Quantitative comparison of Ad-Valorem Equivalents for non-tariff trade measures: WTO vs. TRAINS

Sionegael Ikeme
Amanda M. Countryman
Dale T. Manning
Diane Charlton

March 29, 2025

Abstract

Non-tariff measures (NTMs) play a growing role in trade policy. Economists quantify NTMs by calculating their Ad Valorem Equivalents (AVE), typically employing one of two global NTM datasets: TRAINS from UNCTAD or WTO notifications. While TRAINS data better measure NTMs, their limited coverage over countries and time is inadequate for many types of analyses, and thus researchers often resort to using WTO data, which is not as accurate as the TRAINS but has better coverage. We compare the two datasets to assess whether AVEs from WTO notifications align with those from TRAINS, given similar country and time coverage. We find significant AVE discrepancies between datasets, sometimes reversing signs, warranting caution when using WTO notifications in the absence of TRAINS data to construct NTMs.

Introduction

Economists have increasingly studied non-tariff measures (NTMs) as crucial elements of international trade policy. While the use of traditional tariffs has declined in general, implementation of NTMs has surged (Orefice, 2017; Niu et al., 2020). Recognizing their growing significance to global trade, numerous studies have estimated the effects of NTMs not only on trade but also on outcomes such as impacts on supply chains (Raimondi et al., 2023), labor (Porto, 2018), and the environment (Mao, Liu, and Wang, 2023). However, measuring NTMs and their impacts can be challenging. Economists typically quantify the effects of NTMs using Ad Valorem Equivalents (AVE). An AVE is the tariff rate that would have an equivalent effect on imported quantities as an NTM, making heterogeneous NTMs more comparable (Disdier, Fugazza et al., 2020). The literature has mostly relied on either the Trade Analysis Information System (TRAINS) database of the United Nations Conference on Trade and Development (UNCTAD) or World Trade Organization (WTO) notifications from the Integrated Trade Intelligence Portal (ITIP) to estimate AVEs. TRAINS data provide more accurate information on NTMs compared to WTO notifications (De Melo and Nicita, 2018; UNCTAD, 2018; Rau and Vogt, 2017) but their limited temporal and country coverage requires the use of WTO data in some contexts (Grübler and Reiter, 2021). TRAINS provides a detailed snapshot of NTMs implemented by select countries, with the UNCTAD actively compiling data from various sources, including domestic regulations, customs agencies, and other relevant bodies (De Melo and Nicita, 2018). In contrast, WTO data relies on a more passive approach, collecting information through self-reported notifications of newly implemented NTMs and not reporting when NTMs are withdrawn. This method often leads to the omission of pre-existing or withdrawn measures and can introduce self-notification bias since the WTO data suffers from underreporting when countries fail to accurately notify NTMs. Despite these limitations, WTO data offer significantly broader country and time coverage at a lower cost of collection compared to TRAINS. Many studies rely on TRAINS data, but its limited country and time coverage often prohibits estimating AVEs over extended periods (Beghin, Maertens, and Swinnen, 2015; Bratt, 2017; Cadot and Gourdon, 2016; Cadot, Gourdon, and Van Tongeren, 2018). Since UNCTAD collects data over successive years for few countries, it is difficult to obtain time-varying AVEs from the TRAINS. To date, Niu et al. (2018) is the only published study to our knowledge that estimates time-varying AVEs using the TRAINS database, but only includes three-year intervals for a limited number of countries. On the other hand, WTO notification data has been used in many studies (Ghodsi et al., 2017; Ghodsi and Stehrer, 2022; Grübler and Reiter, 2021;

Adarov and Ghodsi, 2023) due to its broader country and time coverage. This expanded coverage not only facilitates the estimation of time-varying AVEs (Adarov and Ghodsi, 2023) but also allows for analysis of a greater number of countries. This is particularly important because AVEs are estimated based on cross-country variation, and a lack of country variation can introduce bias. Despite their common use, it remains unclear if using WTO data produces similar AVE estimates as the TRAINS dataset, conditional on similar time and country coverage. This research helps fill this gap in the literature.

In this paper, we compare the two primary global NTM databases and assess how similar AVEs are when derived from WTO notifications and TRAINS, examining only countries and years where the databases have comparable coverage. We assume that differences in AVEs across databases are likely due to measurement errors inherent in WTO data because of the WTO's known limitations. We assess whether differences in AVEs vary by importing country, product, and time, which could provide insights into sources of measurement error. Finally, we explore whether certain factors related to the importer, exporter, and product can explain the observed differences in AVEs since some countries might be more likely to correctly notify the WTO of NTMs. Specifically, we assess whether the importance of imported goods for both the importer and exporter, the GDP of both countries and the overall NTM intensity imposed by the importer can explain the magnitude of AVE differences. Assessing the sensitivity of AVE estimates to data sources is critical for several reasons. First, this quantitative comparison reveals whether WTO notification data can be substituted for TRAINS data in years and countries for which TRAINS data do not exist. This would suggest that the WTO can provide an alternative to using TRAINS and allow for the estimation of AVEs across more countries and over longer time periods. Alternatively, if estimated AVEs differ between the two datasets, it can serve as a warning of potential biases or inconsistencies inherent in the WTO data. Finally, identifying shortcomings of each dataset can help optimize their use. By understanding the sources of discrepancies and the measurement errors associated with WTO notification data, we can lay the groundwork for future research to improve the use of NTM data to estimate AVEs.

To estimate AVEs, we employ a trade quantity-based approach proposed by Kee and Nicita (2022). This method enables us to derive bilateral AVEs at the yearly Harmonized System (HS) 6-digit product level from 2013 to 2016 while addressing the endogeneity of NTMs and tariffs by using the NTMs and tariffs imposed by neighboring importers as instrumental variables (IVs). We concentrate on nine South American countries given their geographic proximity and relatively good coverage in both datasets. In addition, we focus on two types of NTMs: Technical Barriers to Trade (TBT) and Sanitary and Phytosanitary (SPS) measures. We prioritize these measures because they are the most frequently used types of NTMs (Ghodsi et al., 2017) and are particularly dominant in agri-food sectors (Raimondi et al., 2023). We find that while some AVEs yield relatively similar values, this is not the case for all estimates at the product-importer-exporter-year level. In fact, there are instances where the differences in AVEs are substantial, sometimes even resulting in opposite signs. This is particularly concerning since it implies that the estimated effect of NTMs on trade flows can be reversed in some cases depending on the data source used.

While previous studies have explored the qualitative differences between the two NTM databases (Rau and Vogt, 2017; De Melo and Nicita, 2018), to our knowledge, no research has quantitatively compared the resulting AVE estimates. This study bridges that gap by analyzing the statistical differences in the estimated AVEs derived from both sources for similar time and country coverage. This paper contributes to the literature by offering insights into meaningful differences between the datasets employed when estimating AVEs for NTMs, thereby providing valuable guidance for future analyses, especially those exploring the effects of trade policy on international trade and other economic outcomes. We also quantify the measurement error in WTO notification data. We find that WTO and TRAINS data do not yield similar AVEs even when examining countries and time periods where both datasets have high coverage. This highlights the need for improved data quality and collection methods for studying NTMs. Future data collection should focus on both accuracy and coverage. Furthermore, our findings suggest measurement errors in WTO notification data lead to significant miscalculations of AVEs, underscoring the importance of using extreme caution when using the WTO data. Before delving into the quantitative analysis, we provide background on NTM data and their use in trade policy. Then, we qualitatively compare the two databases and outline how the literature has quantified NTMs through AVEs. Following this, we explain our methods and provide an overview of the data used to estimate AVEs in this study. We then present results and conclude with a discussion of key findings and policy implications.

