Nontariff Barriers and the Elasticity of Customs Duties Evasion*

Rafael Bressan† Enlinson Mattos ‡

2023-07-10

Abstract

This paper leverages an exogenous tariff reform in Brazil with rich administrative data to document the elasticity of customs duties evasion (ECDE) for the universe of importers in Brazil. Our contributions are twofold. First, as our data reflects the effectively applied tariff rates on each import, our estimates are larger than that in the literature. We find an increase of 1.98% in evasion of value for each percentage point increase in the tariff rate and 0.97% for evasion of quantities. Second, we document that nontariff barriers (NTB) act to reduce ECDE up to 0.76 and -0.09, for value and quantity respectively.

keywords: elasticity of custom duties evasion, nontariff barriers, trading companies, tax enforcement

1 Introduction

Modern tax and customs administrations are increasingly concerned about the importance of policies that promote business development and, at the same time, ensure voluntary compliance with tax obligations. Of course, not all import declarations are inspected, as this would not only be infeasible due to limited state capacity but would also imply unnecessary audits on low-risk importers. Promoting voluntary tax compliance requires

*We have benefited from comments and suggestions from Annika Nivala, Bruno Giovannetti, Emanuel Ornelas, Evelina Gavrilova-Zoutman, Gabriel Ulyssea, João Paulo Pessoa, Joel Slemrod, Vladimir Ponczek, Vladimir Teles, and participants at the 78th Annual Congress of the IIPF. Rafael Bressan thanks the support of the Brazilian Revenue Service (RFB). All data work for this project involving confidential taxpayer information was made at RFB facilities on RFB computers by RFB employees, and at no time was confidential taxpayer data ever outside of the RFB computing environment. The views and opinions presented in this paper reflect those of the authors. They do not necessarily reflect the views or the official position of the Brazilian Revenue Service. All results have been reviewed to ensure that no confidential information is disclosed. Any errors are our own.

†Sao Paulo School of Economics, FGV/EESP and Brazilian Revenue Service, RFB. E-mail: Rafael.Bressan@rfb.gov.br

‡Sao Paulo School of Economics, FGV/EESP. E-mail: enlinson.mattos@fgv.br

developing a risk-based approach to the selection of which imports to audit. While the literature on customs duties evasion has explored the level of tariffs as a relevant risk factor (Fisman and Wei 2004), little has been investigated about other characteristics of imports that may have a modulating effect on the elasticity of evasion in relation to tariff rates. In particular, the questions we ask are as follows: (i) Is there an impact of customs policies such as nontariff regulation, on customs duties evasion? and (ii) What are the mechanisms through which this regulation operates for firms willing to take the risk of misreporting quantities/values of imports?

To answer these research questions, we use for the first time a rich confidential administrative dataset, and we exploit the exogenous implementation in 2004 of two taxes levied on Brazilian imports to document the elasticity of customs duties evasion. Our data constitute a panel with three dimensions: product, country of origin, and year. We explore the within variation in a fixed effects panel data model to identify our parameter of interest. This empirical specification, based on our theoretical model presented in section 2, is able to identify an ECDE of 2.43 (1.45) for values (quantities). At the same time, we identify that nontariff barriers are a strong reducer of this elasticity. The ECDE decreases to 0.91 (values) and -0.06 (quantities) after interacting NTB with tariff rates. Furthermore, we consider a machine learning model with causal forests to corroborate our choice of NTB as the mains source heterogeneity in the ECDE.[1] We extend the model proposed in Mishra, Subramanian, and Topalova (2008) to include the tariff rate as part of the evasion cost. This is consistent in the Brazilian case, where the penalty for misclassifying imported products is proportional to the evaded value of taxes. From this theoretical model, we predict that the elasticity of evasion must be positive with respect to tariffs. More importantly, this elasticity may be different for the same tariff level if, for instance, an importer faces a higher cost of evasion due to a greater perception of enforcement (lower elasticity), or if its evasion technology is worse.

To empirically investigate the predictions of the theoretical model and interpret our candidates for evasion risk factors we use administrative data containing information on import values, quantities, importing and exporting companies, tariff rates effectively applied, and audits. We explore two margins: reported values and quantities. We show how the use of our detailed data on all taxes levied on imports is crucial for the estimation of evasion elasticity. For example, UNCTAD/TRAINS data for Brazil do not include some taxes on imports and may not capture the true evasion incentives importers have[2].

[1]The presence of trading companies is estimated to be one of these salient sources, but without a statistically significant effect in our linear model specification. Our previous version of the paper brings the results for trading companies and are available upon request.

[2]We are referring to data extracted from the World Integrated Trade Solution (WITS), which contains TRAINS tariff data. Although the TRAINS database makes available 'applied tariff' data, those are incomplete in Brazil's case. other taxes are levied at border arrival but not contemplated by TRAINS. These taxes are imposed in cascade, that is, any error/omission in a tax rate can produce a very large bias in the estimation of effective applied tariff, as explained in our identification strategy.

Our identification assumption for the ECDE relies on heterogeneous tariff exposure of imports after the introduction of two new taxes, PIS and COFINS. These taxes affected virtually all imports of goods and services and were the main component of the 2003 - Tax Reform[3](Werneck 2006). They were introduced abruptly; and took effect less than six months after the tax reform was approved at rates of 1.65% and 7.6%, respectively. The creation of these taxes in such a short period, implied a very limited room for tax avoidance.[4] Considering the wide tax base of PIS and COFINS, all products regardless of the country of origin were affected by this tax reform which reinforces the robustness of our empirical strategy to the selection problems.

Our results for the ECDE have magnitudes comparable to the literature for evasion in values.[5] In addition, our findings confirm that the possibility of reclassification into products with a lower tariff rate contributes to the explanation of evasion. A 1 percentage point reduction in the rate of a similar product leads to an increase of 0.54% in evasion of values and 0.36% in quantities. In this context, an ECDE of 2.43 for value would lead us to a revenue-maximizing rate of 41.15%, whereas in 2021 approximately 22.03% of all Brazilian imports were subject to tariffs higher than this optimal value.

Next, we investigate the main risk factor in the evasion of customs duties that have a modulating effect on the ECDE. We explore a specific tax treatment present in approximately 25% of Brazilian imports, import licensing. An import license is an electronic form the importer must file before the shipment of goods in case of selected goods. The government authorizes imports upon verification of compliance with legal and administrative regulations. Although there are some tax treatments for imports related to tariffs (tariff agreements, drawback regime, and exemptions), import licenses are considered a nontariff barrier (NTB) exposed to a constant share of imports in our period. Second, we document how the elasticity of evasion is affected by the prominent presence of trading companies in the import process.

Our last set of results shows that exposure to import licensing reduces the elasticity of evasion by approximately 1.84 and 1.91 percentage points in value and quantity, respectively. Importers perceive the bureaucratic process of providing information on the product to be imported as an extra level of enforcement on imported goods and services. We also investigate the tax authority's auditing policies on imports.

We contribute to the literature in three main aspects. First, even though estimating ECDE is not a novelty (Fisman and Wei 2004; B. S. Javorcik and Narciso 2008; Mishra,

[3]Constitutional Amendment No. 42, of December 19, 2003. Available at planalto.gov.br Taxes began to be collected on May 1, 2004.

[4]Only a few products had a zero tax rate (hydraulic motors, cylinders, parts of harvesting machines, and others have zero rate and machines, equipment, and films without national equivalents for the cinematographic industry) or were exempt (books, periodicals, and papers with no national equivalent, for printing books, newspapers, and periodicals).

[5]Using only UNCTAD/TRAINS data to calculate tariff rates and tax amounts for 6-digit HS instead of our administrative data would lead to a lower and non-significant effect on quantity evasion.

Subramanian, and Topalova 2008; Bouet and Roy 2012), this paper explores rich administrative data with information on all 8-digit encoded products (importer, legal purchaser, effective tariff rates, country of origin, border entry point, and administrative treatments), which is matched with the 6-digit data from the UN COMTRADE Harmonized System (HS) to perform a granular evasion-gap analysis.[6] Moreover, we investigate an exogenous tariff shock (Demir and B. Javorcik 2020) implemented in 2003 in Brazil to identify the evasion responses.

