Trade Frauds, Trade Elasticities and Non-Tariff Measures*

Hiau Looi Kee† Alessandro Nicita‡

World Bank UNCTAD

June 2016

Abstract

This paper studies the role of burdensome non-tariff measures (NTM) in inducing large trade frauds and affecting trade flows. We develop a methodology to estimate bilateral ad valorem equivalent (AVE) of NTMs at detailed product level for a wide range of importing and exporting countries. Results show that tariff and NTMs are substitutes which highlight the importance of including AVEs when estimating trade elasticities. In addition, exporters or products that have higher AVEs tend to have larger trade discrepancies, suggesting firms mis-declare product codes or country of origin to circumvent the cumbersome and opaque NTM.

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

JEL Classification Numbers: F10, F14

*We thank Ana Fernandes, Russell Hillberry, Chad Bown, Aaditya Mattoo, Bob Rijkers, Pol Antras and seminar participants in World Bank for feedback and comments. The results and opinions present in this paper are our own, and do not necessarily represent the views of the World Bank, its 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.

‡UNCTAD. Email: Alessandro.Nicita@unctad.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 (NTM) 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 mis-declared as other commodities upon importation into the United States and trans-shipped through other countries to evade anti-dumping duties."

– World Customs Organization, *Illicit Trade Report* 2012

1 Introduction

In an ideal world, when firms export a product from China to the US, the reported export value and volume of China should match up with the reported import value and volume of US. Unfortunately, this rarely happens in practise. At the aggregate level, the US recorded US$467 billion of imports from China, while China only recorded US$397 billion of exports to the US. That is a trade discrepancy of US$70 billion. In fact, the presence of trade discrepancies are wide spread and is not unique to only the developed or developing countries. In most cases, import statistics are found to be larger than export statistics. There are several possible reasons behind such trade discrepancies. Transshipment and re-exports of products going through third countries could cause large gaps between import data and export data[1]; Different valuation of imports, which normally include freight and insurance costs, could explain why import values are usually larger than the corresponding export values.[2] Finally, fraudulent reporting practise in order to evade tariffs or duties could also cause import

[1] Ferrantino and Wang, 2008.

[2] Hummels and Lugovskyy (2006).

statistics to be smaller than export statistics[3]. What remains puzzling is why would import quantity be significantly larger than export quantity, particularly at detailed product level.

An equally puzzling phenomenon in another area of the empirical trade literature is the wide ranging size of trade elasticities estimates, despite the fact that trade elasticities are essential for the calculation of welfare gains from trade liberalization or tariff reduction. Many researchers have relied on time series or cross sectional variations of trade costs, with estimates from nearly 0 to -10.[4] Not all the estimates use variations in tariffs to estimate the elasticities, let alone other trade policies that potential could impact trade flows and affect the estimates.

This paper studies the role of non-tariff measures (NTM) in affecting trade flows and inducing trade frauds that causes large trade discrepancies. We compare detailed product level trade statistics of a wide range of importing and exporting countries to show that there are systematic and large discrepancies that can only be explained by the presence of burdensome NTMs of the importing countries. NTMs to be considered include Sanitary and Phytosanitary Measures, such as requirements on labelling, hygiene, maximum pesticide residue limits and testing; as well as Technical Barriers to Trade, such as requirements on labelling, product quality, packaging, and certifications. Compliance of these requirements is financially costly and time consuming which increases trade cost and hence potentially reduces trade, much like a tariff. In some cases, when NTMs are very restrictive, it may even undermined the trade liberalization impacts of tariff reductions. High compliance costs also give firms the incentives to misrepresent the country of origin or product codes in order to circumvent the NTMs and hence commit trade frauds.[5]

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

[4] Baier and Bergstrand (1999), Broda and Weinstein (2006), Hummels (2001), Hillberry and Hummels (2013), Simonovska and Waugh (2014).

[5] It should be noted that this paper only focuses on large discrepancies, whereby import value or volume exceeds export value or volume by more than 40% and the reverse, in order to side steps differences in valuations and small statistical errors.

To study the effects of NTMs on trade flows and large trade discrepancies, this paper heeds the message of Goldberg and Pavcnik (2016): "*Measure before you estimate!*", first measures the Ad Valorem Equivalent (AVE) of NTMs, based on UNCTAD-TRAIN database. This paper develops a methodology to estimate these AVEs at product level with bilateral variations, given that compliance costs likely to vary across exporting countries and products, even if the NTMs of the importing countries are not country or product specific. In addition, in order to address the issue of close to 90 percent of country pairs has zero trade at HS 6 digit product level, this paper adopts a zero-inflated negative binomial regression specification, together with other specifications such as zero-inflated Poisson, negative binomial and Poisson regressions in our estimations, guided by model identification tests to select the estimates. We also use instrumental variables to address the endogeneity issues of tariff and NTMs. Our results show that the estimated AVEs are well defined, with a sample mean of 11.5 percent. The results further show that NTMs and tariff are substitutes in our sample suggesting that it is important to separately include NTMs in the studying of trade policies and trade discrepancies. Specifically, exporters that face more restrictive NTMs often have lower tariffs. This result sheds light on the recent papers on the estimations of trade elasticities for the calculation of gains from trade.[6] Omitting NTMs in any trade flows regressions will likely lead to an underestimation of the trade elasticity given that both NTMs and tariffs are negatively correlated with trade flows while simultaneously the two policy variables are themselves negatively correlated.[7]

Using the estimated AVEs, this paper shows that burdensome NTMs are associated with large trade discrepancies. Specifically, controlling for importer-product pair, exporters that face higher AVE are those that have larger trade discrepancies, consistent with the

[6] Hillberry and Hummels, 2013; Arkolakis et at., 2010; Baier and Bergstrand, 2001; Goldberg and Pavcnik (2016).

[7] In a recent paper, Sequeira (forthcoming), shows that trade responded weakly to large tariff liberalization between Mozambic and South Africa due to customs corruptions.

hypothesis of origin frauds, whereby exporters mis-declared their country of origin to gain access to the importer's market. Similarly, controlling for importer-exporter pair, products that have higher AVEs are those that have larger trade discrepancies, consistent with the hypothesis of product frauds, whereby exporters mis-declared the product codes to avoid NTMs compliance. These results are robust to simultaneity and measurement errors in AVEs, which we address with the use of instrumental variables. Overall, a 10 percentage points increase in AVE may cause a six percentage points increase in trade discrepancies, which is about 6% of the average trade discrepancies at the sample mean. Comparatively, a 10 percentage points increase in tariff may only lead to a two percentage points increase in trade discrepancies. This suggests that a significant part of the widespread trade discrepancies observed between countries can be attributed to NTM-induced trade frauds.