Background

Current NTM data

There are multiple sources of NTM data available for analysis. Rau and Vogt (2017) provide a comprehensive overview covering inventories, legislative reviews, notifications, surveys, import refusals, and other sources. Among these, national legislation inventories emerge as the most reliable. These inventories meticulously catalog national legislation, allowing measures to be classified into different types of NTMs using the MAST classification system[1]. Notably, the TRAINS database, jointly developed by UNCTAD and the International Trade Centre (ITC), stands out as a primary resource for researchers investigating NTM effects. This database meticulously gathers and classifies data for NTMs from all legal texts issued by countries, including bilateral and plurilateral agreements. The process involves active data collection from multiple sources, which vary depending on the type of NTM and is subject to independent regulatory review by UNCTAD[2]. The TRAINS database provides HS-6 level NTM data for 103 countries, detailing the implementing country, the affected countries, descriptions of the regulation and product, and the dates of NTM implementation and withdrawal. While data collection began in 1994, there was a major restructuring in 2006, and consequently, the most current NTM records were collected post-2008 (UNCTAD, 2018). Key strengths of the database in-

[1] The Multi-Agency Support Team group (MAST), established under UNCTAD, has been instrumental in shaping the International Classification of NTMs. This classification system is extensively used to categorize and grasp the diverse forms of NTMs impacting global trade.

[2] For instance, SPS-related information can be gathered from ministries of agriculture, standardization agencies, and sometimes even from WTO notifications (UNCTAD, 2018).

clude its traceability, cross-country comparability, transparency, and open accessibility. However, the inconsistent frequency of data collection across countries presents challenges in reliably estimating AVEs over time. This limitation is particularly restrictive in panel analyses, where only countries with multiple years of data can be effectively included, reducing the overall coverage of AVE estimates.

Another valuable source of NTM data is WTO notifications. As WTO members, countries are required to report the implementation of NTMs to enhance the transparency and predictability of trade policies. The current WTO notification dataset covers NTMs for 140 member countries since 1979, providing information at the HS-6 level on the type of NTM, product and regulation details, notification date, and, in some cases, the withdrawal year. While this dataset covers more countries and periods than TRAINS[3], its passive data collection relies on self-reported submissions from member states, causing severe limitations. Countries do not always notify the WTO when an NTM is implemented, or in a timely manner, causing discrepancies between actual implementation and notification years (De Melo and Nicita, 2018). WTO members are only required to notify the WTO of NTMs when they may have a nonmarginal impact on other trading countries or when the regulation is not based on "relevant international standards" (UNCTAD, 2018). This results in delays, omissions, and self-notification bias since there is no external verification process. Additionally, missing HS codes pose challenges for analyzing NTMs at the product level. To address this, we follow Ghodsi, Reiter, and Stehrer (2015) in retrieving missing codes using product information, as detailed in the Data section. Another limitation is the inconsistent recording of NTM withdrawals, particularly for SPS and TBT measures, where no withdrawal records exist. Furthermore, the dataset sometimes lacks implementation dates, requiring the use of notification dates as a proxy. These limitations introduce measurement errors in AVE estimation. Consequently, this database is generally less favored compared to TRAINS data for studies where country and temporal coverage in the TRAINS are adequate. In addition to direct notifications, the WTO dataset includes reverse notifications, where WTO members report NTMs im-

[3] Appendix section I. provides a comparison of data coverage between WTO notification data and TRAINS data.

posed by other members(Rau and Vogt, 2017). These reverse notifications fall under the Special Trade Concerns (STCs) database. While combining STC data with direct WTO notifications could help account for under-reporting by certain countries, matching STC cases with actual notifications is challenging (Ghodsi, Reiter, and Stehrer, 2015).

Business surveys and complaint portals serve as other possible sources of NTM data and offer insights from the perspective of businesses facing trade barriers. Examples include the ITC NTM surveys and the EU market access database. However, survey data tend to be country-specific and often represent potentially biased reports from firms affected by NTMS. Consequently, survey results tend to overlook NTMs that have minimal perceived impact (Rau and Vogt, 2017). Hence, such data must be interpreted with caution, considering the inherent biases and limitations associated with the perspective of surveyed businesses. Accordingly, we focus exclusively on NTM data from the UNCTAD TRAINS and WTO notification databases for this analysis. They are the most commonly used sources for estimating AVEs, and their broad coverage and relatively reliable information make them arguably the most suitable databases for AVE estimation.

Global NTM data limitations

Despite the extensive amount of information on NTMs, these databases share common limitations when used in quantitative studies. These challenges stem not only from the way NTMs are collected but also from the inherent nature of NTMs themselves. De Melo and Nicita (2018) highlight six key limitations to consider when using databases such as TRAINS or WTO notifications. Limitations include limited comprehensiveness, lack of precision, data dimensionality, time dimension, product-specific requirements, and endogeneity.

Comprehensiveness refers to both omissions and double counting. As discussed previously, there are multiple sources for NTM databases, making data collection challenging due to information being scattered across various sources, which often leads to missing data. Double counting can arise when identical NTMs from different sources refer to the same regulation. In addition, cross-country comparisons are difficult because data availability and collection efforts vary across countries. Notification-based data also face challenges, as reporting requirements are not always strictly followed and overlapping measures may be counted twice. The lack of precision stems from the qualitative nature of NTM data, making it difficult to determine the stringency of trade regulations. Data often fail to specify how regulation impacts trade costs or the economic sectors it affects. Additionally, the same regulation can have varying effects across countries, with trade costs differing by industry and exporting country. To address this challenge, economists use NTM data to infer stringency by estimating AVEs. Dimensionality poses challenges in disentangling the econometric effects of individual NTMs on products due to the heterogeneous nature of measures classified as NTMs, particularly at the HS-6 level. Collinearity among NTMs is a common issue, as multiple NTMs are often implemented on the same product, making it difficult to isolate the impact of any single measure. When multiple NTMs are applied simultaneously, their effects may overlap, creating statistical challenges in identifying their distinct contributions. As a result, analyses often aggregate NTMs into broader categories, such as SPS and TBT, to improve reliability of AVE estimations. Issues related to time dimensions stem from data collection. While NTM databases often include implementation or notification dates, this information is not always precise enough for time-series or event-study types of analyses. In notification-based datasets, measures that have existed unchanged for a long time are generally not reported, as countries are only required to notify newly implemented NTMs or modifications to existing policy. As a result, long-standing historical NTMs may be missing from the data. Similarly, in the case of the UNCTAD TRAINS database, data is not collected continuously for all countries across years. This irregularity makes it particularly challenging to perform time-series analyses. Product-specific requirements refer to NTMs that affect products at the HS-6 level, covering approximately 5,000 traded goods. However, NTMs do not impact all products uniformly. Certain categories, such as food and chemicals, are more heavily regulated and more likely to be subject to NTMs compared to raw materials, for instance. This variation requires caution when estimating the effects of NTMs, ensuring that the analysis is sufficiently disaggregated to capture these differences. However, the more disaggregated the analysis, the more complex it becomes, posing additional methodological challenges. Finally, endogeneity presents a major challenge in assessing the causal effects of NTMs on trade flows and prices, with reverse causality being a key concern. Countries may impose stricter NTMs on sectors that are economically or politically significant, leading to simultaneity problems. This can be an especially important issue when estimating AVEs where we would like to isolate and quantify the true impact of NTMs. Often, analysis employs IVs to address endogeneity (Bratt, 2017; Kee and Nicita, 2022; Adarov and Ghodsi, 2023).

Due to the various challenges arising from data collection and the inherent nature of NTMs, quantifying their effects is not straightforward. In particular, inconsistencies in data collection can lead to discrepancies between sources. As a result, AVE estimates derived from WTO notifications and UNCTAD TRAINS data are likely to differ, highlighting the importance of carefully considering limitations across datasets when analyzing the impact of NTMs.