Second, unlike the literature on nontariff measures (NTM), (WTO 2012), which focuses on their role as a substitute for tariffs in the welfare gains from trade liberalization (Grübler and Reiter 2021; Egger et al. 2015; Ferrantino 2006; Kinzius, Sandkamp, and Yalcin 2019; Bao and Qiu 2010), we discuss how the import licensing can modulate customs duties evasion as an evasion risk factor, connecting to the literature on pre-shipment inspections (Yang 2008; Anson, Cadot, and Olarreaga 2006).[7]

Third, different from previous literature, we control for the nature of importers in our regression models. Out of more than 20 thousand importers in the postshock period, 548 represent more than 70% of all transactions and more than 54% of the total Free-On-Board (FOB) value. As we have very detailed information on the importers, we can identify the trading companies that operate in the country. Trading companies can serve as intermediaries, specialize in international trade, and presumably have a body of experts on their staff (Ahn, Khandelwal, and Wei 2011). A common aspect of this literature is that these trading companies choose the market to establish operations based on trade costs, with tariffs being a prominent example of such a cost (Antras and Helpman 2004).[8]

Our results speak to the modern tax and customs authority, which focuses its audits on high-risk importers. We provide a robust identification of the elasticity of customs duties evasion and the main important characteristic that can be used for risk assessment, import licensing. And we do control for the share of trading companies in charge of each import. Based on our results, customs inspections could be focused on imports other than those subject to import licensing due to their negative effect on the ECDE, but imports associated to a trading company does not seem to be associated with larger evasion.

The remainder of this paper is organized as follows. The next section provides a theoretical framework to support our empirical findings. In section 3, we detail the datasets used in

[6] Kume, Piani, and Miranda (2011) and Sousa, Tannuri-Pianto, and Santos (2008) consider COMTRADE data for Brazil.

[7] *Ad Valorem* Equivalents (AVE) are estimated for NTM to derive trade elasticities which in turn are the key statistics to make the calculations of gains from trade. Kee and Nicita (2016) is one of the few studies that relate NTM and trade frauds. They study the role of nontariff measures in affecting trade flows and frauds that cause large trade discrepancies.

[8] As we investigate evasion of customs duties, we focus on goods and services quantities and prices (aggregated at 8-digit HS system) reporting responses to tariff changes rather than firm-level data analysis as in Ahn, Khandelwal, and Wei 2011. This does not allow us to identify the role of trading companies in evasion activities, but to control for their presence on the customs duties evasion of the goods and services.

this article and provide descriptive statistics. Section 4 presents the empirical specification, while the results found are presented in section 5. Finally, we conclude in section 6.

2 Theoretical Model

We present a simple model that captures both the impact of enforcement and tax-related knowledge on tariff evasion elasticity. Our model draws largely from Mishra, Subramanian, and Topalova (2008), Slemrod (2001), and Yang (2008) and assumes that importers are rational agents who optimize behavior as in Allingham and Sandmo (1972). This framework guides us to interpret the effect of NTB on evasion but controlling for trading companies with different compliance costs related to tax evasion.

Assume a cost of evasion as follows

$$C = C(\gamma, E)$$

where γ is the smuggling fraction and E is a measure of enforcement quality. The following assumptions are made about this cost[9]: i) $C(\gamma, E) \geq 0$, cost of evasion is non-negative; ii) $C_1(\gamma, E) > 0$, the marginal cost of evasion is positive; iii) $C_2(\gamma, E) > 0$, the cost is increasing in enforcement; iv) $C_{11}(\gamma, E) > 0$, the convexity of costs to evade and; v) $C_{12}(\gamma, E) > 0$, the marginal cost of evasion is increasing in enforcement[10].

One concern we have with that specification is whether the cost of evasion depends on the tariff rate itself, as in Yitzhaki (1974). One example of such a scenario would be if customs authorities devote a greater deal of effort to auditing high-taxed products. Brazilian legislation establishes that if an importer is caught illegally undervaluing an imported good, for example by forging the invoice or smuggling the good, the penalty depends only on the amount by which the good is undervalued. However, if evasion takes place through misclassification, which we aim to control, the penalty is proportional to the evaded tariff.

To investigate this possibility, we bring Figure 1 that shows no sign of a structural relationship between import audits and total tariff exposure with an average overall audit rate of 18.99%, thus, the probability of evasion detection is not related to the tariff rate.

Considering the probability of detection and punishment is not related to tariffs, we can specify a cost function that adds linearly a cost-term proportional to the tariff rate. With a fixed probability of punishment, $\bar{p}$, and a penalty parameter θ, the cost an importer's

[9]Such assumptions in cost of evasion are common in the literature, (Cowell 1990; Kaplow 1990; Mayshar 1990; Hines and Rice 1994; Slemrod 2001; Yang 2008; Chetty 2009; Mishra, Subramanian, and Topalova 2008)

[10]We use the following notation to represent partial derivatives, $\partial C/\partial \gamma = C_1$, since γ is the first argument to the cost function. Second-order derivatives will have the same notation with an additional index, $C_{12} = \partial^2 C/\partial \gamma \partial E$.

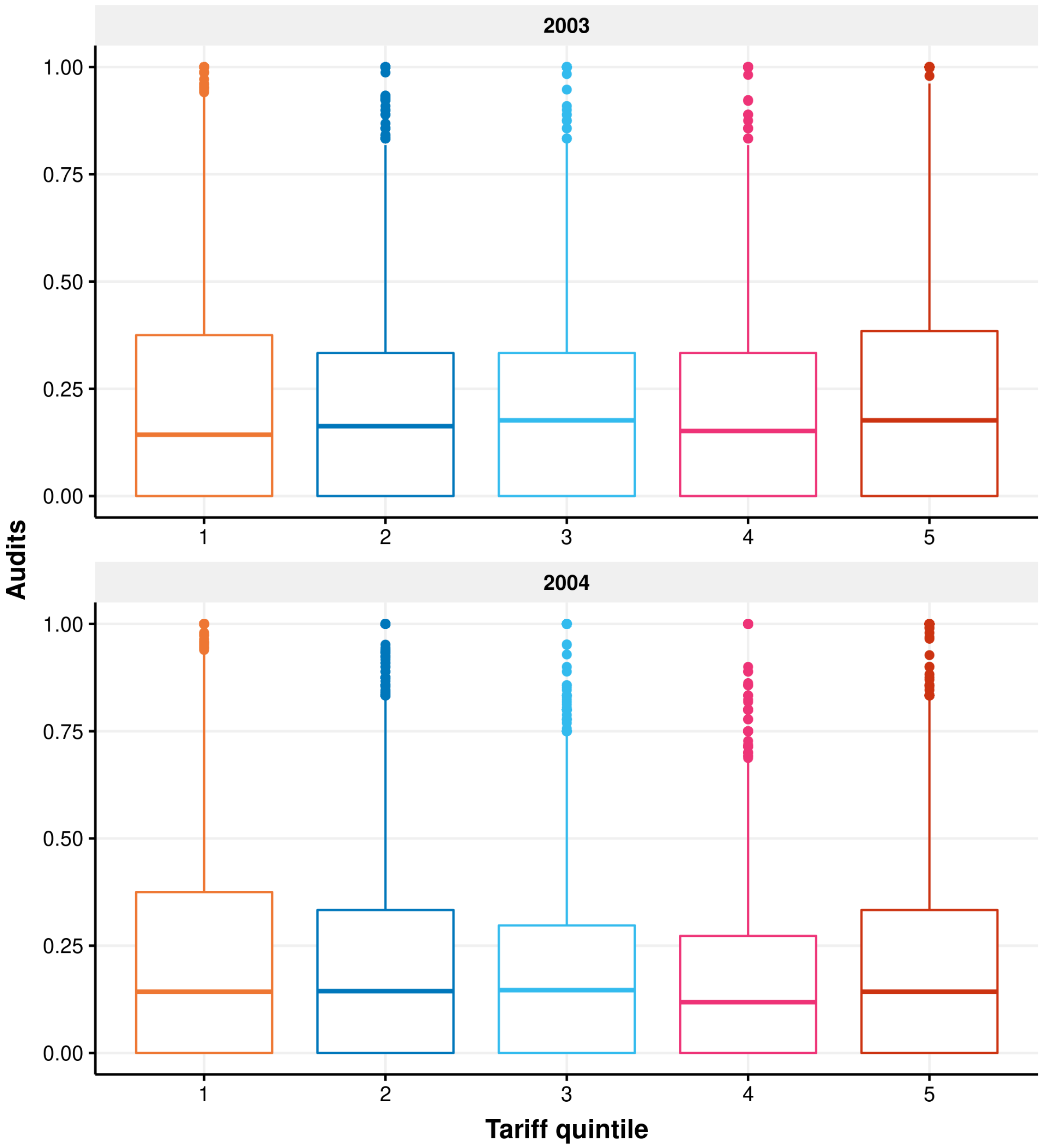

Figure 1: Proportion of imports audited by tariff quintile and year. Distribution by partner country and HS 6-digit products. Box-plot histogram that captures conditional quantiles from 25th to 75th quantiles. Vertical lines denote the conditional median inside those quantile boxes.