To illustrate the effect of NTMs on the estimated trade elasticities, this paper runs gravity regressions at HS 6 digit level, controlling for importer-product and importer-exporter fixed effects. Results show a statistically significant increase in magnitude of the estimated trade elasticity of tariff, from -2.7 to -3. Jointly, when we allow NTMs to be part of the variable trade costs, trade elasticity of trade policy could be as high in magnitude as -6.

This paper is related to several strands of literature. It relates to the literature on trade discrepancies and trade frauds.[8] Unlike the previous papers which mostly focus on a single country imports or exports, this paper uses a detailed trade data from a wide range of importing and exporting countries, which allow us to detect origin frauds and product frauds. Most importantly, previous papers only focus on tariff evasion, while this paper studies the effect of NTM-induced trade frauds, controlling for the effects of tariffs. It should be noted

[8] This literature starts with Bagwati (1964) looking at the trade discrapencies between Turkey and its trading partners. More recent papers include Fisman and Wei (2004) studies how tariff evasion lead to trade discrepancies between Hong Kong and China. Ferrantino and Wang (2008), Ferrantino, Liu and Wang (2012) find strong statistical evidence of under-reporting exports at the Chinese border to avoid paying value-added tax. Javorcik and Narciso (2008) focuses on export of Germany to 10 Eastern European countries and finds evidence that differentiated products have greater tariff evasion.

that, overall, the marginal impact of AVEs on trade discrepancies is larger than that of tariffs. This paper is also related to the literature on estimating the effects of trade policies, particularly non-tariff measures, on trade flows.[9] Similar to the previous papers, our results confirm existing findings that NTMs are an important part of trade policies and should not be overlooked when evaluating the impact of any trade agreements. Unlike the previous papers, our AVE estimates are highly detailed and disaggregated, at HS 6 digit level that varies by importers and exporters. Such detailed variations is very useful for future Computational General Equilibrium (CGE) models to evaluate the impact of trade liberalizations. It is also more informative for policy makers and trade negotiators who would like to know which products and which trade partners to focus on for trade negotiations. This paper further relates to the literature on evaluating trade elasticities and gains from trade.[10] The negative relationship between NTMs and tariffs found in this paper provides an explanation to the existing low trade elasticities estimates of the previous papers and our estimated AVEs could be used jointly with tariff to evaluate the welfare gains from trade. Finally, this paper uses novel econometric techniques in the estimation of AVEs which made it related to the literature on the estimation of gravity models.[11] Unlike the previous papers which used aggregate bilateral import value as the dependent variables, our dependent variable for the AVE estimations is product level bilateral import quantity, which is a count variable with a large presence of zeros and huge overdispersion. This makes zero-inflated negative binomial model more suitable for our current application. We also address the endogeneity of tariff and NTMs in the estimation of AVEs with valid instruments.

The paper proceeds as follows. Section 2 provides a simple model to fix ideas and define

[9] Lee and Swegal (1997), Broda and Weinstein (2006), Kee, Nicita and Olarreaga (2008, 2009) and Kee and Nicita (2013).

[10] Baier and Bergstrand (1999), Arkolaskis et al. (2010), Caliendo and Parro (2015), Goldberg and Pavcnik (2016).

[11] Santos Silva and Tenreyro (2006, 2011), Martin and Cong (2015).

trade discrepancies. Section 3 provides a data description for the paper. Section 4 presents the empirical model to estimate AVEs of NTMs. Section 5 presents the results on the estimation of AVEs. Section 6 shows the empirical evidence on NTM-induced trade frauds, while Section 7 illustrates the bias in trade elasticities when NTMs are omitted. Section 8 concludes the paper.

2 Theoretical Model

In this section, we describe a simple model to show how the differences between imports and exports statistics could be related to tariff and non-tariff measures.

The observable country $i's$ import of product k from country j, is denoted as M_{ij}^{k}, which equals to the observable country $j's$ export of product k to country i, E_{ij}^{k}, and an error term, ε_{ij}^{k} :

$$M_{ij}^{k}\left(t_{ij}^{k}\right)=E_{ij}^{k}\left(t_{ij}^{k}\right)+\varepsilon_{ij}^{k}, \tag{1}$$

with t_{ij}^{k} being the trade policies of country $i's$ import of product k from country j. Restrictive tariff and NTMs will both decrease imports and exports, such at

$$M_{ij}^{k'}\left(t_{ij}^{k}\right)<0 \text{ and } E_{ij}^{k'}\left(t_{ij}^{k}\right)<0.$$

If the error term is purely idiosyncratic and random, we would expect $E\left(\varepsilon_{ij}^{k}\right)=0$. This will be the case when

$$M_{ij}^{k'}\left(t_{ij}^{k}\right)=E_{ij}^{k'}\left(t_{ij}^{k}\right),$$

that is when trade policy changes induce the same responses in imports and exports. Thus even if $\varepsilon_{ij}^{k}\neq 0$ for some observations, on average, $E\left(\varepsilon_{ij}^{k}\right)=0$ if ε_{ij}^{k} is truly random.

However, in this paper, we postulate that trade policy changes may not induce the same responses in imports and exports

the error term is increasing with tariff and non-tariff measures, t_{ij}^t, such that

$$E\left(\varepsilon_{ij}^k t_{ij}^t\right) > 0.$$

In other words, when face with a higher tariff or more restrictive non-tariff measures, exporting firms have incentives to misrepresent or underreport their exports and caused a wedge between M_{ij}^k and E_{ij}^k.

When there is no fraudulent behavior, we expect

$$E\left(\varepsilon_{ij}^k t_{ij}^t\right) = 0.$$

Equation (1) shows that the discrepancy between imports and exports depends on an error term that could be determined by trade policies:

$$GAP_{ij}^k\left(t_{ij}^k\right) \equiv M_{ij}^k\left(t_{ij}^k\right) - E_{ij}^k\left(t_{ij}^k\right) = \varepsilon_{ij}^k\left(t_{ij}^k\right).$$

In the empirical section, we will measure $GAP_{ij}^k\left(t_{ij}^k\right)$ relative to the average size of trade, in minimize heteroskedasticity and the influence of large outliers:

$$Discrepancy_{ij}^k = \frac{M_{ij}^k\left(t_{ij}^k\right) - E_{ij}^k\left(t_{ij}^k\right)}{0.5 * \left(M_{ij}^k\left(t_{ij}^k\right) + E_{ij}^k\left(t_{ij}^k\right)\right)} \tag{2}$$

3 Data

Data used in this study come from UN-Comtrade for HS 6 digit level import and export data, and UNCTAD TRAINS for tariff and NTM data. Most of the NTM data were collected around 2011, so that determine the year coverage of the sample used in the analysis.

