432	251,934	245,615	245,615
R-squared	0.164	0.166	0.166	0.163	0.164	0.165
Importer Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Exporter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Product Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively.

Table 4: First Stage Instrumental Variable Regressions

Dependent variables	AVE		AVE+Tariff	AVE		AVE+Tariff
	(1)	(2)	(3)	(4)	(5)	(6)
Average AVE of exporter	0.634***	0.608***		0.650***	0.626***	
	(0.030)	(0.031)		(0.030)	(0.024)	
Tariff		0.025***			0.024***	
		(0.004)			(0.004)	
Average AVE of exporter			0.986***			0.982***
+Tariff			(0.018)			(0.018)
SPS/TBT NTMs			6.907***			6.829***
			(0.369)			(0.368)
Other NTMs			-0.321			-0.271
			(0.445)			(0.450)
Importer Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Exporter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Product Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistics	441.13	392.79	2938.74	472.38	419.60	2850.57
Observations	256,711	250,432	153,084	251,934	245,615	150,582

Notes: Robust standard errors in parentheses are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively.

Table 5: Second Stage Instrumental Variable Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVEborder	1.257***	1.456***		1.354***	1.537***	
	(0.218)	(0.234)		(0.208)	(0.222)	
tariff		0.805***			0.682***	
		(0.0427)			(0.0436)	
avetar			0.890***			0.798***
			(0.0525)			(0.0532)
SPS/TBT NTMs			-9.221***			-8.764***
			(1.302)			(1.272)
Other NTMs			-3.621***			-3.598***
			(1.267)			(1.256)
Observations	256,711	250,432	153,084	251,934	245,615	150,582
Importer Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Exporter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Product Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively.

Table 6: Origin Frauds: Downward Bias Least Squares Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVEborder	0.691***	0.638***		0.571***	0.523***	
	(0.104)	(0.107)		(0.102)	(0.105)	
Tariff		2.026***			2.007***	
		(0.0669)			(0.0699)	
AVEborder+Tariff			1.706***			1.670***
			(0.0559)			(0.0576)
SPS/TBT NTMs			-16.02***			-24.97***
			(4.489)			(4.501)
Other NTMs			14.97**			13.84**
			(5.843)			(5.566)
Constant	95.41***	79.59***	83.19***	98.40***	82.63***	91.84***
	(0.185)	(0.544)	(3.301)	(0.182)	(0.570)	(3.124)
Observations	251,184	244,689	244,689	246,373	239,854	239,854
R-squared	0.280	0.284	0.283	0.280	0.284	0.284
Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Product FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively.

Table 7: Origin Frauds: Second Stage IV Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVEborder	1.413***	1.672***		1.367***	1.600***	
	(0.250)	(0.269)		(0.237)	(0.252)	
Tariff		2.007***			1.990***	
		(0.0668)			(0.0698)	
AVEborder + Tariff			2.272***			2.306***
			(0.102)			(0.102)
SPS/TBT NTMs			-18.06**			-29.03***
			(8.601)			(8.398)
Other NTMs			7.159			15.92
			(11.12)			(10.50)
Observations	251,184	244,689	146,609	246,373	239,854	144,111
Exporter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Product Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively.

Table 8: Origin Frauds: First Stage Instrumental Variable Regressions

Dependent variables	AVE		AVE+Tariff		AVE	AVE+Tariff
	(3)	(4)	(5)	(6)	(7)	(8)
Average AVE of exporter	0.597***	0.571***		0.613***	0.591***	
	(0.028)	(0.029)		(0.028)	(0.029)	
Tariff		0.021***			0.019***	
		(0.001)			(0.001)	
Average AVE of exporter			0.899***			0.900***
+Tariff			(0.067)			(0.017)
SPS/TBT NTMs			4.601***			4.300***
			(0.941)			(0.940)
Other NTMs			1.466			1.464
			(1.034)			(0.992)
Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Product FE	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistics	444.17	388.29	2422.37	465.06	409.57	2871.97
Observations	251,184	244,689	146,609	246,373	239,854	144,111

Notes: Robust standard errors in parentheses are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively.

Table 9: Product Classification Frauds: Downward Bias Least Squares Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVEborder	0.342***	0.381***		0.320***	0.360***	
	(0.0654)	(0.0670)		(0.0633)	(0.0649)	
Tariff		0.131***			-0.0141	
		(0.0459)			(0.0476)	
AVEborder+Tariff			0.215***			0.103***
			(0.0363)			(0.0371)
SPS/TBT NTMs			-2.048**			-1.303
			(0.973)			(0.939)
Other NTMs			-3.620***			-3.550***
			(0.937)			(0.925)
Constant	95.90***	94.70***	97.01***	98.73***	98.69***	100.4***
	(0.318)	(0.478)	(0.756)	(0.309)	(0.482)	(0.741)
Observations	256,565	250,285	250,285	251,785	245,466	245,466
R-squared	0.227	0.226	0.226	0.233	0.232	0.232
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively.