faces when caught evading is proportional to the duty evaded. Our cost of evasion has the following specification:

$$\tilde{C}(\gamma, E, T) = C(\gamma, E) + \bar{p}\theta\gamma M \cdot T \tag{1}$$

where the "smuggling" fraction is $0 \leq \gamma \leq 1$, E is a measure of enforcement quality other than the detection probability, T is the tariff rate and, $C(\gamma, E)$ is the cost of evasion function.[11] Therefore, quality of enforcement, E, captures any additional enforcement capacity to the tax authority not directly related to tariffs. For instance, E can be associated with product characteristics that facilitate its valuation, (Rauch 1999; B. S. Javorcik and Narciso 2008), or information concerning the product, producer, and terms of trade provided prior to the shipment, i.e., import licensing process or pre-shipment inspection.

Assume a representative firm that imports inelastically an amount M.[12] This firm maximizes the net benefit of evasion, the facilitate between the duties avoided and the cost of evasion, given by: $B = \gamma M \cdot T - \tilde{C}(\gamma, E, T)$, by choosing an optimal "smuggling" fraction γ to evade customs duties. The first-order condition, which implicitly defines the level of optimal evasion as a function of the tariff rate T, and enforcement E, is given by:

$$M \cdot T(1 - \bar{p}\theta) = C_1(\gamma, E). \tag{2}$$

This equation represents the demand side for evasion and has the simple interpretation of evasion's marginal cost must equal its marginal benefit. Differentiating equation (2) with respect to T we have the slope of this demand function:

$$\frac{\partial \gamma}{\partial T} = \frac{M(1 - \bar{p}\theta)}{C_{11}(\gamma, E)} > 0 \tag{3}$$

and the assumption of convex costs in evasion implies that $\partial\gamma/\partial T > 0$, that is, the semi-elasticity of duties evasion with relation to the tariff rate is non-negative. Therefore, as long as $\bar{p}\theta < 1$, which is a plausible assumption according to our data, the evasion elasticity is damped by a factor of $1 - \bar{p}\theta$ and, is positively related to the tariff exposure.[13]

Next, if we differentiate equation (2) with relation to enforcement we have how the optimal evasion level is affected by enforcement quality. The assumptions (i) of convex evasion

[11] This smuggling can occur in different forms: (i) Underpricing, (ii) misclassification or even (iii) concealing quantities.

[12] This assumption is made for simplicity, as in Mishra, Subramanian, and Topalova (2008). In their work, they show that as long as the absolute value of the elasticity of imports with respect to *tariffs* is less than one the results are unchanged. In Appendix A, we show that, with minor assumptions, if imports respond to enforcement policy, our results also hold.

[13] Using our administrative dataset, the probability of audit is close to 19.0%, while the penalty term in Brazil is typically 75%. This leads to a typical $\bar{p}\theta \approx 0.14$.

costs and (ii) increasing marginal costs with enforcement quality (i.e. $C_{12}(\gamma, E) > 0$) ensure that $\partial\gamma/\partial E < 0$.

$$\frac{\partial\gamma}{\partial E} = \frac{-C_{12}(\gamma, E)}{C_{11}(\gamma, E)} < 0 \tag{4}$$

Therefore, evasion is increasing in tariff rate and decreasing in the quality of the enforcement. This framework can also be used to understand how enforcement modifies the elasticity of evasion with respect to the tax rate. Taking the derivative of evasion elasticity in relation to enforcement, we have

$$\frac{\partial^2\gamma}{\partial T\partial E} = -\frac{M(1-\bar{p}\theta)}{C_{11}^2(\gamma, E)}\frac{\partial C_{11}(\gamma, E)}{\partial E} \tag{5}$$

with a sufficient condition for $\partial^2\gamma/\partial T\partial E < 0$ is $\partial C_{11}(\gamma, E)/\partial E > 0$, i.e., the higher the level of enforcement, more convex is the cost of concealing imports. Hence, if we further assume that the marginal cost of smuggling increases more (convex), the higher the enforcement level, we also expect that the elasticity of evasion is reduced by better government monitoring. This result is crucial in interpreting the role of NTB on evasion.

2.1 Nontariff barriers

Nontariff measures (NTM) comprise all policy measures other than directly applied tariffs that have an economic effect on international trade. An NTB is a specific case of NTM where the effect is known to reduce trade, thus, imposing a barrier on free trade. These measures can be broadly divided into two groups. Technical measures include regulations such as standardization, testing, and certification procedures. They are primarily sanitary and phytosanitary (SPS) and Technical Barriers to Trade (TBT) measures. There are nontechnical measures, which include quantitative restrictions (quotas, non-automatic import licensing), price measures, forced logistics or distribution channels, among others.[14]

Import licensing is a nontariff barrier (NTB) levied upon approximately 25% of Brazilian imports, as shown in Table 4.[15] Our choice for this specific NTB relies on the fact that it is the prevailing barrier in Brazil during the period of our analysis. Moreover, the only other nontariff barrier present in our administrative dataset is the anti-dumping measure, which represents only 0.003% of all imports (Table 4). We are interested in investigating how the imposition of this NTB interacts with tariff rates to potentially modify the elasticity of evasion in the same way our theoretical framework modeled quality of enforcement.

The import license is an electronic document through which the government authorizes

[14]Details on NTM can be found at https://www.oecd.org/trade/topics/non-tariff-measures/.

[15]In this text, whenever we refer to import licensing we mean non-automatic import licensing. In Brazil, all imports are subject to import licensing and the default procedure is automatic unless stated otherwise in regulation.

the import upon verification of compliance with legal and administrative rules. It is required when the import is subject to the approval of one or more consenting bodies (Verification Authorities), and it depends on the product, country of origin, condition (used or brand new), and taxation regime. This is a bureaucratic process that may take up to 60 days to be completed and before that, the importer is not authorized to start the shipment process. The importer is subject to penalties up to 30% of the import CIF (costs, insurance, and freight) value if this rule is not accomplished.[16]

Henceforth, the imposition of a nontariff barrier could plausibly result in an increase in the quality of enforcement, facilitated by the provision of detailed information access. In fact, importers may perceive the entire process of obtaining an import license, contingent upon the submission of comprehensive information to the government, as an elevated level of enforcement.

3 Data

We consider two primary sources of data. The first brings reported imports (and exports) by Brazilian firms (from all of its partners) for 2003 and 2004, aggregate at HS 6-digit, from WITS/COMTRADE database. We label this dataset the *"Trade sample"*. Second, we use a confidential administrative dataset that comprises all Brazilian imports at the transaction level, our *"Administrative sample"*. This data comes from the Ministry of Economy, extracted from the SISCOMEX information system, which registers any import or export process in Brazil with 8-digit product codes following the NCM standard.[17] The import values are FOB, thus, directly comparable to export data from the trade sample. Using the detailed information of the administrative sample we can compute the effective average tariff rate for 6-digit HS codes by country-product-year.

The administrative sample contains information on the US dollar import values, quantities, importer firm, purchaser, exporter, most favored nation tariff level, effectively applied tariff rate, transaction characteristics such as administrative treatment (e.g. exemption status, preferential trade agreement, import license, transaction audit, etc.).[18] Moreover, import duty is not the only tariff that imports are subject to in Brazil. The administrative sample we have is the only source that allows us to compute all other border duties, such as the tax on industrialized products, PIS, and COFINS on imports.