Trade discrepancies is constructed by matching import data with export data at HS 6 digit level according the equation (2) , for both quantity and value. Quantity data have been thoroughly examined and we only include observations when that both importer and exporter are using the same quantity unit in their reporting, to make sure that the discrepancies are not due to differences in quantity units. In addition, quantity units remain the same for any given HS 6 digit product. For the regression analysis below, we always control for product fixed effects so that we are not relying on differences in quantity units across product to estimate the coefficients of interests. Figure 1 presents the sample distribution of trade discrepancies. At the sample mean, trade discrepancies is about 85 percent suggesting that most import records are larger than export records for both quantity and value.

Table 1 presents the list of importing countries included in our sample, together with their coverage ratios of the two types of NTMs, namely the SPS/TBT measures and the non-SPS/non-TBT measures. There are a total of 34 importing countries with 96 exporting countries. Sample is largely determined by the availability of NTM data in TRAINS. There are a wide range of variation in terms of coverage ratios across these different importing countries. Some countries, such as Sri Lanka and Nepal, impose at least one NTMs on all products from all exporting countries, while other countries, such as Peru, only impose at least one NTM on less than 3 percent of its imported products. We exploit these wide range of variation across countries to estimate the AVEs at HS 6 digit level. It should be noted that in the UNCTAD focused mostly on SPS/TBT measures in this round of NTM data collection. So even though both types of NTMs are collected, the quality of data on SPS/TBT is higher. We will focus primarily on the SPS/TBT measures in our estimation of AVEs and will include the other type of NTM as a control variable in our regressions.

4 Estimation of the Ad Valorem Equivalent of NTMs

To estimate the ad valorem equivalent of NTM, we modified Kee, Nicita and Olarreaga (2009), to run the separate quantity-based gravity regressions for each of the nearly 5000 HS 6 digit level for 2011, based on cross-country bilateral import quantity data of 34 importing countries and 96 exporting countries. Moreover, given that compliance costs of NTMs could depend on the exporting country even if the NTMs are mostly multi-lateral in nature, we use interaction terms to obtain bilateral AVE estimates.

4.1 Econometrics Issue I: Large Presence of Zeros and Over-Dispersion

To estimate the bilateral AVE of NTMs at HS 6 digit level, we first need to estimate how much trade is reduced due to the presence of NTMs in a gravity-type regression at HS 6 digit level, based on bilateral trade flow between country pairs. To convert the estimated coefficients to AVE, we will also estimate the import demand elasticity, which is the percentage of trade reduced due to the presence of tariff, and use the estimated import demand elasticity to "re-scale" the NTM coefficient to obtain AVE. For this purpose, we use quantity imported as the dependent variable. However, unlike Kee, Nicita and Olarreaga (2009), we treat quantity imported as a discrete, non-continuous variable that is count in nature and use count data regression technique for our estimations. This is necessary, because nearly 25 percent of the observations has "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 has "Weight in kilograms" as the unit of measurement, which is often recorded as a discrete variable. As such, nearly 100 percent of our observations recorded imported quantity as discrete integers. So it is appropriate to use count data regression technique, such as Poisson or Negative Binomial models for our estimation. Recent papers of Santos Silva and Tenreyro (2006,

2011) have shown that 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.

However, in our current data set, unlike Santos Silva and Tenreyro (2006, 2011), we are focusing on HS 6 digit bilateral trade between country pair. For each HS 6 digit product, there is an excessive number of bilateral country pair that do not trade. In fact, the average share of positive quantity imported is only 6.8% in the sample. This leads to a large presence of zero observations in the bilateral data set as shown in Figure 2. In this regards, the zero inflated models could be more appropriate than traditional count regression models, such as PPML, for our regressions. Moreover, for country-pair that do trade, the volume of their trade have a very large range which leads to extreme over-dispersion (large ratio of variance over mean) as shown in Figure 3. The ratio of the variance of positive quantity and its mean 4.83e+08 in our data set. For this, negative binomial model could be more appropriate than Poisson model, since Poisson model do not allow for over-dispersion. The large presence of zero's together with over-dispersion led us to use Zero-Inflated Negative Binomial (ZINB) Model for our main estimation specification. Nevertheless, for each HS 6 digit product, in addition to ZINB, we also run Negative Binomial (NB), Zero-inflated Poisson (ZIP), OLS in log and Poisson Models, and will use model specification test to picks the best fitted regressions for our AVE estimates.

We assume that the observe 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, and θ includes both α and β. Here $\mathbf{X}$ is a vector of variables to explain positive import quantity, which are standard gravity variables, $\mathbf{Z}_{ij}$, as well as tariff and NTMs:

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

So the import quantity density distribution can be model as

$$pr\left(k\right) = \begin{cases} p_{nij} + \left(1 - p_{nij}\right) h\left(Q_{nij} = k, \theta | \mathbf{X}\right), \text{ if } k = 0 \\ \left(1 - p_{nij}\right) h\left(Q_{nij} = k, \theta | \mathbf{X}\right), \text{ if } k = 1, 2, ... \end{cases} \tag{3}$$

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

$$p_{nij} = \frac{\exp\left(\mathbf{W}\gamma\right)}{1 + \exp\left(\mathbf{W}\gamma\right)} \in \left(0, 1\right). \tag{4}$$

In general, same set of variables could be included in both $\mathbf{W}$ and $\mathbf{X}$ and we do not need to identify additional variable for the participation equation (see Cameron and Trivedi, 1998).

4.2 Econometric Issue II: Endogeneity

Tariff and NTMs are set by the importers and could be endogenous to trading volume. Including tariff and NTMs in the regressions will lead to inconsistent estimates. To address this endogeneity issue, we use the tariff and NTMs presence of the 3 closest neighboring countries as instruments. Trade policy of neighboring countries could be correlated with the trade policy of the importing country due to regional trading agreement or common cultural/history background, but trade policy of neighboring countries should not affect bilateral imports of the own country and hence satisfy the exclusion restrictions for instrumental variables. For tariff, we use the fitted value of the first stage IV regression in the ZINB regressions. For the SPS/TBT NTMs, we ran a probit (selection) regression for the NTM variable and include the inverse Mill's ratios in the second stage ZINB regressions (because NTM variables are discrete).