Quantification of NTM effects

To understand how the effects of NTMs are quantified, we first examine their impact on trade. When an importing country implements an NTM on a product from an exporting country, the exporter may incur additional costs to comply with the regulation, which can be passed on to the consumers of the importing country. Exporters incur higher fixed and or variable costs to comply with NTMs, effectively imposing compliance costs that vary across countries. Those with strong infrastructure and regulatory alignment can adapt more easily, while others lacking such capacity face significantly higher costs, making compliance a greater burden. However, NTMs can also foster market creation. Specifically, TBTs help reduce asymmetric information by signaling uniform standards, reassuring consumers, and boosting demand (Cadot, Gourdon, and Van Tongeren, 2018; Gourdon, Stone, and van Tongeren, 2020). As a result, NTMs generate two opposing effects: compliance costs and market creation. Figure 6 from section II. in the appendix illustrates these mechanisms.

Although various methods exist to quantify the effects of NTMs, the econometric approach enables us to estimate the price and/or quantity changes induced by their presence. Within this framework, two main methodologies emerge: the price-based approach, which captures only compliance costs, and the quantity-based approach, which accounts for both compliance costs and market creation effects. The quantity-based method examines how NTMs affect trade volumes and translates these effects into trade costs using elasticities of substitution. This approach identifies instances where import quantities rise (indicating that the demand creation effect outweighs compliance costs, leading to negative AVEs) and cases where they decline, reflecting the opposite effect. In contrast, the price-based approach focuses on how NTMs influence prices, typically expecting an increase when compliance costs rise (Cadot and Gourdon, 2016).

Both approaches rely on gravity models, which regress bilateral trade flows on importer-exporter characteristics and trade barriers, including tariffs and NTMs. Past studies have employed both quantity-based methods (Kee, Nicita, and Olarreaga, 2009; Essaji, 2008; Beghin, Maertens, and Swinnen, 2015; Ghodsi and Stehrer, 2022) and price-based methods (Cadot and Gourdon, 2016; Ing et al., 2016; Cadot, Gourdon, and Van Tongeren, 2018; Kravchenko et al., 2022) to estimate AVEs, with the choice often depending on data availability and model specifications. Cadot, Gourdon, and Van Tongeren (2018) compare these approaches, highlighting key differences in their measured effects. While both offer valuable insights, we adopt the quantity-based method in this paper, as it captures both positive and negative AVEs, which is particularly relevant for SPS and TBT measures, where the incidence of negative AVEs is substantial (Gourdon, Stone, and van Tongeren, 2020). Among quantity-based methods, we focus on bilateral AVE estimation, which accounts for AVE variation across importing and exporting countries at the product-year level. Indeed, compliance costs for NTMs vary across countries, even under uniform regulations (Bratt, 2017; Kee and Nicita, 2022), with exceptions for SPS measures that target specific exporters. To capture this variation, we follow Kee and Nicita (2022), modeling imports as a function of gravity variables, tariffs, and NTMs. A key advantage of this method is its simultaneous estimation of import elasticities, eliminating the need for pre-estimated values from studies like Bratt (2017).

Methods

To estimate AVEs, we use a gravity model at the HS-6 digit level to assess trade effects at the importer-exporter-year level from 2013 to 2016. We compare AVEs derived from two NTM data sources by estimating the gravity model twice: once with data from WTO notifications and once with data from TRAINS. In addition, we focus on nine South American importers and 82 exporters, selected based on data availability across multiple years in both databases. In this section, we first address potential endogeneity in estimating tariff and NTM effects on trade, and then we present the gravity framework. Finally, we explain how the estimated coefficients are used to calculate AVEs.

Accounting for Endogeneity

The gravity model estimates the impact of NTMs on import quantities by regressing NTMs on trade flows. This approach quantifies how an additional NTM affects the volume of imports while controlling for gravity characteristics and other trade barriers. However, before regressing NTMs on import quantity, we must address the issue of endogeneity of trade volume and prices inherent in both NTMs and tariffs (Kee, Nicita, and Olarreaga, 2009; Bratt, 2017; Kee and Nicita, 2022). Endogeneity arises from reverse causality since the importance of a traded good might influence the implementation of trade policies. Specifically, factors such as import penetration in industries can drive NTM protection (as import penetration increases, governments may impose NTMs as protective measures (Trefler, 1993; Lee and Swagel, 1997)). Prior studies have addressed the endogeneity issue using IVs, a common approach being the use of NTMs imposed by neighboring countries as instruments (Kee, Nicita, and Olarreaga, 2009; Cadot, Gourdon, and Van Tongeren, 2018; Bratt, 2017; Kee and Nicita, 2022). Following this approach, we adopt a two-step regression method to correct for potential bias. In the first step, we regress an importer's

NTM implementation on the mean implementation of its three closest neighbors[4]. In the second state, we then use the predicted NTMs from this regression in the gravity equation. We apply the same instrumental variable strategy for tariffs since tariffs have the same endogeneity concerns. Specifically, we estimate the level of tariffs (equation 1) and the number of NTMs (equation 2) imposed by country i on product k, defined at the HS 6-digit level, as a function of the average of NTMs and tariffs of three neighboring countries of i for each year from 2013 to 2016.

$$t_{ijk} = \widehat{\alpha_i}\bar{t}_{ijk} + Z_{ij}\hat{\alpha} + \epsilon_{ijk} \tag{1}$$

$$NTM^n_{ijk} = \widehat{\gamma_i}\overline{NTM}^n_{ijk} + Z_{ij}\hat{\alpha} + \epsilon_{ijk} \tag{2}$$

Where the term $\bar{t}_{ijk}$ denotes the average applied tariff imposed by the 3 neighboring countries of i on imports from j for good k. Z_{ij} represents gravity variables, including distance, contiguity, landlocked status, common language, regional trade agreements (RTA), and colonial relationships, as well as the logarithm of GDP for both the importer and exporter. $\overline{NTM}^n_{ijk}$ refers to the average count of non-tariff measures (NTMs) of type n (such as SPS, TBT, or other NTMs) imposed by the three neighboring countries of i on imports from j for good k. We apply Equation 2 to all types of NTMs considered in our study, including SPS measures, TBTs, and other types of NTMs[5]. We then use the estimated NTMs and tariffs at the product-importer-exporter level to assess their impact on trade. However, employing NTMs from other countries as IVs introduces a potential concern: it may exacerbate the measurement error inherent in WTO notification data. This approach risks amplifying existing biases, as NTMs from neighboring countries—which may themselves be subject to inaccuracies—are used to predict NTMs imposed by others.

[4]Neighboring countries are defined based on the geographical distance between their capitals.

[5]NTMs other than SPS and TBT are categorized in one variable, but, as this is not the focus of our study, we will mostly focus on the results for SPS and TBT.