This administrative dataset also allows us to control for the nature of importing firms. We

[16] Cost, insurance, and freight. The free-on-board value is added by the insurance and freight costs.

[17] The Integrated Foreign Trade System - SISCOMEX is an administrative information system that integrates the activities of registration, monitoring, and control of foreign trade operations in Brazil. It was instituted by Decree No. 660, of September 25, 1992. However, the SISCOMEX-imports module went live only on January 1, 1997. NCM stands for *"Nomenclatura Comum do Mercosul"*, Mercosur Common Nomenclature and is an extension of the 6-digit codes from the Harmonized System

[18] The figure of the purchaser is prescribed by Brazilian law as the firm or person who is the ultimate merchandise purchaser and uses a trading company as an intermediary to access international markets. In that case, the importer is the trading company.

consider two criteria to characterize them as trading companies operating in Brazil. First, Brazilian customs regulation deems as a trading company any firm that has a registered purchaser in its SISCOMEX profile. We can identify this rule in the administrative data whenever the importer's identification number is different from the purchaser's. In that case, the import transaction is flagged as being executed by an *intermediary* company. Those companies operate on behalf of their clients, which may be many and from different industries and this leads to a very distinct profile of transactions. The second aspect we consider to characterize a firm as trading relates to its pervasiveness in international trade. An *expert* company is identified if, for a given year, it imported more than twenty different products at the 8-digit classification code and each product imported more than twenty times.[19] A trading company is defined as either an intermediary or an expert, with the transactions flagged as being carried out by an expert only if the importer is not operating in the intermediary capacity.

3.1 Evasion quantities

Inspired by the work of Fisman and Wei (2004), we compute two measures of value evasion and two of quantity evasion of customs duties.[20] For each of those measures, we combine our two datasets. We aggregate our administrative sample at 6-digit codes to match the export data from the UN COMTRADE data in order to compute four evasion measures as in Mishra, Subramanian, and Topalova (2008). We define value evasion as the trade gap:

$$Ev_{cpt} = \log(X_{cpt}) - \log(M_{cpt}) \quad (6)$$

where X_{cpt} are the exports to Brazil reported by partner country c, of product p at year t and M_{cpt} are the import values as declared in Brazil upon goods arrival. Notice that for this measure of evasion, the sample is restricted to those transactions for which there are matched non-zero exports and imports, that is, for every export transaction with a value greater than zero there is a corresponding import. For the quantity counterpart of this measure, we just replace quantity with trade values, being careful to measure trade in the same units.

The trade gap can capture import concealment (i.e., tax base evasion) as long as importers and exporters do not diverge in their incentives to report the transaction's true value. Specifically, exporters should not have the incentive to overstate their sales and, at the

[19] The expert definition is *ad hoc*. We create two other definitions to assess robustness, high and low pervasiveness (25/35 and 10/15, respectively), and our results are not materially affected by those definitions.

[20] We drop approximately 1.94% and 2.99% of value and quantity observations respectively, as a result of non-matched exports. This trade gap measure introduced in Fisman and Wei (2004) has since been the standard proxy variable for evasion in border duties literature, (Mishra, Subramanian, and Topalova 2008; B. S. Javorcik and Narciso 2008; Bouet and Roy 2012; B. S. Javorcik and Narciso 2017).

same time, the importer is also evading. Indeed, it is unreasonable that exporters have an incentive to over-report export proceeds since they would incur corporate income taxes, and rebates on an eventual VAT collected during the production chain are limited to what has previously been paid.

Another source of concern would be the related-party transactions and the practice of transfer pricing. We believe intrafirm trade would not interfere with our trade gap measure because such practice entails export and import values to move *in tandem*, so the multinational enterprise can benefit from profit-shifting. Therefore, there would be no systematic relationship between the gap between intrafirm trade and duties rates. Notice that, in the same way, transfer pricing does not seem to affect our identification strategy, we cannot identify its presence. Thus, evasion through transfer pricing is not being evaluated in this paper. Our focus is on customs duties evasion through under-reporting of quantities, values, or misclassification on the part of importers.

The second evasion measure makes the assumption that, for any reported export that does not have a matching import, the value (or quantity) imported is assumed to be zero, that is, the product has been completely smuggled upon arrival. This extreme assumption needs a slight modification in the equation (6):

$$Evx_{cpt} = \log(1 + X_{cpt}) - \log(1 + M_{cpt}) \tag{7}$$

This measure captures the extreme value (quantity) evasion. Thus, this sample includes those observations for which exports are recorded with a value greater than zero but no counterpart import transaction is recorded. This sample subsumes the one used in Equation (6).[21]

3.2 Summary statistics

We present summary statistics of our measures of trade gap for both samples, Trade and Administrative, in Table 1. It is reasonable to believe that importers have the incentive to conceal their trade values or quantities. Therefore, we expect trade gaps to be positive or close to zero. Table 1 provides those summary statistics and they mostly confirm our hypothesis. Quantity and extreme quantity evasion have a small negative value in the trade sample. The administrative sample also presents a small negative gap for extreme value and a more pronounced negative gap for extreme quantity. All other evasion measures are positive as expected.

[21] Note that whenever the Administrative sample is used, the values (or quantities) imported, M_{cpt}, comes from this dataset, while exports figures always come from the Trade sample.

Table 1: Descriptive statistics

	Trade data			Administrative data		
	Mean	SD	Obs.	Mean	SD	Obs.
Value Evasion						
log(expo)	4.1330	2.7503	60930	4.0263	2.8113	63465
log(impo)	4.0285	2.9064	60930	3.9153	3.0517	63465
Gap	0.1045	2.0895	60930	0.1109	2.1876	63465
Extreme Value Evasion						
log(expo)	4.1576	2.4841	64722	4.1576	2.4841	64722
log(impo)	4.0810	2.5806	64722	4.1888	2.5435	64722
Gap	0.0766	1.8187	64722	-0.0312	1.7725	64722
Quantity Evasion						
log(expo)	8.1934	3.7511	60248	8.0883	3.7984	57464
log(impo)	8.1994	3.8207	60248	8.0510	4.0104	57464
Gap	-0.0061	2.5577	60248	0.0374	2.7322	57464
Extreme Quantity Evasion						
log(expo)	7.8745	3.9167	64461	7.8792	3.9317	59234
log(impo)	7.8833	4.0177	64461	8.0681	3.9030	59234
Gap	-0.0088	2.9931	64461	-0.1888	2.9018	59234

Notes: Import figures in administrative data are based on the transaction-level, 8-digit administrative sample aggregated by country, year and product code at 6-digit. Trade data uses both import and export figures from COMTRADE.

We show the products with tariff rates above the year's median and those below this value in Table 2. The trade gap is higher for products whose tariffs are above the median and we observe a negative gap for below-the-median products. This fact may be the first indication of misclassification toward products with lower tariff rates. Either the values or quantities missing from high-taxed goods resurface at low-taxed codes.

Table 2: Trade gap by tariff rate

	Below Median			Above Median		
	Mean	SD	Obs.	Mean	SD	Obs.
Value Evasion						
log(expo)	4.3729	2.8283	33792	3.6317	2.7392	29670
log(impo)	4.4337	3.0508	33792	3.3251	2.9434	29670

Table 2: Trade gap by tariff rate *(continued)*

	Below Median			Above Median		
	Mean	SD	Obs.	Mean	SD	Obs.
Gap	-0.0608	2.0643	33792	0.3066	2.3045	29670
Extreme Value Evasion						
log(expo)	8.5377	3.9487	31038	7.5609	3.5416	26423
log(impo)	8.6333	4.1641	31038	7.3678	3.7063	26423
Gap	-0.0955	2.7320	31038	0.1931	2.7233	26423
Quantity Evasion						
log(expo)	4.5013	2.5292	34187	3.7728	2.3743	30532
log(impo)	4.6532	2.6105	34187	3.6689	2.3602	30532
Gap	-0.1518	1.6920	34187	0.1039	1.8492	30532
Extreme Quantity Evasion						
log(expo)	8.3991	4.0310	31669	7.2823	3.7258	27562
log(impo)	8.6435	4.0513	31669	7.4076	3.6137	27562
Gap	-0.2444	2.7779	31669	-0.1253	3.0364	27562

Notes: Information about tariffs comes from the administrative sample and are simple averages. The aggregated at 6-digit HS codes administrative sample is used for import figures, while exports come from the trade sample.