4.3 Econometric Issue III: Bilateral Coefficients

While NTMs are mostly multilateral and not exporting country specific, the impact of NTMs on bilateral trade flows are likely to vary across exporting countries, due to compliance costs and other importing, exporting country specific factors. To capture this in the estimation of AVEs, we interact our NTM variable with variables that are specific to importing countries and exporting countries. In addition, we also allow import demand elasticity, which depends on the coefficient on tariff to have bilateral variations. To this end, we interact tariff and SPS/TBT measure with the following two variables: the share of the exporter in the world trade of this HS 6 product, and the share of the importer in the world trade of this HS 6 product. Both variables capture the market powers of the exporter and importer in the world market for each HS 6 product. We expect compliance cost to be lower if the exporter has a larger share in the world trade of the product, which results in smaller trade impact due to the presence of importing country's SPS/TBT measures. However, it is also possible that the exporter could easily divert their export to other market when face with 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 market share of a product, it could be more difficult to comply with its SPS/TBT measures and the compliance cost for exporting countries could be higher, and that lead to a larger trade impact. However, it is also possible that if the market share of the importer is large, it is more difficult for exporting countries to divert their products to other markets, so the trade impact of SPS/TBT measures could be lower. The overall trade effects of the market shares of importer and exporter in the world market due to the presence of SPS/TBT depends on the specific products and country-pair and we will let the data reveal the dominating force. Similar argument can be made for tariff and the resulting bilateral import elasticities.

4.4 Overall Estimation Procedure and the Construction of AVE

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

4.4.1 First Stage Regressions

For the first stage regression for tariff, we use the average tariff of 3 closest countries, $\bar{t}_{nij}$, as an instrument for tariff, t_{nij}, to get fitted tariff, $\hat{t}_{nij}$:

$$\hat{t}_{nij} = \mathbf{Z}_{ij}\hat{\nu} + \hat{\nu}^t \bar{t}_{nij}, \tag{5}$$

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

For the first stage regression for NTM, because NTM is a dummy variable that equals one when there is at least one NTM presence, the first stage regression is a probit regression with density function, $f(.)$, which we use the average presence of NTM in the 3 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 Mill ratio, $invMill_{nij}$, constructed based on $\hat{\delta}$ and $\hat{\delta}^{NTM}$, the coefficients of the second stage control variables, $\mathbf{Z}_{ij}$ and the average NTM presence, $\overline{NTM}_{nij}$.

4.4.2 Second Stage Regressions

The main specification of second stage regression for zero-inflated negative binomial model (ZINB), negative binomial model (NB), zero inflated Poisson model (ZIP), Poisson model is

$$
\begin{aligned}
\ln E\left(Q_{nij}|\mathbf{X}\right) &= \beta_n+\beta^t_{nij}\hat{t}_{nij}+\beta^{NTM}_{nij}NTM_{nij}+\beta^M invMill_{nij}+\mathbf{Z}_{ij}\beta+\varepsilon_{nij}, \text{ with} \quad (6) \\
\beta^t_{nij} &= \beta^t_n+\beta^t_1 share_{ni}+\beta^t_2 share_{nj}, \\
\beta^{NTM}_{nij} &= \beta^{NTM}_n+\beta^{NTM}_1 share_{ni}+\beta^{NTM}_2 share_{nj}.
\end{aligned}
$$

Control variables included in $\mathbf{Z}_{ij}$ are the standard gravity variables: the log of GDP of importer and exporter, bilateral distance between importer and exporter, landlocked indicators for importer and exporter, and common border indicator. Bilateral coefficients of tariff and NTMs, β^t_{nij} and β^{NTM}_{nij}, are obtained by using the interaction terms based on the share of importer in the world market, $share_{ni}$ and the share of the exporter in the world market, $share_{nj}$. For log-OLS, the second stage regression is similar to (6) but with the log of quantity imported, $\ln\left(Q_{nij}\right)$, as the dependent variable.

For the zero-inflated specifications, such as ZINB and ZIP, we also include a logit regression to explain the presence of excessive zeros in the bilateral trade. While 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, $\mathbf{Z}_{ij}$ to form $\mathbf{W}$ in equation (4).

4.4.3 Construction of AVEs

Once all five versions of Equation (6) are estimated, we use Vuong tests to select the best fit models to retrieve $\hat{\beta}^{t}_{nij}$ and $\hat{\beta}^{NTM}_{nij}$ for the construction of AVE. For all five models, $\hat{\beta}^{t}_{nij}$ equals the log difference in quantity imported (expected quantity for count models) if tariff increases by 1 percentage point, while $\hat{\beta}^{NTM}_{nij}$ equals the log difference in quantity imported (expected quantity for count models) if NTM increases from 0 to 1. The proportionate change in quantity imported (or expected quantity imported for count models) due to NTM is

$$\begin{aligned}
\ln E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=1\right) - \ln E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=0\right) &= \hat{\beta}^{NTM}_{nij} ==> \qquad (7)\\
\ln\left(\frac{E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=1\right)}{E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=0\right)}\right) &= \hat{\beta}^{NTM}_{nij}\\
\frac{E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=1\right)}{E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=0\right)} &= \exp\left(\hat{\beta}^{NTM}_{nij}\right)\\
\frac{E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=1\right) - E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=0\right)}{E\left(Q_{nij}|\mathbf{X}, NTM_{nij}=0\right)} &= \exp\left(\hat{\beta}^{NTM}_{nij}\right) - 1.
\end{aligned}$$

Likewise, the proportionate change in quantity imported (or expected quantity imported for count models) due to 1 percentage point increase in tariff is

$$\begin{aligned}
\ln E\left(Q_{nij}|\mathbf{X},t_{nij}=t+1\right)-\ln E\left(Q_{nij}|\mathbf{X},t_{nij}=t\right) &= \hat{\beta}_{nij}^{t} ==> \\
\ln\left(\frac{E\left(Q_{nij}|\mathbf{X},t_{nij}=t+1\right)}{E\left(Q_{nij}|\mathbf{X},t_{nij}=t\right)}\right) &= \hat{\beta}_{nij}^{t} \\
\frac{E\left(Q_{nij}|\mathbf{X},t_{nij}=t+1\right)}{E\left(Q_{nij}|\mathbf{X},t_{nij}=t\right)} &= \exp\left(\hat{\beta}_{nij}^{t}\right) \\
\frac{E\left(Q_{nij}|\mathbf{X},t_{nij}=t+1\right)-E\left(Q_{nij}|\mathbf{X},t_{nij}=t\right)}{E\left(Q_{nij}|\mathbf{X},t_{nij}=t\right)} &= \exp\left(\hat{\beta}_{nij}^{t}\right)-1.
\end{aligned}$$

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 quantity imported as the presence of NTM_{nij}, or*[12]

$$AVE_{nij}=\frac{\exp\left(\hat{\beta}_{nij}^{NTM}\right)-1}{\exp\left(\hat{\beta}_{nij}^{t}\right)-1}\cong\frac{\hat{\beta}_{nij}^{NTM}}{\hat{\beta}_{nij}^{t}} \text{ for small } \hat{\beta}_{nij}^{NTM},\hat{\beta}_{nij}^{t}. \tag{8}$$

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

5 Results on AVE Estimations

Figures 4 and 5 present the distribution of the first stage F-stats for tariff and NTMs. It is clear that these IVs work very well, with a sample mean of 1718.692 for tariff and 6243.708

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

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

for NTM, indicating these are not weak instruments.