Quantity based gravity model

Once we obtain the predicted NTMs and tariffs from our stage 1 regression, we include them in the second stage gravity model, where we express bilateral import value as a function of bilateral gravity characteristics, NTMs, and tariffs[6]. We derive the following gravity equation for each good k, capturing trade flows from importer i to exporter j.

$$\begin{aligned}\ln Q_{ijk} = \beta_k + \left(\beta_k^t + \beta_{k1}^t \,\text{share}_{ki} + \beta_{k2}^t \,\text{share}_{kj}\right) t_{ijk} \\ + \sum_n \left(\beta_k^{NTM^n} + \beta_{k1}^{NTM^n} \,\text{share}_{ki} + \beta_{k2}^{NTM^n} \,\text{share}_{kj}\right) NTM_{ijk}^n + Z_{ij}\beta + \varepsilon_{ijk} \quad (3)\end{aligned}$$

In the gravity equation, Q_{ijk} represents the value of imports of product k by importer i from exporter j. The variables $share_{ki}$ and $share_{kj}$ denote the importer's and exporter's share in the world market for product k, respectively. The term t_{ijk} refers to the predicted applied tariff imposed by importer i on imports of product k from exporter j, while NTM_{ijk}^n indicates the predicted count of non-tariff measures imposed by importer i on imports of product k from exporter j. The NTM effect is estimated at the importer-exporter level by interacting the NTM count of type n with the market share of the importer and exporter for that good in global trade. Z_{ij} consists of all gravity variables, such as distance, contiguity, landlocked status, language, and regional trade agreements, as well as the logged GDPs of both the importer and exporter. We also adjust some of the gravity variables by GDP to account for multilateral resistance following Bratt (2017); Baier and Bergstrand (2009). We provide more details in section III. in the appendix. To account for zero trade flows, we employ a Pseudo-Poisson Maximum Likelihood (PPML) specification. Following Kee and Nicita (2022), we estimate this model separately for

[6] Kee and Nicita (2022) use import quantity instead of import value; however, this approach often results in missing data, particularly for goods and services with inconsistent or undefined units of quantity. As noted by Ferrantino (2006) and Kee, Nicita, and Olarreaga (2009), import value is commonly used as a proxy for quantity.

each year and at the HS-6 digit product level. We denote the obtained coefficients as:

$$\beta^{t}_{ijk} \equiv \beta^{t}_{k} + \beta^{t}_{k1}\, share_{ki} + \beta^{t}_{k2}\, share_{kj} \tag{4}$$

$$\beta^{NTM^n}_{ijk} \equiv \beta^{NTM^n}_{k} + \beta^{NTM^n}_{k1}\, share_{ki} + \beta^{NTM^n}_{k2}\, share_{kj} \tag{5}$$

We use the coefficients β^{t}_{ijk} and $\beta^{NTM^n}_{ijk}$ to obtain AVEs at the product-importer-exporter-year level. Finally, we cluster standard errors at the importer level to address known issues related to heteroscedastic standard errors in the gravity model (Fell and Duver, 2024).

Estimation of Ad Valorem Equivalents

To derive the AVEs, we follow Kee and Nicita (2016, 2022), where we determine the marginal effects of NTMs on quantity traded from equation 5. From the gravity equation (1), we know that the proportionate change in import quantity due to an additional NTM from m to $m + 1$ is:

$$\ln(Q_{ijk}|NTM = m + 1) - \ln(Q_{ijk}|NTM = m) = \widehat{\beta}^{NTM}_{ijk} \tag{6}$$

$$\frac{Q_{ijk}|NTM = m + 1}{Q_{ijk}|NTM = m} = \exp(\widehat{\beta}^{NTM}_{ijk}) \tag{7}$$

$$\frac{Q_{ijk}|NTM = m + 1 - Q_{ijk}|NTM = m}{Q_{ijk}|NTM = m} = \exp(\widehat{\beta}^{NTM}_{ijk}) - 1 \tag{8}$$

Similarly, the proportionate change in quantity imported due to a one percentage point increase in an import tariff is:

$$\frac{Q_{ijk}|t_{ijk} = t + 1 - Q_{ijk}|t_{ijk} = t}{Q_{ijk}|t = t} = \exp(\widehat{\beta}^{t}_{ijk}) - 1 \tag{9}$$

The AVE of an NTM on product k 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 the NTM (Kee and Nicita, 2016). Thus, we compute the

$AVE_{ijk}^{NTM^n}$ for NTM type n, using the following expression:

$$\text{AVE}_{ijk}^{NTM^n} = \frac{\exp(\widehat{\beta}_{ijk}^{NTM^n}) - 1}{\exp(\widehat{\beta}_{ijk}^{t}) - 1} \tag{10}$$

This represents the mean AVE for an additional NTM on product k imposed by importer i on exporter j. Using equation 10, we can obtain the effect of the total number of NTMs imposed by country i by multiplying the NTM coefficient in equation 6 by the total count of NTMs to obtain the total AVE for all NTMs imposed as follows (Kravchenko et al., 2022; Fell and Duver, 2024).

$$\text{AVE}_{ijk}^{NTM^n} = \frac{\exp(\widehat{\beta}_{ijk}^{NTM} * NTM_{ijk}^{n}) - 1}{\exp(\widehat{\beta}_{ijk}^{t}) - 1} \tag{11}$$

Data

Trade and tariff data

This study uses a dataset covering trade flows between 2013 to 2016 for nine South American importing countries and 82 exporting countries. The data are disaggregated to the HS-6 digit product level, following HS4 revisions[7]. Trade values (in U.S. dollars) are obtained from the World Integrated Trade Solution (WITS) database, while gravity variables (log of distance, common language, contiguity, colonial ties, how landlocked importer and exporter are) come from the Center for Prospective Studies and International Information (CEPII) (Conte, Cotterlaz, and Mayer, 2022), and regional trade agreement (RTA) data are sourced from Blöthner and Larch (2022). GDP values are taken from the World Bank. For tariff data, we use Most-Favored-Nation (MFN) tariffs from the World Integrated Trade Solution (WITS) database. After merging all data by product-importer-exporter-year, the final dataset forms a balanced panel that is then subset by year and product for gravity analysis. The dataset, which includes nine importing coun-

[7]The importing and exporting countries used in this analysis can be found in the appendix section IV - I.

tries, 82 exporting countries, and 5,152 products over four years (2013–2016)[8], is then merged with the NTM data. Summary statistics for the variables used in this analysis are provided in the appendix section IV - II.

NTM data

We use the UNCTAD TRAINS database and WTO notification data available through the Integrated Trade Intelligence Portal (I-TIP). However, due to missing HS codes, we rely on the dataset provided by Ghodsi et al. (2017), which is derived from I-TIP data. Both datasets provide information on NTM types (e.g., SPS, TBT), the count of applied NTMs, the implementing country, the affected country, relevant regulations, and affected product details. While the TRAINS database can be readily employed, WTO notifications require certain assumptions and adjustments to align with our analysis. We follow the recommendations of Ghodsi, Reiter, and Stehrer (2015) to transform the WTO data. First, the WTO notification data has missing product classifications. To address this, Ghodsi et al. (2017) retrieved missing HS codes by cross-referencing product descriptions with other datasets. However, if an HS code is inaccurately retrieved, it may introduce measurement errors. Second, we address the lack of withdrawal information by adopting the approach proposed by Ghodsi et al. (2017); Ghodsi and Stehrer (2022), which assumes that NTMs, once implemented, remain in place indefinitely. This assumption enables us to construct a cumulative count of NTMs per year. However, it also imposes an upward bias in implemented NTMs, likely overstating the number of active NTMs over time, at least among NTMs that were ever recorded in the dataset. Finally, since the WTO notification database only includes notified NTMs, we assume that all notified NTMs represent the total number of NTMs implemented, although, in reality, some measures may exist without formal notification. Additionally, the WTO might be notified of certain NTMs with a delay, which introduces measurement errors regarding the actual year of implementation. After processing the NTM data, we merge each NTM dataset with bilateral trade and country characteristics to create a bilateral panel dataset at the

[8]This is equivalent to 15,023,232 observations. We have $4\times5{,}152\times9\times81$ observations. Although there are 82 exporters, we exclude observations where the importer is the same as the exporter.