Next, we present descriptive statistics for the administrative sample. The preshock period ranges from 2003-01-01 to 2004-04-30 for a total of sixteen months and the postshock covers the dates from 2004-05-01 to 2004-12-31, eight months, for all tables in this subsection. We start in Table 3 by showing the number of observations, unique importers, partner countries, products at 8-digit code, and entry points (e.g. ports, airports, etc.) in each period for all importers in the sample.

Table 3: Administrative sample statistics

Period	Obs.	Importers	Partners	Products	Ports	Avg. FOB	Med. FOB	FOB month
Pre (16 m.)	5,867,562	25,121	211	8,682	103	13.2801	0.749	4,870,097
Post (8 m.)	3,569,589	20,838	201	8,396	104	14.4423	0.799	6,444,137

Notes: The preshock period ranges from 2003-01-01 to 2004-04-30 and postshock covers the dates from 2004-05-01 to 2004-12-31. Values expressed in thousand dollars.

In the preshock period, we have more than 5.8 million transactions available from about

25 thousand different importers and 211 partner countries. During that period, Brazil imported 8,682 different products,[22] arriving at the border in 103 customs facilities. In the postshock period, we have more than 3.5 million transactions from over 20 thousand importers.

Table 3 shows, per period, the values in thousands of US dollars involved in those transactions for all firms. The typical import has a FOB value of $13.28 and $14.44 for the pre and postshock periods respectively. The distribution is highly skewed, as shown by a median well below the average values. The preshock period import amount is close to 78 billion US dollars, while the postshock value is close to 52 billion dollars.

Administrative treatments are shown in Table 4 for each period. The proportion of imports subject to a given treatment and its standard deviation are tabulated. We have approximately 0.03% of imports that are subject to some kind of anti-dumping measure. At the same time, about 25% of all imports in Brazil during the period under analysis had a non-automatic import license process linked to them. The import license is the most disseminated nontariff barrier an importer may face and it is the *de facto* NTB we study in this paper. Other administrative treatments include the drawback regime[23], Ex-II, Ex-IPI, exemption of PIS and COFINS on imports, and tariff agreement (e.g. MERCOSUR).[24]

Table 4: Summary statistics

	2003			2004		
Variable	Mean	SD	Obs.	Mean	SD	Obs.
Value Gap	0.1019	2.0715	29,619	0.1069	2.1065	31,311
Extreme Value Gap	0.0824	1.7975	31,399	0.0712	1.8384	33,323
Quantity Gap	0.0209	2.5547	29,278	-0.0315	2.5602	30,970
Extreme Quantity Gap	0.0206	2.9745	31,266	-0.0365	3.0103	33,195
Tariff Exposure	0.1671	0.1162	55,476	0.2278	0.1230	57,809
Import License	0.2522	0.3790	55,476	0.2493	0.3785	57,814
Anti-dumping	0.0003	0.0142	55,476	0.0002	0.0127	57,814
Audits	0.2422	0.2883	55,476	0.2272	0.2830	57,814
Drawback	0.0462	0.1638	55,476	0.0497	0.1671	57,814
Tariff Agreement	0.0512	0.2077	55,476	0.0503	0.2058	57,814
Ex-II	0.0046	0.0523	55,476	0.0057	0.0578	57,814
Ex-IPI	0.0038	0.0487	55,476	0.0031	0.0428	57,814
PIS/COFINS exemption	0.0167	0.1274	55,473	0.0167	0.1276	57,808

[22]The harmonized system does not discriminate perfectly each product, thus, different but very similar products may be classified at the same NCM code.

[23]The drawback is a special customs regime that allows for the suspension or elimination of taxes levied on the acquisition of inputs used in the industrialization of exported products.

[24]Ex-II and Ex-IPI refer to exception tax rate cases for these taxes (0.03% and 0.07% respectively preshock.). PIS and COFINS on imports exemption status is based on the authors' legislation interpretation to set some specific HS codes as exempt. For the preshock period, the exemption status is counterfactual, imports that would be exempt had PIS and COFINS been implemented at import registration.

Table 4: Summary statistics *(continued)*

Variable	Mean	SD	Obs.	Mean	SD	Obs.

Notes: Trade gaps are given in log-dollars for value or log-quantity. Tariff exposure is the total tariff exposure, comprising the II, IPI, PIS and COFINS taxes. All other variables from Import License to PIS/COFINS exemption represent the proportion of imports subject to that administrative treatment.

Finally, Table 5 shows all federal duties charged upon a typical import. The rates are expressed in decimals and we show a simple average, FOB value-weighted average, and the median value for both periods. Only the tax on imports (II) and the tax on industrialized products (IPI)[25] are present in the preshock period and, on average, those rates vary little across periods. The bulk of the customs duties burden increase relied on the inception of PIS and COFINS on imports with nominal rates of 1.65% and 7.6%, respectively.

Table 5: Summary statistics on border duties

	Preshock (16 months)				Postshock (8 months)			
Duty	Avg.	W. Avg.	Median	Std. Dev.	Avg.	W. Avg.	Median	Std. Dev.
II	0.0901	0.0408	0.105	0.0786	0.0841	0.0386	0.0960	0.0876
IPI	0.0431	0.0210	0.000	0.0599	0.0421	0.0198	0.0000	0.0588
PIS	0.0000	0.0000	0.000	0.0000	0.0160	0.0147	0.0165	0.0028
Cofins	0.0000	0.0000	0.000	0.0000	0.0737	0.0675	0.0760	0.0129

Notes: Rates in decimals. Weighted average is based on FOB values.

3.3 NTB summary statistics

Table 6 provides summary statistics comparing the imports subject to licensing during 2003 and 2004 and those exempt from such a procedure.

Table 6: Summary statistics separated by NTB status.

	Preshock (16 months)			Postshock (8 months)		
Variable	Mean	SD	Obs.	Mean	SD	Obs.
No NTB						
Tariff Exposure	0.1802	0.1234	4,231,864	0.2773	0.1452	2,568,464
Anti-dumping	0.0000	0.0059	4,231,864	0.0000	0.0054	2,568,469

[25] II stands for *imposto de importação* and IPI for *imposto sobre produtos industrializados*. II is truly a tariff on imports, while IPI is a VAT-like tax with incidence only on industrialized products (e.g. manufacturers) and is charged along the production chain.

Table 6: Summary statistics separated by NTB status. *(continued)*

Variable	Mean	SD	Obs.	Mean	SD	Obs.
Audits	0.2046	0.4034	4,231,864	0.1675	0.3734	2,568,469
Drawback	0.0000	0.0000	4,231,864	0.0000	0.0000	2,568,469
Tariff Agreement	0.0277	0.1641	4,231,864	0.0278	0.1645	2,568,469
Ex-II	0.0012	0.0342	4,231,864	0.0031	0.0558	2,568,469
Ex-IPI	0.0082	0.0902	4,231,864	0.0035	0.0588	2,568,469
PIS/COFINS exemption	0.0189	0.1362	4,231,857	0.0198	0.1392	2,568,464
NTB						
Tariff Exposure	0.0355	0.0942	1,635,698	0.1240	0.1043	1,001,117
Anti-dumping	0.0003	0.0185	1,635,698	0.0002	0.0143	1,001,120
Audits	0.1988	0.3991	1,635,698	0.1691	0.3748	1,001,120
Drawback	0.2689	0.4434	1,635,698	0.2915	0.4545	1,001,120
Tariff Agreement	0.0623	0.2418	1,635,698	0.0588	0.2352	1,001,120
Ex-II	0.0079	0.0886	1,635,698	0.0057	0.0751	1,001,120
Ex-IPI	0.0064	0.0796	1,635,698	0.0044	0.0663	1,001,120
PIS/COFINS exemption	0.0573	0.2323	1,635,680	0.0552	0.2284	1,001,117

Notes: Tariff exposure is the total tariff exposure, comprising the II, IPI, PIS and COFINS taxes. All other variables from Import License to PIS/COFINS exemption represent the proportion of imports subject to that administrative treatment. The preshock period ranges from 2003-01-01 to 2004-04-30 and postshock covers the dates from 2004-05-01 to 2004-12-31.