For second stage regressions, overall, ZINB works best for 62 percent of the products, NB works best for 25 percent of the products, and the remaining 6 percent are using ZIP; 3 percent for OLS and only one percent for Poisson regressions. Figure 6 presents the distribution of the estimated bilateral AVEs. At sample mean, when SPS/TBT measures are present, the average AVE is 11.5%, whereas the mean tariff is 7.1%.

Moreover, NTMs and tariff appear to be gross substitutes in our sample. We pool all AVE estimates together and regress them on the bilateral tariffs of the country-pairs. Results are presented in Table 2. Column (1) shows that controlling for importer-exporter, importer-product and exporter-product fixed effects, lower bilateral tariffs are associated with higher AVEs. Column (2) further shows that when compare across exporting countries, controlling for importer-product and exporter fixed effects, exporters that face more restrictive NTMs often has lower tariffs. These results suggest that as a policy instrument, tariff and NTMs are substitutes and should be considered jointly when we want to evaluate the impact of trade policy on any economic outcomes.

6 Results on NTMs Induced Trade Frauds

6.1 Least Squares Downward Bias and Instrumental Variable Estimations

Table 3 presents the least squares results when we regress trade discrepancies on NTMs and tariffs. Dependent variables for Columns (1) - (3) are the quantity discrepancies, while the dependent variables for Columns (4)-(6) are the value discrepancies. Both set of discrepancies are constructed according to equation (2) . 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 the full sets of fixed effects: importer, exporter and product, and we cluster the standard error by importer-product, which is the level of aggregation of most NTM variables. Clustering is particularly necessary, since AVEs are estimated with errors.

Columns (1) and (4) regress trade discrepancies on the estimated AVEs of SPS/TBT measures. Columns (2) and (5) include bilateral tariff in the regressions. Columns (3) and (6) sum up AVE and tariff together as one variable, and included the presence of the other NTMs as a control variable in the regressions. In all specifications, the least squares estimated coefficients on AVEs are very small and not statistically significant.

However, there are reasons to believe that these least squares estimates for AVEs are bias. First, larger trade frauds could be a result of corrupt customs officers and procedures, which may also cause less restrictive NTMs due to the lack of enforcements. Such negative simultaneity between trade discrepancies and AVE will bias the least squares estimates towards zero. Moreover, a large part of trade discrepancies could be due to idiosyncratic and random recording errors and not NTM-induced trade frauds which make it harder for us to find any systematic relationship between trade discrepancies and NTMs, and will further push the least squares estimates towards zero. Finally, there may be a concern that the AVE 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 AVE 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, $\overline{AVE}_{n\tilde{}ij}$, 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 the similar SPS/TBT measures in the other markets. So 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.[13]

Table 4 presents the first stage regression results for the IV estimations. Similar to the previous table, Columns (1)-(3) are for the quantity discrepancies, Columns (4)-(6) are for the value discrepancies, and standard errors are clustered by importer-product. Across all specifications, the coefficients on the average AVE of exporting country, $\overline{AVE}_{n^{\sim}ij}$, is positive and statistically significant. Together with the high first stage F-statistics, confirming that $\overline{AVE}_{n^{\sim}ij}$ is a valid instrument.

Table 5 presents the second stage regression results. Similar to Table 3, dependent variables for Columns (1) - (3) are the quantity discrepancies, while the dependent variables for Columns (4)-(6) are the value discrepancies, and we cluster the standard errors by importer-product. Compare to Table 3, the results in this table is much stronger and better, confirming the downward bias in the previous least squares estimates. The coefficients for AVE are consistently larger and statistically significant across all columns. Overall, a 10 percentage points increase in AVE could lead to a 6 percentage points increase in trade discrepancies, while a 10 percentage points increase in tariff only lead to a 2 percent increase in trade discrepancies. This suggests that a significant part of the trade discrepancies are in fact NTM-induced trade frauds.

6.2 Sources of Trade Frauds: Origin Frauds and Product Frauds

What can explain the positive relationship between trade discrepancies and AVEs? It could be that firms mis-declared the country of origin of the products in order to avoid the burdensome NTMs. This would be consistent with origin frauds. If this is the case, we will 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

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

6 and 7 test the origin frauds 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. As such we rely on the variation of AVE of exporters in the same importing market with the same product to identify the coefficient of AVE. Similar to the previous tables, dependent variables for Columns (1) - (3) are the quantity discrepancies, while the dependent variables for Columns (4)-(6) are the value discrepancies, and we cluster the standard errors by importer-product. Both LS and IV results show that comparing among exporters within importer-product pair, exporters that face more burdensome AVEs are those that have larger trade discrepancies with the importers. This is consistent with origin frauds, suggesting that firms from exporting countries with high AVEs mis-declared their country of origin as countries with low AVEs and caused the large discrepancies.

Another type of trade frauds is when products crossing borders are mis-declared as other products in order to avoid facing burdensome NTMs. If this type of trade frauds exists, we will expect that, 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 frauds 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 of AVE across products within importer-exporter pair to identify the coefficient on AVE. Similar to the previous tables, dependent variables for Columns (1) - (3) are the quantity discrepancies, while the dependent variables for Columns (4)-(6) are the value discrepancies, and we cluster the standard errors by importer-product. Both tables show that comparing among products within importer-exporter pair, products that face more burdensome AVEs are those that have larger trade discrepancies. This is consistent with product frauds, suggesting that firms mis-declared products with high AVEs with products with low AVEs and caused the large discrepancies.

Overall, evidence presented in Tables 3 to 10 are consistent with the NTM-induced trade frauds, with firms either mis-declared their country of origin or their product codes in order to avoid the burdensome NTMs of the importing countries. Based on the results of Table 5, one standard deviation increase in AVE could lead to 3%~6% increase in trade discrepancies. This is not a very large impact in terms of size, nevertheless it is statistically different from zero, indicating that part of the widespread and large trade discrepancies can be attributed to firms deliberately misrepresent themselves to avoid burdensome NTMs which lead to trade frauds.

7 Trade Elasticities and NTMs

In this section, we will demonstrate that omitting the AVE of NTMs from a trade flows regression may lead to downward bias in the trade elasticities of tariff, given that AVEs and tariffs are negatively correlated and both variables are themselves negatively correlated with trade. A few caveats before we proceed. First, unlike most papers in this area, our current data set is based on bilateral trade flows at HS 6 digit level for a wide range of importers and exporting. The estimated trade elasticities at this disaggregate level may be significantly different from any estimates based on aggregate trade flows which are used for welfare gains calculations. Second, we only focus on the positive trade flows and are ignoring the zeros in trade flows. Third, we rely on fixed effects to control for variables that are specific to country-pairs and do not include all the country specific geography variables in the regressions.