HS-6 digit level. This dataset includes the count of NTMs imposed by an importer on an exporter. Since our analysis focuses on SPS and TBT measures, all other NTMs are grouped into a single aggregated category. To maintain consistency across analyses, we use TRAINS data for non-SPS and non-TBT NTMs in both regressions (the gravity model with TRAINS data and with WTO notification data). This approach ensures that the only variation in estimated AVEs is due to differences in data sources for the SPS and TBT measures, allowing for a more direct comparison. Table 3 in section IV - III of the appendix presents the summary statistics for the observed count of NTMs in the finalized dataset. The average number of applied NTMs does not seem to vary between the two datasets, but the third quantile reveals that the WTO data for SPS measures contains more zeros than the TRAINS data, as shown in tables 4 and 5, which report the percentage of observations with an NTM count of zero. Additionally, the maximum number of NTMs is considerably higher in the WTO data for both SPS and TBT cases. which might be due to unreported withdrawals of NTMs in the WTO.

We also compare NTM coverage across HS-6 codes and differences in NTM counts between the two datasets, specifically by looking at the variations in NTM coverage at the HS-6 digit level and differences in the total number of NTMs imposed in TRAINS and WTO data. Figures 7 and 8 in appendix section IV - IV display the number of distinct HS codes affected by NTMs over the years for each importing country. We observe discrepancies in the number of HS-6 codes covered by at least one NTM and note that these discrepancies vary both across importers and NTM types. Another perspective is provided in figures 9 and 10, which, instead of focusing on HS coverage, represent the total number of NTMs. Since we account for NTM stringency through total NTM counts, this metric is equally important. Interestingly, for some countries, even when the number of distinct HS codes covered is similar across datasets, the total NTM count still differs. These observed discrepancies likely stem from measurement errors in WTO notification data. These figures highlight two main factors contributing to differences in AVE estimates between the datasets: either the absence of certain HS-6 codes from NTM coverage or discrepancies in the total number of NTMs applied to a given HS code. A wide gap in

NTM coverage could contribute to significant differences in AVE estimates. However, if the impact of NTMs on trade is relatively small, the resulting difference in AVEs could be marginal, so estimating AVEs in each dataset remains essential. Finally, to compare NTM coverage across HS sections, appendix section IV - V provides a detailed comparison of NTM coverage across each HS section, with additional details on HS sections also given in appendix section IV - VI.

Results

Differences in AVE Estimates

We obtained gravity coefficients from the 20,608 regressions (5,152 products × 4 years), which are summarized in the appendix section V - I. We then estimated AVEs at the product-importer-exporter-year level, yielding 15,023,232 AVEs for each dataset. To meaningfully compare AVE estimates between the two datasets, we match each AVE from the TRAINS dataset to its equivalent from the WTO dataset at the HS-6 digit product-importer-exporter-year, and NTM type level. Tables 10 and 11 in the appendix section V - III present summary statistics for the AVE estimates obtained for SPS and TBT measures across both datasets. Given that we estimate multiple regressions at the product-year level, we present the results for the obtained AVEs while also removing extreme values following Bratt (2017)[9].

We begin with a global comparison of AVE values due to convergence issues[10]. Tables 1 and 2 compare AVEs from TRAINS and WTO notifications across all estimates, showing the proportion of AVEs in TRAINS relative to their corresponding values in WTO. The two tables classify observations into three groups: AVEs equal to 0, NA (due to lack of convergence in regression estimation), or nonzero/non-NA across both NTM datasets. Comparing these categories helps assess dataset alignment and identify initial discrepan-

[9]Specifically, we drop AVE values below the 10th percentile and above the 90th percentile as some estimates reach extreme magnitudes. This issue was also noted by Bratt (2017), who similarly excluded extreme AVEs in their analysis, attributing these outliers to the estimation process, which relies on a combination of six different coefficients to compute AVEs.

[10]Details on convergence issues are provided in the appendix section V - II.

cies. Tables 1 and 2 show that AVEs are zero in both datasets for approximately 20% of observations, often due to NTMs being recorded as zero. However, NTMs are recorded as zero more frequently compared to the proportion of zero AVEs (about 62% for SPS and 31% for TBT as shown in appendix section IV - III). This discrepancy arises from convergence issues, which often stem from a lack of variation in the NTM variables, particularly when NTMs are uniformly recorded as zero across all importers in the gravity equation. Indeed, we observe that AVEs are often missing (NA) in both datasets, as shown in table 3. This table illustrates how the presence of NTMs is reflected in estimated AVEs across datasets and NTM types, recording the percentage of observations in different NTM value categories and their corresponding AVE values. For both SPS and TBT, a substantial number of observations have NTMs equal to zero while AVEs remain NA. This suggests that the presence of NAs is not an issue, as it is likely that the NTM count is zero and the corresponding AVE should be zero as well. A notable discrepancy arises when AVEs based on WTO data are nonzero while those from TRAINS are zero (occurring in 13.93% of SPS observations and 3.36% for TBT). Since AVEs are the product of the NTM gravity coefficient and count, they are zero when the NTM count is null. This suggests that these cases occur when NTMs are recorded in the WTO dataset but not in TRAINS. A likely explanation is that the WTO dataset assumes NTMs are never withdrawn, meaning some measures persist in its records even when absent from TRAINS. Conversely, in 12–15% of SPS and TBT cases, TRAINS produces nonzero AVEs while WTO produces AVES of zero. This indicates that some NTMs may be missing from the WTO dataset but are captured in TRAINS, possibly due to incomplete WTO notifications by importers or the exclusion of NTMs implemented before 1995, which are not included in our WTO data. Finally, in 13.30% of SPS and 31.32% of TBT observations, both datasets yield nonzero AVEs. Within this subset for SPS measures, 6.7% have the same sign for AVEs, while 6.5% have opposite signs. For TBT measures, 17% of the AVEs share the same sign across both datasets, suggesting a consistent directional effect on trade. However, in 14.5% of TBT cases, AVEs have opposite signs. This noteworthy share of AVEs with opposite signs raises concerns, as it suggests conflicting interpretations of NTMs' impact on trade. Such divergence highlights inconsistencies in the NTM data, which could significantly influence policy analysis. The subsequent analysis focuses on AVEs that are neither NAs nor equal to zero in both datasets. The subsequent analysis focuses on AVEs that are neither NAs nor equal to zero in both datasets. Specifically, we examine 30.02% of the original SPS observations (12.36% + 13.30% + 3.36% from table 1) and 60.7% of the TBT observations (15.45% + 31.32% + 13.92% from table 2). By narrowing our focus to these subsets, we can more effectively assess AVE discrepancies among NTMs that have an effect on trade and identify key factors driving the differences between the datasets.

To examine the differences in AVE values between TRAINS and WTO notification data, we plot the density distribution of the percentage point difference in AVEs (computed as TRAINS AVE minus WTO notification AVE for each product-importer-exporter-year observation), as shown in figure 1. From table 4, the mean AVE difference is 57.35 for SPS measures, with a median of 0.78, while for TBT measures, the mean difference is -178.51, with a median of -0.01. For both NTM types, the distribution of AVE differences exhibits a central peak around 0, along with two additional peaks for values above and below zero. The first and third quantiles in table 4 are notably high for both SPS and TBT cases, indicating substantial variation in AVE discrepancies. To further analyze these differences, we categorize AVE discrepancies based on the cases identified in tables 1 and 2. Specifically, we examine AVE differences across four categories: (i) AVEs with the same sign in both datasets, (ii) AVEs with opposite signs, (iii) cases where AVEs are zero in TRAINS but nonzero in WTO, and (iv) cases where AVEs are nonzero in TRAINS but zero in WTO. Figure 2 illustrates these cases, showing that AVE discrepancies remain substantial. Notably, in categories (i), (iii), and (iv), differences are not always centered around zero, highlighting the extent of measurement error in the WTO data. To illustrate these discrepancies in AVEs, figure 13 in appendix V - V presents the case of SPS measures imposed by Brazil on Argentina for some crop and vegetable related HS-6 digit products. We find that while some products, such as sorghum and saffron, exhibit similar AVEs across the two datasets, others, such as coriander, show much larger discrepancies in AVE estimates, highlighting specific inconsistencies in AVE estimation between datasets. Furthermore, we truncate our data where AVEs below -100% are replaced with -100%, as values lower than this threshold are difficult to interpret in an economic context (Bratt, 2017; Cadot, Gourdon, and Van Tongeren, 2018). The results with truncated values are given in figure 12 in appendix section V - IV and remain consistent with figure 1. From this analysis, it is visible that the AVE estimates from the two datasets are not identical, and in some cases, the differences can be substantial. This discrepancy suggests that the AVEs derived from TRAINS data and WTO notifications differ, which is due to measurement errors in WTO notification data.