Our Table 6 exhibits the following findings: (i) products subjected to nontariff barriers demonstrate a tendency towards lower tariff rates, consistent with the observations made in Kee and Nicita 2016; and (ii) audits have witnessed a comparable decline rate between the NTB and No NTB groups during the preshock to postshock periods. Despite the fact that the drawback regime does not qualify as a NTB, our administrative sample categorizes all drawback imports as NTBs. This results in an artificial correlation between NTBs and drawbacks. Also, the proportion of imports subject to a tariff agreement is slightly higher in the NTB group than No NTB. This is mainly due to agricultural products, that are usually subject to NTB, where Brazil is a large importer among its MERCOSUR partners[26].

[26] As an example, in 2003 Brazil accounted for 85.9% of Argentina's wheat exports according to the The Observatory of Economic complexity (https://oec.world/en).

Table 7: Change in mean tax rates by NTB status.

	NTB			No NTB		
Variable	Preshock	Postshock	Norm. Difference	Preshock	Postshock	Norm. Difference
II	0.0235	0.0223	0.0205	0.1159	0.1081	0.0996
IPI	0.0104	0.0098	0.0166	0.0557	0.0548	0.0159
PIS	0.0000	0.0156	5.8494	0.0000	0.0162	9.9581
COFINS	0.0000	0.0718	5.8494	0.0000	0.0745	9.9581
Tariff Exposure	0.0355	0.1240	0.8901	0.1802	0.2773	0.7204

Notes: Mean tax rates for each tax and total tariff exposure, comprising the II, IPI, PIS and COFINS taxes. Absolute values for normalized differences between pre and postshock is shown in column Norm. Difference.

Table 7 goes deeper in showing how each customs duty that comprises the total tariff exposure behave in the NTB group and otherwise for both pre and postshock periods. The differences between the average tax rates are evaluated using the normalized difference (Guido W. Imbens and Wooldridge 2009). No group shown relevant differences for II and IPI, while PIS and COFINS where inexistent in the preshock period. Across groups the greater differences are found in II and IPI, where NTB products have lower tariff rates. PIS and COFINS, because of their broad base, show only minor differences across groups in the postshock period.

4 Empirical Strategy

4.1 The exogeneity of PIS/COFINS shock

The constitutional amendment 42/2003 and the Federal act 10833/2003 transformed PIS and COFINS into contributions (taxes that do not share revenues with the Federation's States) levied on the value added for most economic sectors. The rate for noncumulative sectors was set at 1.65% and 7.6%, respectively (previous rates were 0.65% and 3% on gross revenue). More important to our investigation, this amendment levied those taxes (PIS and COFINS) on imports, which were previously exempt. Although those are VAT-like taxes, which can be compensated further down the production chain, they have a significant impact on an importer's cash flow. Custom duties in Brazil must be fully paid, at their nominal rates (e.g., 1.65% and 7.6% for PIS and COFINS, respectively) before goods can leave the customs area. This means the importers, upon the arrival of goods in Brazilian territory, must remit the II, IPI, PIS, COFINS, and the State VAT (ICMS)[27] taxes in full before transporting them to their facilities. In the case of a product with

[27] The *Imposto sobre a Circulação de Mercadorias e Serviços* - ICMS - is a state tax that is also levied upon arrival at the border.

the following rates of 5%, 15%, 1.65%, 7.6% and 18%, respectively, this would result in a charge of over 55% on the import value, according to the rules of tax incidence.

Both of our datasets, *administrative and trade*, cover all Brazilian imports during 2003 and 2004 (and 2001-2002 for the placebo exercise) as the PIS and COFINS on imports were established on May first, 2004. Our focus relies on the short-term consequences because, although the inception of the taxes was exogenous, not later than 2005 onward, the Federal Act 10865/2004, which established the taxes, had undergone through several subsequent changes,[28] reducing our conviction on the exogeneity of this shock for periods beyond 2004.

Approximately five months after the Trade Reform law has passed (Laws 42/2003 and 10833/2003), the Federal Act 10865/2004 enabled PIS and COFINS on imports to be charged on May 2004. Although the PIS and COFINS rates levied upon imports were, in general, the same for all products, the differential exposition to this shock comes from previous rates of II and IPI. PIS and COFINS are "cascading" taxes, since their base contains other import duties, namely the import tariff and the tax on industrialized products, in addition to the state tax on merchandise and services. The incidence chain of border duties is as follows: import tax II is applied to the CIF value; then, the tax on industrialized products IPI is levied upon the sum of the CIF value and the amount paid for II; finally, both PIS and COFINS are levied upon the whole sum of the CIF value plus II plus IPI. Therefore, the composite tariff rate after the introduction of PIS and COFINS is given by $T_p = (1 + II_p) \times (1 + IPI_p) \times (1 + PIS + COFINS) - 1$, where we make it clear that II and IPI rates are dependent upon the product p being imported. If we consider that in 2003, PIS and COFINS were nonexistent, we arrive at the expression $\Delta T_p = T_{p,2004} - T_{p,2003} = (Pis + COFINS) \times (1 + II_p) \times (1 + IPI_p)$ for the tariff variation induced by the introduction of the new border duties.

Our identification strategy, therefore, relies on the exogenous variation of total import tariff exposure, introduced by PIS and COFINS. In addition to temporal variation, the shock produces variations across different products, something that we will utilize through the use of fixed effects in our empirical strategy. Our evasion measures, explained in the next section, will be related to this exogenous variation in tariffs and we will be able to estimate a semi-elasticity measure of customs tax evasion concerning the tariff rate.

One concern for our identification strategy is that the rates for II and IPI should not compensate for PIS and COFINS in the postshock period. To assess this possibility, we compute average rates of II and IPI by product, and attribute each product to an ordered quartile. We calculate the average exposure in the postshock period with the same products in those quartiles. The results are presented in Table 8.

[28] Federal Act 10865/2004 consolidated link can be accessed at http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2004/Lei/L10.865.htm.

Table 8: Preshock exposures and postshock changes

	Preshock (16 months)		Postshock (8 months)	
Quartile	II	IPI	II	IPI
All Imports				
1	0.0220	0.0029	0.0240	0.0028
2	0.0803	0.0160	0.0732	0.0150
3	0.1260	0.0285	0.1169	0.0259
4	0.1572	0.0917	0.1452	0.0904
NTB Imports				
1	0.0000	0.0000	0.0027	0.0004
2	0.0002	0.0001	0.0032	0.0005
3	0.0191	0.0044	0.0181	0.0032
4	0.1134	0.0400	0.0993	0.0346

Notes: Values in decimals for tariff rates. Quartiles are based on baseline's total average exposure by aggregate HS 6-digit product code.

There is no change in average rates for all importers or trading companies. The first quartile, the least exposed, saw a small increase in II rates, while other quartiles experienced small reductions. These intraquartile movements do not compensate for the 9.25 percentage point increase in tariff exposure; moreover, we can see that rates in the postshock period are still correctly ordered.

Figure 2 shows the scatter plots for the Most favored nation (MFN) and the effectively applied tariff rates (EAR) tariff burden with their figures plotted on a reference line of slope one. We can observe that most of the products lie above this reference line, representing the burden increase due to the PIS and COFINS shock. Although there is a burden reduction for some products, for example, in the EAR top panel, 3.00% of products had tariff exposure reduction in the postshock period. The correlation of tariff exposure is positive and close to one, reinforcing the argument of an exogenous shock. If compensation in II and IPI rates had fully occurred, we would observe a flattening curve where the 2004 exposure has little correlation to the previous year's tariff rate. However, a perfect and positive correlation indicates that there was no compensation; the most taxed goods in 2003 are still at the top of the rankings after the shock, with a higher total tariff exposure.