Table 12 presents the results of some variations of the following regression:

$$\ln\left(M_{nij}\right) = \beta_{ij} + \beta_{ni} + \beta_t \ln(1 + t_{nij}) + \beta_{AVE} \ln\left(1 + AVE_{nij}\right) + \nu_{nij}. \quad (9)$$

In equation (9), M_{nij} is the value of imports of product n of importing country i from exporting country j, t_{nij} is the applied bilateral tariff of product n in i from j, and AVE_{nij} is the ad valorem equivalent of NTM of product n in i from j. In one of the specifications in Table 12 we also add tariff with AVE to form a new regressor, $\ln(1 + (t_{nij} + AVE_{nij})/100)$. In all regressions, we control for importer-exporter fixed effect, β_{ij}, which will pick up multilateral resistant between i and j, and any other country specific geography variables, such as bilateral distance, language, culture, border, landlocked and institutions/colonial history. We also control for importer-product fixed effect, β_{ni}, which will absorb factors such as preference, market competition, market structure and MFN tariffs. All standard errors are clustered by importer-product.

Column (1) only includes all the fixed effects and t_{nij} in the regression. The coefficient on tariff is -2.7 and is statistically significant. Column (2) adds AVE_{nij} in the regression, which is instrumented with $\overline{AVE}_{n\tilde{\ }ij}$. Now the coefficient on tariff increases in magnitude to -3.0, which is statistically larger than the previous estimate based on the robust standard error. This illustrates that omitting NTMs in the estimation of trade elasticities may lead to estimates that are bias towards zero, given that tariffs and AVEs are negatively correlated and both policy variables are themselves negatively correlated with trade flows. Columns (3) and (4) allow AVEs to be part of the variable trade costs in the estimation of trade elasticity. Column (3) presents the second stage result when we add tariff and AVE together to form a new trade policy variable, while Column (4) imposes the restriction that trade elasticity from tariff and AVE must be the same. In both cases, the estimated trade elasticity is around

-6.6. All the first stage results have the right signs and high F-statistics and are available upon requests.

Overall, these results highlight that it is important to consider the AVEs of NTMs as part of trade policies, which along side with tariff, have significant impacts in deterring trade.

8 Concluding Remarks

This paper studies the wide spread and large discrepancies of detailed product trading between importing and exporting countries. We attribute part of the discrepancies to trade frauds whereby firms deliberately mis-declared their country of origin or product code in order to avoid the burdensome non-tariff measures, such as certifications and inspection requirements of the importing countries. Our approach utilized a detailed product level trade statistics of a wide range of importing and exporting countries. We first estimate the ad valorem equivalents of the non-tariff measures of the importing countries, with variations at product and exporting country level. We then relate the estimated ad valorem equivalents to the existing trade discrepancies, both in terms of quantity and value discrepancies. We found that exporters or products that have larger trade discrepancies tend to face more restrictive non-tariff measures, suggesting mis-declaration of product codes or misrepresentation of the country of origin, in order to circumvent the cumbersome and opaque non-tariff measures.

Our ad valorem equivalent estimates at product level with bilateral variations for a wide range of countries fill an important void in the trade policy literature, which so far has been focusing on tariffs. Given that tariffs have been steadily declining for the past few decades, recent trade negotiations and agreements are mostly focusing on non-tariff measures. Our estimates could inform future works on evaluating the impacts of trade agreements. In addition, our finding that AVEs and tariffs are substitutes suggesting that it is important to include NTMs in the estimation of trade elasticities and the evaluation of gains from trade. Lastly, this paper presents a novel methodology to estimate the ad valorem equivalent of NTMs. With new and more detailed data being systematically collected and made available, the methodology developed in this paper could be refined to estimate the AVEs for subcategories of NTMs, which will be more useful for policy makers and trade negotiators.

References

[1] Arkolakis, Costas, Arnaud Costinot and Andre Rodriguez-Clare (2010). "New Trade Models, Same Old Gains?" *American Economic Review* 102, issue 1, 94-130.

[2] Baier, Scott and Jeffrey Bergstrand (2001). "The Growth of World Trade: Tariffs, Transport Costs and Income Similarities," *Journal of International Economics* 53:1, 1-27.

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

[4] Broda, Christian and David Weinstein (2006). "Globalization and the Gains from Variety," *Quarterly Journal of Economics* 121 (2), 541-585.

[5] Caliendo, Lorenzo and Fernando Parro (2014). "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies* 82 (1), 1-44.

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

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

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

[9] Fisman, Raymond and Shang-Jin Wei (2004). "Tax Rates and Tax Evasion: Evidence from "Missing" Imports in China," *Journal of Political Economy*, vol. 112, no. 2.

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

[11] Helpman, Elhanan, Marc Melitz and Yona Rubinstein (2008). "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics* 123(2), 441-487.

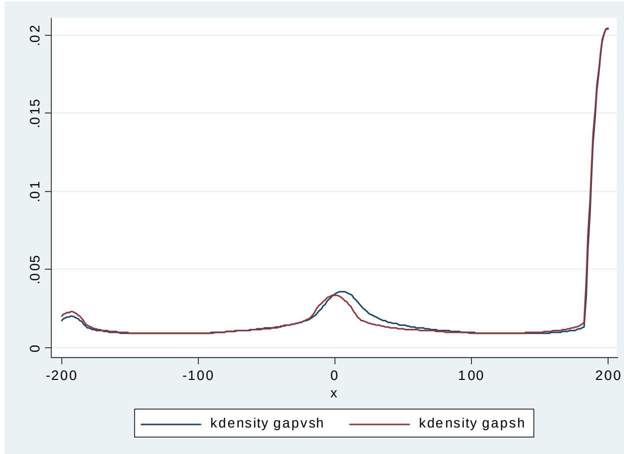
[12] Hillberry, Russell and David Hummels (2013). "Trade elasticity parameters for a computable general equilibrium model." *Handbook of Computable General Equilibrium Modeling* 1, 1213-1269.

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

[14] Javorcik, Beata and Gaia Narciso (2008). "Differentiated Products and Evasion of Import Tariffs," *Journal of International Economics* 76, 208-222.

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

Figure 1: Sample Distribution of Trade Discrepancies



[16] Martin,William J. & Pham,Cong S., (2015). "Estimating the gravity model when zero trade flows are frequent and economically determined," Policy Research Working Paper Series 7308, The World Bank.

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

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

[19] Sequeira, Sandra (forthcoming). "Corruption, Trade Costs and Gains from Tariff Liberalization: Evidence from Southern Africa," *American Economic Review*.