Finally, we examine whether differences in estimated AVEs across datasets vary by importing country, affected product, and the year of NTM implementation. The results are presented in the appendix section V - VI. Key takeaways from this expanded analysis are that the differences in AVEs between the two datasets vary by importer, product, and time. In the next section, wegate these discrepadiscrepancies further by regressing AVE differences on additional factors.

Explaining the difference in AVE estimates

To explore the observed bias and detect systematic errors, we regress the difference in AVEs on various factors that may contribute to measurement errors in the WTO data as follows:

$$|AVE^{n,UNCTAD}_{i,j,k,t} - AVE^{n,WTO}_{i,j,k,t}| = importImp_{i,k,t} + Share_{i,k,t} + Share_{j,k,t} + GDPpercapita_{i,t} + GDPpercapita_{j,t} + totNTM^{n}_{i,t} + \tau_t + \epsilon_{i,j,k,t} \qquad (12)$$

First, we examine the likelihood of countries misreporting or failing to notify NTMs correctly by including several explanatory variables. These include the importance of the imported product in total imports for the importing country ($importImp_{i,k,t}$), measured as the ratio of the good's import value to total imports in a given country. We also account for the importer's importance in the world market using its global market share for that specific good ($Share_{i,k,t}$), as well as the exporter's global market share ($Share_{j,k,t}$). We include the GDP per capita of both the importing ($GDPpercapita_{i,t}$) and exporting countries ($GDPpercapita_{j,t}$). Another key factor is the total number of SPS or TBT measures implemented by the importer in a given year, using TRAINS data ($totNTM^{n}_{i,t}$), which helps to assess whether countries that impose a high number of NTMs overall are more prone to reporting inconsistencies in the WTO notification data. We also include time fixed effects (τ_t) to assess whether the assumption of cumulative implementation of NTMs over time in WTO notifications contributes to measurement errors over time. We estimate the model using clustered standard errors at the product-importer-year level to account for heteroskedasticity and potential correlations within clusters.

The regression results are provided in table 5. For SPS, we find that the export share is statistically significant at the 95% confidence level. As the export share increases, the difference in AVEs decreases. This could indicate that when an exporting country has a larger presence in the global market, importers may be less likely to misreport or fail to notify NTMs correctly. However, attributing this entirely to misreporting remains challenging, as there are other potential sources of measurement error beyond notification discrepancies, such as differences in how NTMs are recorded or interpreted across datasets. In the TBT case, the GDP per capita of the reporter is significant, suggesting that higher reporter GDP per capita is associated with greater differences in AVEs. However, the magnitude of the effect appears small. Additionally, we find that import share is statistically significant at the 95% confidence level, suggesting that as the importer's world market share for a given good increases, the difference in AVE estimates between the two datasets decreases. Similar to the SPS case, the export share is also significant and negatively correlated with AVE differences, reinforcing the idea that larger exporting countries may be less prone to NTM misreporting. We also observe that the signs of the coefficients for import importance, importer market share, and export share remain consistent across the two types of NTMs, suggesting that similar factors may be influencing discrepancies in AVE estimates for both SPS and TBT measures. However, this analysis remains exploratory, and the relatively low adjusted R-squared values indicate that the model explains only a small portion of the variation in AVE differences, implying that additional factors may contribute to these discrepancies.

Conclusion and policy implications

Accurate and comprehensive NTM data is essential for analyzing and understanding the role of NTMs in global trade. NTMs are increasingly shaping trade flows, with some countries substituting tariffs with NTMs as trade barriers (Niu et al., 2018, 2020). Unlike tariffs, NTMs are subject to less stringent WTO oversight, making them more difficult to monitor and regulate. While primarily imposed unilaterally by importing countries, NTMs can have discriminatory effects, disproportionately impacting certain exporters. This lack of transparency underscores the need for reliable NTM data to assess their economic impacts accurately. To achieve this, broad country coverage is crucial. AVE estimates rely on cross-country variation to capture the true effects of NTMs, and without good coverage, bias in AVE estimates is likely to increase. Measurement error and biased AVE estimation limits the reliability of using AVE estimates in policy analysis or to inform policy changes. Since NTMs vary significantly across countries due to differences in regulatory frameworks and enforcement practices, a dataset with extensive country coverage is necessary to account for these discrepancies and provide a more precise evaluation of their trade effects. Measuring time variation in NTMs is also crucial for analyzing the effects of key trade events, such as the implementation of trade agreements and trade wars (Cadot and Gourdon, 2016). When studying trade agreements, it is important to have not only comprehensive country coverage, but also comprehensive time coverage to capture variation in the effects of NTMs over time (Niu et al., 2018). NTMs play an important role in trade wars. For example, Chen, Hsieh, and Song (2022) highlight the impact of NTMs in the U.S.–China trade war and find that NTMs accounted for nearly 50% of the overall decline in Chinese imports from the U.S. However, their study infers the effects of NTMs by attributing trade impacts to unexplained trade costs rather than directly using NTM data. The lack of comprehensive NTM data and subsequent challenges in measuring AVEs introduces additional measurement errors by attributing impacts of changes in trade values to NTMs that might actually arise from other sources. This underscores the need for improved NTM data to ensure more accurate and reliable analysis of trade policy effects. Finally, this study is particularly relevant for research on the agricultural sector, as SPS and TBT measures often significantly affect agricultural trade (Gourdon, Stone, and van Tongeren, 2020; Disdier, Fugazza et al., 2020).

Findings from this research demonstrate that data coverage is crucial, and while the TRAINS database is more accurate than WTO notifications for the data it collects, TRAINS does not always provide sufficient country and time coverage. We assess whether WTO data can produce similar AVE estimates to those derived from TRAINS when conditioned on the same country and time coverage. Our results indicate that even when controlling for similar country and time coverage, AVE estimates frequently do not align across datasets. Among AVEs differing from zero, the meandifference between TRAINS and WTO AVE estimates is approximately 57.35 percentage points for SPS measures and -178.51 percentage points for TBT measures. However, since the mean may be skewed by extreme values, particularly in the TBT case, it does not fully capture the extent of these discrepancies. Additionally, we identify numerous instances where AVEs are null in TRAINS but nonzero in WTO, and vice versa, which also raises concerns. These discrepancies can lead to inconsistent policy interpretations and biased AVE estimates.

In sum, our findings suggest that the assumptions required when using WTO notification data introduce substantial measurement errors in recorded NTMs and ultimately bias AVE estimates. Notably, the magnitude of these differences varies depending on the product, NTM type, importing country, and year. Moreover, for both SPS and TBT measures, discrepancies in AVE estimates tend to decrease with the exporter's share in the global market for that good. Although we attempt to identify conditions under which AVE estimates from both datasets align more closely, this remains challenging due to the multiple sources of measurement error inherent in WTO notification data, which are difficult to fully disentangle. A limitation of this study is low country coverage, as we only include nine importing countries, resulting in relatively low variation across countries in the gravity model. This choice was driven by the limited overlap of countries covered in both datasets across successive years and their geographic proximity, which was important for our instrumental variable approach.