The results presented in Figure 2 indicate that the preshock proportion of NTB imports, represented by lighter colors (i.e., green-yellow), does not exhibit any discernible pattern. This suggests that companies importing NTB subject products have not altered their

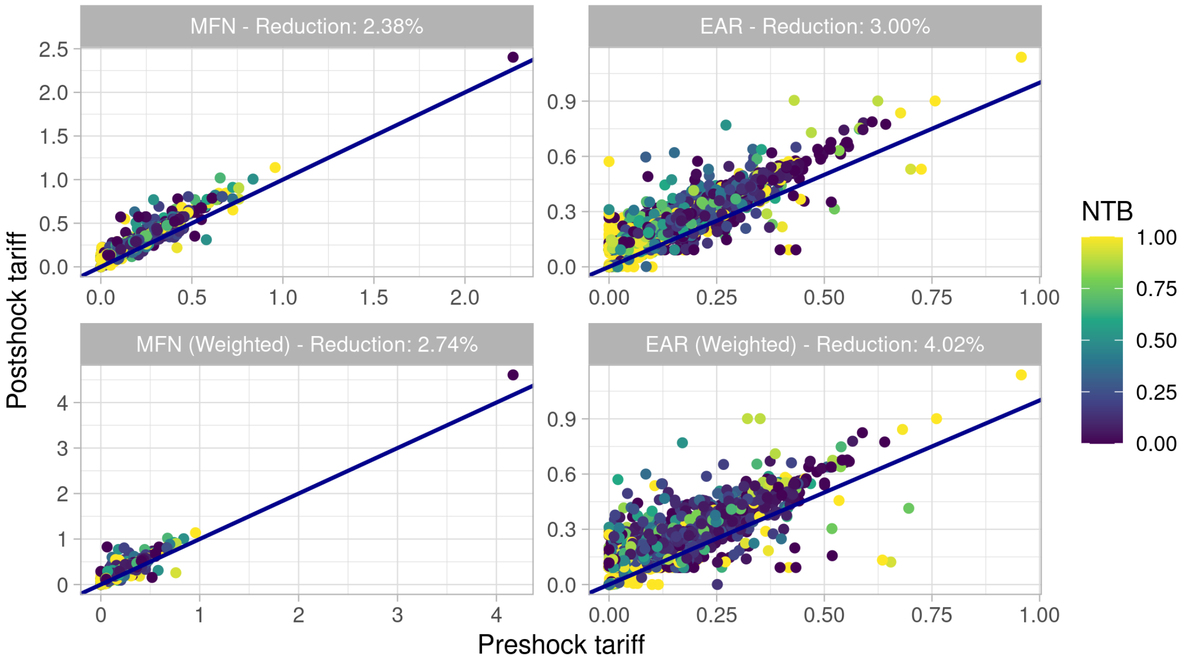

Figure 2: Correlation of pre and postshock tariff burden. Each dot represents an HS 6-digit product. Most favored nation (MFN) and effectively applied tariff rates (EAR). Products above the reference line have seen a postshock total tariff exposure increase. Previously high-taxed products tend to keep their status in the postshock period. The color scheme denotes the preshock proportion of NTB on the number of imports. Top row presents tariff exposure computed from simple averages and weighted averages are used in bottom row. Each panel discriminates the share of observations that reduced exposure pre versus postshock.

product mix in the postshock period.

4.2 Econometric specification

As evasion can take place either by mispricing the unit value of imported goods, smuggling quantities or through misclassification of products in the NCM, we implement a machine learning model using the causal forests algorithm (Athey, Tibshirani, and Wager 2019; Wager and Athey 2018) to investigate potential causes for border duties evasion. We find that a similar product classification is the most relevant feature to explain evasion heterogeneity in three out of four evasion measures.[29]

To introduce product misclassification in our econometric strategy, we follow Fisman and Wei (2004). For each observation at the 6-digit level, country and period, we define the similar group as all products sharing the same 4-digit code and compute the minimum tariff exposure of such a group. Similar products are less taxed than the correct classification, and the importer can take the risk of misclassifying to evade customs duties. That is, by keeping the tariff on a given product constant, lower tariffs on similar products can yield higher levels of evasion. We consider the following econometric specification to address our theoretical framework:

$$y_{cpt} = \beta_1 T_{cpt} + \beta_2 T_{cpt}^{sim} + \lambda' X_{cpt} + \delta_{cp} + \delta_{ct} + \delta_{pt} + \varepsilon_{cpt}, \tag{8}$$

where y refers to the trade gap (our evasion measures for values and quantities), $t \in \{2003, 2004\}$ reflects the PIS and COFINS shock when only 2004 has been affected.[30] The total tariff exposure goods are subject to is designated by T_{cpt} and it is dependent on the originating country since Brazil does have a major trade agreement, MERCOSUR. We use as tariff exposure an average of actually applied tariff rates on each transaction from the Administrative sample. T_{cpt}^{sim} is the minimum tariff exposure of similar products. We also introduce fixed effects for the pairs country-product, country-year, and product-year, δ_{cp}, δ_{ct}, δ_{pt} respectively, to account for unobserved characteristics related to these pairs. For example, the product-time fixed effects account for changes in the productivity of a specific product that is shared among countries, as well as changes in trade policy targeted at the product level. The country-product fixed effects absorb a country's comparative advantage of exporting product p while the pair country-year may capture the overall change in the productivity of an exporter country. We also include in the main specification control variables and administrative treatments which may confound how the tariff exposure affects evasion, X_{cpt}. Administrative treatments comprise the follow-

[29] The machine learning analysis is presented in Appendix D and focuses on explaining evasion heterogeneity through observable import characteristics.

[30] Another interpretation of our empirical strategy is similar to a shift-share approach with only one industry. In that case, we must assume that the units that received stronger shocks would have the same evolution of potential outcomes as those that received weaker tax shocks.

ing: exemption status from PIS and COFINS, tariff agreements, anti-dumping duties, and drawbacks. Other control variables include average exchange rate, proportion of trading companies, proportion of audits, brazilian States where the import was cleared, average time to clearance, insurance and freight costs, number of imported products, number of unique importers and importer's years in activity.[31]

The variation in tariffs induced by the shock is $\Delta T_{cp} = T_{cp,2004} - T_{cp,2003} = (Pis + COFINS) \times (1 + II_{cp}) \times (1 + IPI_{cp})$ due to the incidence nature of PIS and COFINS taxes.[32] Therefore, the variation in customs duties burden faced by importers due to the introduction of those two taxes on imports of goods and services is proportional to the preshock total tariff rate exposure, $(1 + II_{cp}) \times (1 + IPI_{cp})$. Since this exposure is set on a product basis, the shock created variation across products even though PIS and COFINS rates were generally the same for all products.[33] Considering the wide tax base of PIS and COFINS across products and irrespective of origin country, it is not likely that endogeneity issues arise. Even though, for each set of results, we provide several robustness assessments and falsification tests in the Appendix B.

NTB Heterogeneity

Nontariff barriers have become a prominent trade policy instrument in the last few decades (WTO 2012). Grübler and Reiter (2021) reports the evolution of NTM notifications to the WTO, and the results show a jump from approximately one thousand notifications in 1996 to approximately four thousand notifications in 2019 (Grübler and Reiter 2021, Figure 8, p.p. 149). We incorporate this variable along with its interaction with our tariff exposure into Equation (8) to obtain our new econometric specification:

$$y_{cpt} = \beta_1 T_{cpt} + \beta_2 T_{cpt} \cdot NTB_{cpt} + \beta_3 T^{sim}_{cpt} + \beta_4 NTB_{cpt} + \lambda' X_{cpt} + \delta_c + \delta_p + \delta_t + \varepsilon_{cpt} \quad (9)$$

where y_{cpt}, assumes one of the four evasion measures as in the equation (8) and, NTB_{cpt} is the proportion of product p imports, coming from partner country c at year t subject to import licensing, our representative NTB. In addition, we include fixed effects without interaction. Since the NTB designation is associated with a *per* product basis, the product-year fixed effect would absorb most of the effect we are interested in.

[31] Appendix B provides additional robustness checks for this linear specification. Figure 3 shows that the proportion of NTBs did not change from 2003 to 2004 by tariff quintile. Table B.5 starts from the main specification and, incrementally includes the average exchange rate and the proportion of import value by State and re-estimates the model. The results do not qualitatively change.

[32] We are simplifying this expression by assuming that both II and IPI did not change over time. Introducing variation in these taxes would only bring detailed caveats that do not affect our results.

[33] With few exceptions, such as exempt products and petroleum-derived products, which are subject to specific rates.

5 Empirical results

We provide the first set of estimation results for the ECDE in Table 9. We estimate equation (8) considering two types of tariffs and the four evasion measures, in a total of eight models estimated. In the first set of models, the tariff is computed using a simple average of effectively applied tariff rates, information available only in the administrative dataset.[34] The next set of models uses the most favored nation tariff, which is the tariff typically found in studies using data from UNCTAD/TRAINS.