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

Figure 2: Proportion of Positive Quantity

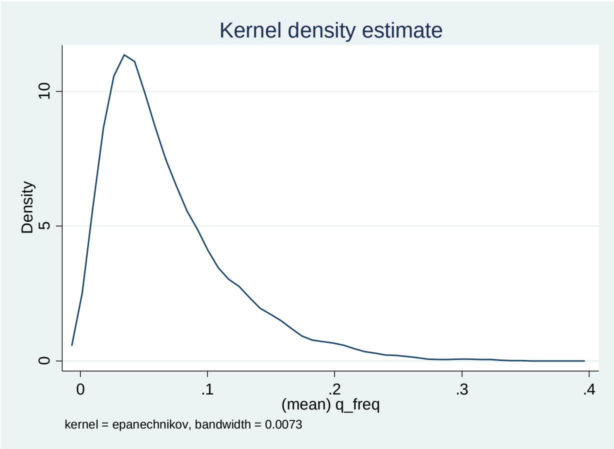


Figure 3: Over Dispersion: Ratio of Variance over Mean of Positive Quantity

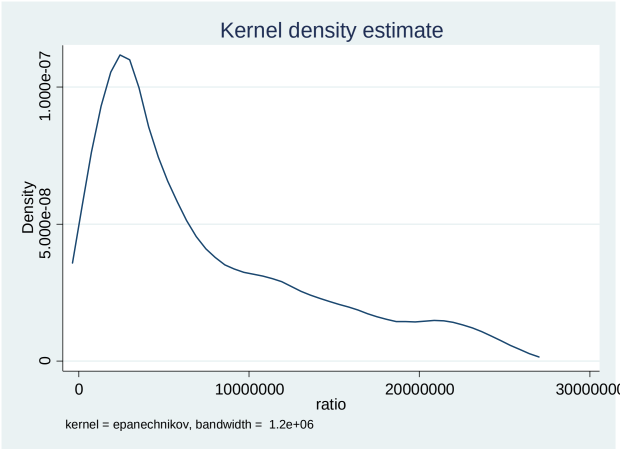


Figure 4: Sample Distribution of First-Stage F-Statistics for Tariff

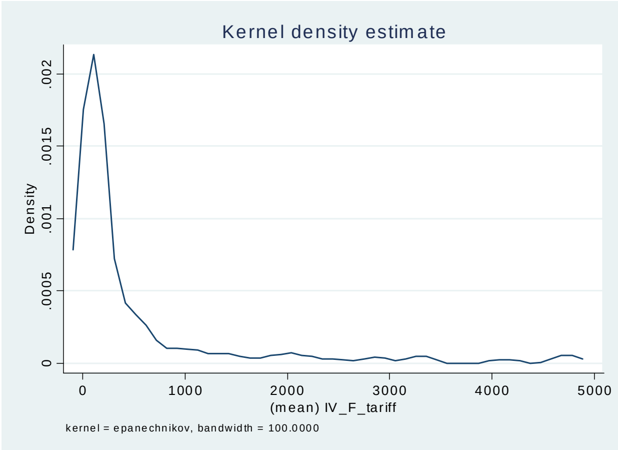


Figure 5: Sample Distribution of First-Stage F-Statistics for SPS/TBT NTMs.

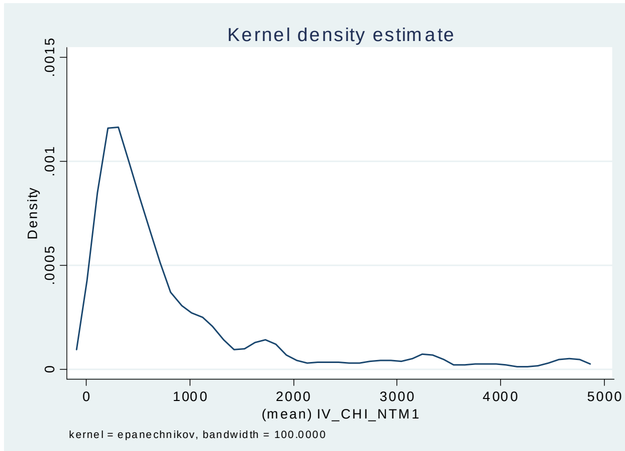


Figure 6: Distribution of the Estimated AVEs for SPS/TBT Measures

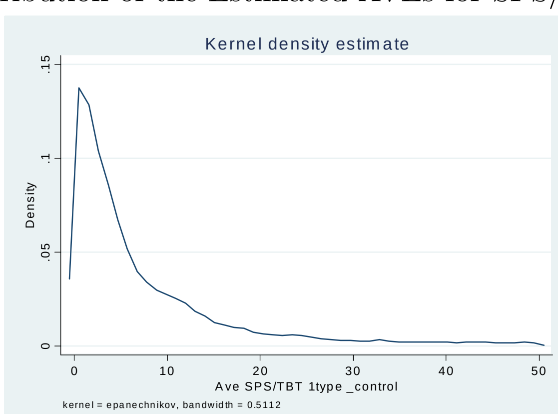


Table 1: Coverage Ratio of Non-Tariff Measures by Importing Countries

Importing Country	SPS/TBT	Others
AFG	100.00	46.85
ARG	47.47	89.44
BOL	2.09	2.35
BRA	55.78	8.34
CHL	33.43	7.63
COL	3.03	4.65
CRI	19.00	25.93
ECU	25.62	27.57
EGY	92.55	47.03
EUN	87.36	13.27
GTM	18.59	95.97
IDN	28.48	24.66
IND	63.01	99.99
JPN	36.86	7.86
KAZ	88.22	75.52
KEN	86.13	55.61
KHM	14.03	24.55
LBN	7.73	1.17
LKA	100.00	100.00
MDG	66.24	22.06
MEX	47.83	33.62
MUS	19.90	15.78
NAM	50.97	51.05
NPL	100.00	100.00
PAK	29.83	99.86
PER	2.58	0.83
PRY	23.28	11.89
SEN	17.42	18.30
SLV	25.59	95.24
SYR	32.97	87.40
TUN	18.59	19.62
TZA	37.16	2.70
UGA	92.32	81.32
URY	3.78	0.29

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

Table 2: NTMs and Tariff are Substitutes

Dependent Variable	Estimated Ad Valorem Equivalent of NTMs	
	(1)	(2)
Tariff	-0.034**	-0.075***
	(0.017)	(0.014)
Fixed Effects		
Importer-exporter	Yes	No
Importer-product	Yes	Yes
Exporter-product	Yes	No
Exporter		Yes
Observations	652,226	707,217

Notes: Robust standard errors are clustered by importer-product;
* and *** indicate that coefficients are significant at 95% and 99% levels.