These findings emphasize the need for improved data collection in NTM reporting to enhance consistency and reliability of AVE estimates. A more comprehensive and accurate dataset is essential not only for understanding the effects of NTMs on trade but also for evaluating their broader economic implications. Given the challenges with NTM data that we explore in this analysis, some studies resort to estimating AVEs of NTMs without using direct NTM data, particularly in the services sectors (Fontagné, Guillin, and Mitaritonna, 2011; Guillin, 2013). They infer AVEs from all unexplained variation in trade costs. However, this approach relies on the assumption that any trade value not explained by the gravity model or other known trade regulations, such as tariffs, is solely attributed to the effects of NTMs. This assumption can be problematic, as it overlooks other potential factors that influence trade.

The findings of this study highlight key challenges when using different NTM data sources to assess their trade effects, underscoring the importance of data accuracy when estimating AVEs. Specificaly, we show that although the WTO notification data has better coverage than the TRAINS, inaccurate reporting in the WTO can lead to significant bias in AVE estimations. Given the central role of NTMs in trade policy, this work raises awareness of data limitations and encourages a more informed use of these datasets. Identifying measurement issues provides a foundation for future improvements in NTM data, contributing to more accurate and reliable trade policy analysis.

References

Adarov, A., and M. Ghodsi. 2023. "Heterogeneous effects of nontariff measures on cross-border investments: Bilateral firm-level analysis." *Review of International Economics* 31:158–179.

Baier, S.L., and J.H. Bergstrand. 2009. "Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation." *Journal of International Economics* 77:77–85.

Beghin, J.C., M. Maertens, and J. Swinnen. 2015. "Nontariff measures and standards in trade and global value chains." *Annu. Rev. Resour. Econ.* 7:425–450.

Blöthner, T., and M. Larch. 2022. "Economic Determinants of Regional Trade Agreements Revisited Using Machine Learning." *Empirical Economics* 63:1771–1807.

Bratt, M. 2017. "Estimating the bilateral impact of nontariff measures on trade." *Review of International Economics* 25:1105–1129.

Cadot, O., and J. Gourdon. 2016. "Non-tariff measures, preferential trade agreements, and prices: new evidence." *Review of World Economics* 152:227–249.

Cadot, O., J. Gourdon, and F. Van Tongeren. 2018. "Estimating ad valorem equivalents of non-tariff measures: Combining price-based and quantity-based approaches." *OECD Trade Policy Papers*, pp. .

Chen, T., C.T. Hsieh, and Z.M. Song. 2022. "Non-Tariff Trade Barriers in the US-China Trade War." *NBER Working Paper*, pp. .

Conte, M., P. Cotterlaz, and T. Mayer. 2022. "The CEPII Gravity database." *CEPII Working Paper*, pp. .

De Melo, J., and A. Nicita. 2018. "Non-tariff measures: Data and quantitative tools of analysis." *Ferdi Working paper*, pp. .

Disdier, A.C., M. Fugazza, et al. 2020. "A practical guide to the economic analysis of non-tariff measures.", pp. .

Essaji, A. 2008. "Technical regulations and specialization in international trade." *Journal of International Economics* 76:166–176.

Fell, J., and A. Duver. 2024. "Non-tariff measures: a methodology for the quantification of bilateral trade effects of policy measures at a product level." *Applied Economics* 56:4374–4388.

Ferrantino, M.J. 2006. "Quantifying the trade and economic effects of non-tariff measures." *OECD Trade Policy Papers*, pp. .

Fontagné, L., A. Guillin, and C. Mitaritonna. 2011. "Estimations of tariff equivalents for the services sectors."

Ghodsi, M., J. Grübler, O. Reiter, and R. Stehrer. 2017. "The evolution of non-tariff measures and their diverse effects on trade." Working paper, Wiiw Research Report.

Ghodsi, M., O. Reiter, and R. Stehrer. 2015. "Compilation of a Database for Non-Tariff Measures from the WTO Integrated Trade Intelligence Portal (WTO I-TIP)." *Paper written in the PRONTO project, work in progress*, pp. .

Ghodsi, M., and R. Stehrer. 2022. "Trade policy and global value chains: tariffs versus non-tariff measures." *Review of World Economics* 158:887–916.

Gourdon, J., S. Stone, and F. van Tongeren. 2020. "Non-tariff measures in agriculture." *OECD Food, Agriculture and Fisheries Papers*, pp. .

Grübler, J., and O. Reiter. 2021. "Characterising non-tariff trade policy." *Economic Analysis and Policy* 71:138–163.

Guillin, A. 2013. "Assessment of tariff equivalents for services considering the zero flows." *World Trade Review* 12:549–575.

Ing, L.Y., S.F. de Cordoba, O. Cadot, et al. 2016. "Non-tariff Measures in ASEAN." In *Economic Research Institute for ASEAN and East Asia and United Nations Conference on Trade and Development*.

Kee, H.L., and A. Nicita. 2022. "Trade fraud and non-tariff measures." *Journal of International Economics* 139:103682.

—. 2016. "Trade frauds, trade elasticities and non-tariff measures." In *5th IMF-World Bank-WTO Trade Research Workshop, Washington, DC, November*. vol. 30.

Kee, H.L., A. Nicita, and M. Olarreaga. 2009. "Estimating trade restrictiveness indices." *The Economic Journal* 119:172–199.

Kravchenko, A., A. Strutt, C. Utoktham, and Y. Duval. 2022. "New Price-based Bilateral Ad-valorem Equivalent Estimates of Non-tariff Measures." *Journal of Global Economic Analysis* 7.

Lee, J.W., and P. Swagel. 1997. "Trade barriers and trade flows across countries and industries." *Review of economics and statistics* 79:372–382.

Mao, R., Y. Liu, and X. Wang. 2023. "Economic and environmental impacts of agricultural non-tariff measures: evidence based on ad valorem equivalent estimates." *International Food and Agribusiness Management Review* 26:379–396.

Niu, Z., C. Liu, S. Gunessee, and C. Milner. 2018. "Non-tariff and overall protection: evidence across countries and over time." *Review of World Economics* 154:675–703.

Niu, Z., C. Milner, S. Gunessee, and C. Liu. 2020. "Are nontariff measures and tariffs substitutes? Some panel data evidence." *Review of International Economics* 28:408–428.

Orefice, G. 2017. "Non-tariff measures, specific trade concerns and tariff reduction." *The World Economy* 40:1807–1835.

Porto, G.G. 2018. "Labor market effects of non-tariff measures in Latin America." *UNCTAD*, pp. 195.

Raimondi, V., A. Piriu, J. Swinnen, and A. Olper. 2023. "Impact of global value chains on tariffs and non-tariff measures in agriculture and food." *Food Policy* 118:102469.

Rau, M.L., and A. Vogt. 2017. "Data concepts and sources of non-tariff measures (NTMs)–an exploratory analysis." *Bern: World Trade Institute. September. http://www. etsg. org/ETSG2017*, pp. .

Trefler, D. 1993. "Trade liberalization and the theory of endogenous protection: an econometric study of US import policy." *Journal of political Economy* 101:138–160.

UNCTAD. 2018. *UNCTAD TRAINS: The Global Database on Non-Tariff Measures User Guide (2017, Version 2).*

Figure 1: Density plot of the difference in AVEs for both SPS and TBT measures (in percentage points).