Table 10: Product Classification Frauds: Second Stage IV Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVEborder	1.868***	1.977***		1.965***	2.051***	
	(0.223)	(0.239)		(0.213)	(0.227)	
Tariff		0.0973**			-0.0479	
		(0.0471)			(0.0487)	
AVEborder + Tariff			0.260***			0.134**
			(0.0511)			(0.0550)
SPS/TBT NTMs			-3.561***			-2.668**
			(1.223)			(1.190)
Other NTMs			-3.105***			-3.063***
			(1.178)			(1.159)
Observations	256,565	250,285	152,919	251,785	245,466	150,418
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses are clustered by importer-product;
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively.

Table 11: Product Classification Frauds: First Stage Instrumental Variable Regressions

Dependent variables	AVE		AVE+Tariff		AVE	AVE+Tariff
	(3)			(6)		
Average AVE of exporter	0.636***	0.610***		0.650***	0.625***	
	(0.030)	(0.031)		(0.030)	(0.031)	
Tariff		0.023***			0.023***	
		(0.005)			(0.005)	
Average AVE of exporter			0.990***			0.984***
+Tariff			(0.021)			(0.022)
SPS NTMs			6.877***			6.828***
			(0.370)			(0.369)
Other NTMs			-0.285			-0.235
			(0.461)			(0.464)
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistics	445.38	396.27	2147.43	472.86	419.99	2088.07
Observations	256,565	250,285	152,919	251,785	245,466	150,418

Notes: Standard errors are clustered by importer-product
, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively

Table 12: Second Stage IV Regressions for Homogeneous and Differentiated Products, Dependent Variable: Quantity Discrepancies

	Origin Frauds						Product Classification Frauds					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
AVEborder	1.681***	2.030***		2.823***	2.554***		1.993***	2.148***		2.375***	2.228***	
	(0.335)	(0.368)		(0.852)	(0.854)		(0.295)	(0.320)		(0.632)	(0.655)	
tariff		2.143***			1.266***			0.0751			0.246	
		(0.0748)			(0.297)			(0.0588)			(0.154)	
avetar			2.272***			1.951***			0.283***			0.327*
			(0.112)			(0.545)			(0.0689)			(0.189)
SPS/TBT NTM			-8.025			-21.63			-3.652**			2.846
			(11.98)			(16.90)			(1.479)			(5.472)
Other NTM			8.341			48.75**			-6.023***			7.765
			(17.47)			(22.18)			(1.424)			(5.350)
Obs	182,481	177,530	108,125	11,213	10,820	6,350	184,666	179,869	110,887	11,887	11,507	7,042
Fixed Effects:												
Importer-Product	Yes	Yes	Yes	Yes	Yes	Yes						
Exporter	Yes	Yes	Yes	Yes	Yes	Yes						
Importer-Exporter							Yes	Yes	Yes	Yes	Yes	Yes
Product							Yes	Yes	Yes	Yes	Yes	Yes

Notes: Columns (1)-(3) and (7)-(9) consist of trading of differentiated products.
Columns (4)-(6) and (10)-(12) consist of trading of homogeneous products.
Standard errors are clustered by importer-product.
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively

Table 13: Second Stage IV Regressions for High-Income Importing Countries, Dependent Variable: Quantity Discrepancies

	Origin Frauds						Product Classification Frauds					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
AVEborder	1.530***	1.819***		1.409	1.544		1.947***	2.038***		1.427	1.752	
	(0.250)	(0.268)		(1.398)	(1.511)		(0.227)	(0.242)		(1.085)	(1.226)	
tariff		2.029***			-0.224			0.0731			-1.054***	
		(0.0692)			(0.355)			(0.0474)			(0.302)	
avetar			2.321***			0.0342			0.225***			-0.423
			(0.104)			(0.517)			(0.0511)			(0.407)
SPS/TBT NTMs			-19.29**			0			-4.316***			11.90**
			(8.626)			(0)			(1.262)			(5.750)
Other NTMs			5.471			0			-3.145***			-4.569
			(11.19)			(0)			(1.210)			(6.672)
Observations	216,921	210,940	130,235	32,687	32,160	14,580	222,594	216,837	136,785	33,647	33,120	15,574
Fixed Effects												
Importer-Product	Yes	Yes	Yes	Yes	Yes	Yes						
Exporter	Yes	Yes	Yes	Yes	Yes	Yes						
Importer-Exporter							Yes	Yes	Yes	Yes	Yes	Yes
Product							Yes	Yes	Yes	Yes	Yes	Yes

Notes: Unless otherwise specified, samples consist of high-income importing countries only.
Columns (1)-(3) and (7)-(9) consist of trading between high-income importing countries and developing exporting countries.
Columns (4)-(6) and (10)-(12) consist of trading between high-income importing countries and high-income exporting countries.
Standard errors are clustered by importer-product.
*, ** and *** indicate that coefficients are significant at 90%, 95% and 99%, respectively