Table 3: Downward Bias Least Squares Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVE	0.006	0.005		0..003	0.002	
	(0.005)	(0.005)		(0.005)	(0.005)	
Tariff		0.204***			0.193***	
		(0.067)			(0.060)	
AVE+Tariff			0.024***			0.020***
			(0.005)			(0.005)
Other NTMs			-2.892***			-1.570**
			(0.662)			(0.654)
Fixed Effects						
Importer	Yes	Yes	Yes	Yes	Yes	Yes
Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Product	Yes	Yes	Yes	Yes	Yes	Yes
Observations	583,929	575,823	575,823	571,778	563,671	563,671

Notes: Standard errors 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	0.082***	0.081***		0.087***	0.086***	
	(0.021)	(0.022)		(0.023)	(0.023)	
Tariff		0.011*			0.012*	
		(0.007)			(0.006)	
Average AVE of exporter			0.407***			0.421***
+Tariff			(0.073)			(0.074)
Other NTMs			4.383***			4.312***
			(0.405)			(0.403)
Fixed Effects						
Importer	Yes	Yes	Yes	Yes	Yes	Yes
Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Product	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Product	No	No	No	No	No	No
Importer-Exporter	No	No	No	No	No	No
F-Statistics	14.53	13.94	30.85	14.74	14.11	32.40
Observations	583,929	575,823	575,823	571,778	563,671	563,671

Notes: Standard errors 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)
AVE	0.573***	0.580***		0.584***	0.593***	
	(0.201)	(0.206)		(0.202)	(0.209)	
Tariff		0.200***			0.186***	
		(0.065)			(0.058)	
AVE+Tariff			0.252***			0.244***
			(0.074)			(0.069)
Other NTMs			-3.981***			-2.631***
			(0.751)			(0.730)
Fixed Effects						
Importer	Yes	Yes	Yes	Yes	Yes	Yes
Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Product	Yes	Yes	Yes	Yes	Yes	Yes
Observations	583,929	575,823	575,823	571,778	563,671	563,671

Notes: Standard errors 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)
AVE	0.023***	0.024***		0.022***	0.023***	
	(0.006)	(0.006)		(0.005)	(0.005)	
Tariff		1.766***			1.782***	
		(0.201)			(0.206)	
AVE+Tariff			0.056***			0.054***
			(0.007)			(0.006)
Other NTMs			28.159***			24.759***
			(6.036)			(5.623)
Fixed Effects						
Importer-Exporter	No	No	No	No	No	No
Importer-Product	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-Product	No	No	No	No	No	No
Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Observations	568,699	595,094	595,094	591,017	582,438	582,438

Notes: Standard errors 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)
AVE	0.398***	0.422***		0.445***	0.466***	
	(0.136)	(0.137)		(0.136)	(0.138)	
Tariff		1.803***			1.817***	
		(0.222)			(0.228)	
AVE+Tariff			0.991***			1.014***
			(0.129)			(0.131)
Other NTMs			7.716			1.935
			(7.026)			(6.446)
Fixed Effects						
Importer-Exporter	No	No	No	No	No	No
Importer-Product	Yes	Yes	Yes	Yes	Yes	Yes
Exporter-Product	No	No	No	No	No	No
Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Observations	568,699	560,407	560,407	556,431	548,151	548,151

Notes: Standard errors 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)			(5)		
Average AVE of exporter	0.106***	0.106***		0.111***	0.111***	
	(0.018)	(0.018)		(0.019)	(0.019)	
Tariff		-0.058***			-0.055***	
		(0.014)			(0.014)	
Average AVE of exporter			0.161***			0.168***
+Tariff			(0.021)			(0.022)
Other NTMs			12.269***			11.602***
			(1.268)			(1.156)
Fixed Effects						
Importer	No	No	No	No	No	No
Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Product	No	No	No	No	No	No
Importer-Product	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Exporter	No	No	No	No	No	No
F-Statistics	35.02	34.38	59.34	35.20	34.56	59.44
Observations	568,699	560,407	560,407	556,431	548,151	548,151

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

Table 9: Product Frauds: Downward Bias Least Squares Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVE	0.008*	0.005*		0.006	0.005	
	(0.004)	(0.004)		(0.004)	(0.005)	
Tariff		-0.018			-0.020	
		(0.017)			(0.021)	
AVE+Tariff			0.009***			0.005
			(0.004)			(0.004)
Other NTMs			-3.403***			-2.117***
			(0.600)			(0.569)
Fixed Effects						
Importer-Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Product	No	No	No	No	No	No
Exporter-Product	No	No	No	No	No	No
Product	Yes	Yes	Yes	Yes	Yes	Yes
Observations	618,555	610,140	586,200	606,043	597,624	573,672

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

Table 10: Product Frauds: Second Stage IV Regressions

Dependent variables	Quantity Discrepancies			Value Discrepancies		
	(1)	(2)	(3)	(4)	(5)	(6)
AVE	0.771***	0.781***		0.759***	0.773***	
	(0.233)	(0.239)		(0.230)	(0.239)	
Tariff		-0.039*			-0.041	
		(0.021)			(0.028)	
AVE+Tariff			0.083**			0.081***
			(0.033)			(0.029)
Other NTMs			-3.334***			-2.225***
			(0.663)			(0.646)
Fixed Effects						
Importer-Exporter	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Product	No	No	No	No	No	No
Exporter-Product	No	No	No	No	No	No
Product	Yes	Yes	Yes	Yes	Yes	Yes
Observations	583,722	575,607	575,607	571,578	563,461	563,461

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

Table 11: Product Frauds: First Stage Instrumental Variable Regressions

Dependent variables	AVE		AVE+Tariff		AVE	AVE+Tariff
	(3)			(6)		
Average AVE of exporter	0.083***	0.082***		0.088***	0.087***	
	(0.022)	(0.022)		(0.023)	(0.023)	
Tariff		0.012*			0.013*	
		(0.007)			(0.007)	
Average AVE of exporter			0.397***			0.410***
+Tariff			(0.076)			(0.077)
Other NTMs			4.318***			4.225***
			(0.412)			(0.411)
Fixed Effects						
Importer	No	No	No	No	No	No
Exporter	No	No	No	No	No	No
Product	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Product	No	No	No	No	No	No
Importer-Exporter	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistics	14.86	14.27	27.24	15.04	14.40	28.59
Observations	583,722	575,607	575,607	571,578	563,461	563,461

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

Table 12: Trade Elasticities of Tariff and NTMs

Dependent Variable	ln(imports)			
	(1)	(2)	(3)	(4)
ln(1+tariff)	-2.654***	-2.960***		-6.604***
	(0.153)	(0.161)		(0.266)
ln(1+AVE)		-9.856***		-6.604***
		(0.595)		(0.266)
ln(1+tariff+AVE)			-6.734***	
			(0.276)	
Fixed Effects				
Importer-Product	Yes	Yes	Yes	Yes
Importer-Exporter	Yes	Yes	Yes	Yes
Observations	657,510	657,510	657,510	657,510

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