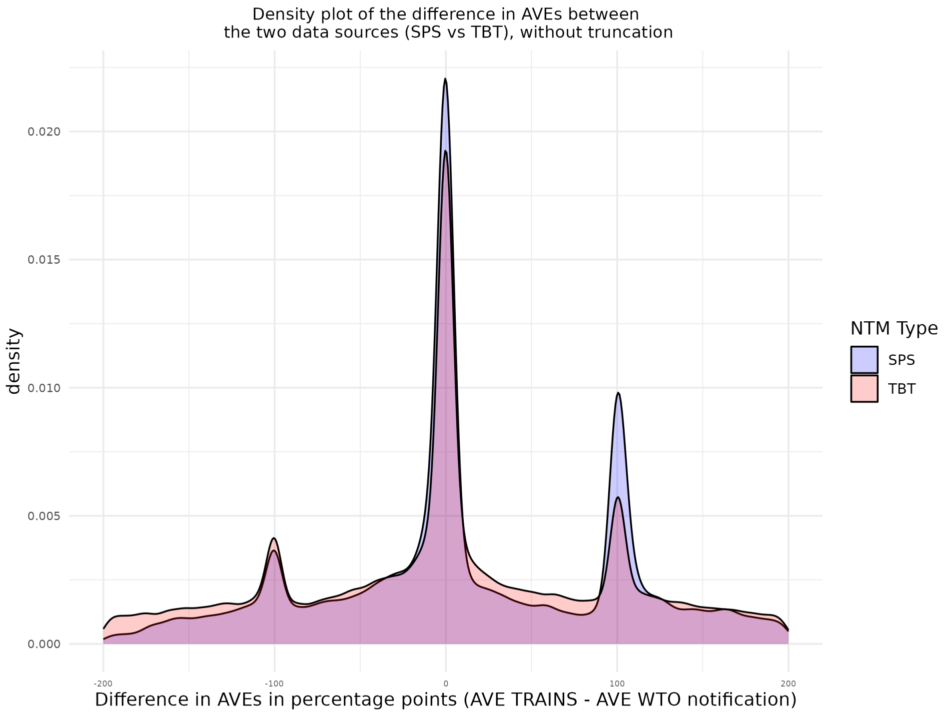

Figure 2: Density plot of the difference in AVEs for both SPS and TBT measures (in percentage points).

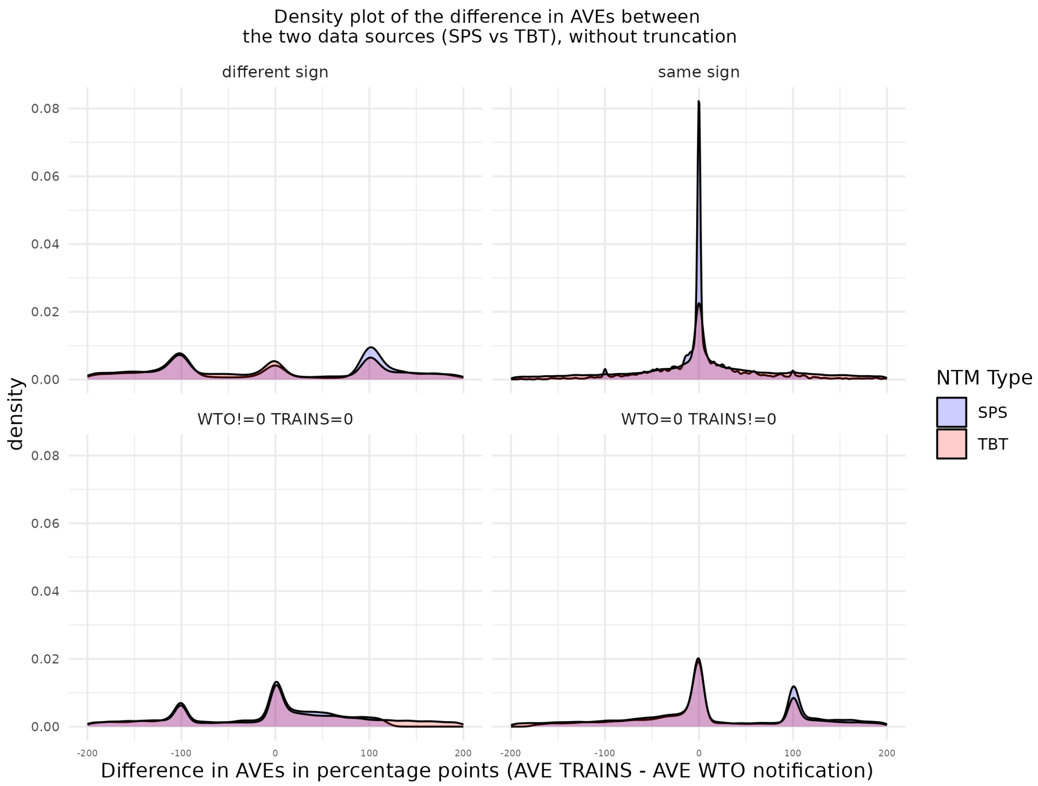

Table 1: Comparison of WTO and TRAINS AVE estimates for SPS measures (0, NAs, and ≠0).

SPS AVEs		TRAINS		
		0	NA	≠ 0
WTO	0	20.10%	2.47%	12.36%
	NA	31.75%	10.48%	5.29%
	≠ 0	3.36%	0.88%	13.30%

Table 2: Comparison of WTO and TRAINS AVE estimates for TBT measures (0, NAs, and ≠0).

TBT AVEs		TRAINS		
		0	NA	≠ 0
WTO	0	23.96%	6.42%	15.45%
	NA	0.85%	3.22%	1.15%
	≠ 0	13.93%	3.69%	31.32%

Table 3: Comparison of WTO and TRAINS AVEs and corresponding NTM count values.

Data	NTM	NTM = 0 AVE = NA	NTM = 0 AVE = 0	NTM = 0 AVE ≠ 0	NTM >0 AVE = NA	NTM >0 AVE = 0	NTM >0 AVE ≠ 0
TRAINS	SPS	10.83%	54.90%	0%	3.00%	0.32%	30.95%
	TBT	9.83%	38.26%	0%	3.49%	0.48%	47.93%
WTO	SPS	45.73%	34.71%	0%	1.80%	0.22%	17.54%
	TBT	2.46%	45.46%	0%	2.76%	0.37%	48.95%

Table 4: Summary Statistics for the Difference in AVEs between WTO Notifications and TRAINS.

Diff in AVE	Mean	Median	Sd	1st Qu.	3rd Qu.
SPS	57.35	0.78	157.67	-20.36	107.22
TBT	-178.51	-0.01	1181.73	-248.10	154.38

Note: The summary statistics are calculated from the obtained AVEs, excluding observations where AVEs are null or NA in both datasets.

Table 5: Regression results: difference in AVE estimates for SPS and TBT measures.

	SPS	TBT
Dependent Var.:	abs(diff_AVE_SPS)	abs(diff_AVE_TBT)
Constant	140.3*** (7.141)	528.4*** (27.47)
Year2014	-15.89** (1.286)	-59.17*** (4.573)
Year2015	-16.54** (1.539)	6.446 (3.521)
Year2016	-8.257* (1.433)	6.837 (5.089)
tot_SPS	-6.43e-7 (1.02e-5)	-
Reporter_GDP_per_cap	-0.0019 (0.0010)	0.0089* (0.0022)
Partner_GDP_per_cap	-1.82e-5 (1.35e-5)	2.25e-5 (3.28e-5)
import_importance	53.88 (48.47)	204.8 (166.4)
import_share	-0.8226 (0.4901)	-9.553* (2.323)
export_share	-0.661* (0.1352)	-3.185** (0.3955)
tot_TBT	-	-1.57e-6 (1.08e-5)
--------------------	------------------	------------------
S.E.: Clustered	by: Repo. & Prod. & Year	by: Repo. & Prod. & Year
Observations	1,391,364	4,776,680
R2	0.00798	0.00234
Adj. R2	0.00797	0.00234

Coefficient values and their standard errors (in parentheses) for the regressions in Equation 12. The first column presents the dependent variable as the difference in AVE for the SPS type, while the second column presents the dependent variable as the difference in AVE for the TBT type. Significance levels are denoted as follows: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1