[34]We also run regressions with FOB value-weighted tariffs. Both aggregation methods have advantages and disadvantages. Since our dependent variable relies on FOB values, the weighted average tariff may raise an endogeneity concern, while a simple average gives equal weight to products with different import value shares but the same number of transactions. Weighted average tariff results are qualitatively the same and are available upon request.

Table 9: Regression results comparing effectively applied against MFN tariff rates

	Administrative Sample - Effectively Applied Tariff				Trade Sample - MFN Tariff			
Dep. Var.:	Value	Ext. Value	Quantity	Ext. Quantity	Value	Ext. Value	Quantity	Ext. Quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tariff	1.9840***	1.5444***	0.9657**	0.8679**	1.1282**	1.0570***	-0.1171	-0.1856
	(0.3082)	(0.2382)	(0.4401)	(0.4059)	(0.5614)	(0.3687)	(0.6897)	(0.5927)
Similar	-0.4960**	-0.3894**	-0.2878	-0.2546	-0.0314	-0.0440	-0.0067	0.0075
	(0.2131)	(0.1494)	(0.3208)	(0.2875)	(0.2264)	(0.1431)	(0.3168)	(0.2803)
Num.Obs.	63226	64481	57267	59023	63226	64481	57267	59023
R2 Adj.	0.538	0.570	0.546	0.619	0.537	0.569	0.545	0.619
FE: cp	✓	✓	✓	✓	✓	✓	✓	✓
FE: ct	✓	✓	✓	✓	✓	✓	✓	✓
FE: pt	✓	✓	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product or the most favored nation accordingly. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

The results are quite distinct between data sources. Higher and statistically significant elasticities are estimated when the administrative data are used. For example, the value ECDE estimated using the administrative data is 1.98 (i.e. for a one percentage point increase in total tariff exposure, evasion raises by 1.98%). When regressions are carried out on the MFN tariff rate, the resulting elasticities are lower, with larger standard errors and evasion for quantities are nonsignificant at the 10% level. The lack of significant results with the MFN tariff is expected since importers care about *effectively* applied tariff rates on their imported goods. The results in Table 9 reinforce our claim that administrative data should be used to assess evasion whenever available, especially for developing economies where customs duties are not restricted to the MFN tariff only.[35]

Our finding that quantity evasion responses to tariffs are statistically significant contrasts with previous literature (B. S. Javorcik and Narciso (2008) and Fisman and Wei (2004) when not controlling for similar products).[36] Although it is easier for a customs official to identify smuggling quantities by comparing the reported versus the true weight of shipment containers, a possible explanation for finding quantity evasion responses in Brazil might come from the combination of tight regulation of prices through the Customs Valuation Agreement (CVA) and high tariff rates (II, IPI, PIS, and COFINS). B. S. Javorcik and Narciso (2017) show that this combination could trigger quantity evasion (i.e., smuggling) by importers.

Table 9 shows that the tariff rate on similar products is relevant to determining the evasion in value terms.[37] When the minimum tariff rate of the same 4-digit products is considered, the $Tariff$ coefficients are slightly larger and still significant at the 1% level, while the coefficients on similar products have the proper negative sign. These results suggest that importers misreport their high-taxed goods as low-taxed varieties, and they choose the classification of the imported goods with the lowest tariff possible within a group of similar HS codes.

Validity check

In order to check the parallel trends assumption of our differences-in-differences identification strategy, we compute placebo-like regression estimates using trade gaps from the years 2001 and 2002. For this exercise, we only use the observations where the product-partner is present in all four years, 2001 to 2004, such that a 2003 trade gap for a given product coming from an origin country is replaced by the respective 2001 trade gap. The

[35] For illustration purposes, we provide in Table B.1 the difference of averages in effectively applied tariff rate and the most favored nation tariff rate for the top five most imported products (HS 6-digit) in 2003 and 2004. We observe clear differences in tariff levels, more than six percentage points for electronic integrated circuits with mostly zero tariffs on petroleum-related products.

[36] We conjecture that previous studies may not have found positive elasticities because of the use of MFN tariffs for developing economies. Mishra, Subramanian, and Topalova (2008) find positive and statistically significant quantity evasion in India, but with a much smaller magnitude.

[37] We also performed several robustness checks considering the average tariff rate of similar products, excluding the product at hand, instead of the minimum rate (Table B.2), the simple and weighted averages of similar products on evasion (Table B.2). The results are in Appendix B.

results are shown in Table 10. Two results deserve attention. First, the columns 1-4 show that the estimated coefficients are mostly unchanged when compared to our estimations using the full sample in Table 9. Second, the once statistically significant coefficients on *Tariff* and *Similar* explanatory variables loose significance due to the reduction in the estimated magnitude for all models, while standard errors do not materially change, reinforcing our empirical strategy.

Table 10: Placebo results including minimum tariff rate on similar products

	Original				Placebo			
Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)	Value (5)	Ext. Value (6)	Quantity (7)	Ext. Quantity (8)
Tariff	1.7152***	1.4299***	0.7151	0.6640	0.4499	0.2317	-0.2021	0.0550
	(0.4371)	(0.3273)	(0.5087)	(0.4728)	(0.4268)	(0.3554)	(0.4814)	(0.4816)
Similar	-0.6331***	-0.4813***	-0.3154	-0.3293	-0.1074	-0.0568	-0.0500	-0.0687
	(0.2207)	(0.1525)	(0.3166)	(0.3048)	(0.2493)	(0.2053)	(0.2982)	(0.2961)
Num.Obs.	46379	47351	42308	43366	45405	46334	41555	42460
R2 Adj.	0.515	0.593	0.523	0.626	0.472	0.540	0.454	0.567
FE: cp	✓	✓	✓	✓	✓	✓	✓	✓
FE: ct	✓	✓	✓	✓	✓	✓	✓	✓
FE: pt	✓	✓	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariffs on similar products and its own are calculated as simple average from transaction-level data by origin country, year and 6-digit product. All standard errors are two-way clustered at country and product. All regressions include additional control variables. We include only observations where product-country is present in all four years, 2001 through 2004. Original results are those for years 2003 and 2004, while placebo results use trade gaps taken from years 2001 and 2002.

Cross section Models. Also, we explore four years of data, from 2001 to 2004 individually, to estimate the ECDE. This specifications uses only the cross-section of product-origin country within a given year. The results follow in Table 11.

Table 11: Regression results including minimum tariff rate on similar products. 2001.

Dep. Var.:	Value	Ext. Value	Quantity	Ext. Quantity
	(1)	(2)	(3)	(4)
Tariff	2.4494***	2.1040***	1.6507***	1.9080***
	(0.3486)	(0.2602)	(0.5084)	(0.5404)
Similar	-0.3897	-0.3836**	-0.5553	-0.6110
	(0.2768)	(0.1905)	(0.3768)	(0.3975)
Num.Obs.	30603	31211	28042	28666
R2 Adj.	0.113	0.169	0.135	0.259
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariffs on similar products and its own are calculated as simple average from transaction-level data by origin country, year and 6-digit product. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Table 12: Regression results including minimum tariff rate on similar products. 2002.

Dep. Var.:	Value	Ext. Value	Quantity	Ext. Quantity
	(1)	(2)	(3)	(4)
Tariff	3.2602***	2.5110***	2.3956***	2.4014***
	(0.3942)	(0.3055)	(0.4077)	(0.3908)
Similar	-0.5566	-0.4543*	-1.2247***	-1.2078***
	(0.3650)	(0.2485)	(0.3164)	(0.3395)
Num.Obs.	30191	30784	27561	28208
R2 Adj.	0.113	0.178	0.118	0.263
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariffs on similar products and its own are calculated as simple average from transaction-level data by origin country, year and 6-digit product. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Table 13: Regression results including minimum tariff rate on similar products. 2003.

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	2.6929***	2.2537***	1.8070***	1.9097***
	(0.3335)	(0.2701)	(0.5025)	(0.4364)
Similar	-1.1487***	-0.8334***	-1.1582***	-1.0932***
	(0.2854)	(0.2231)	(0.3917)	(0.3673)
Num.Obs.	30710	31267	27795	28620
R2 Adj.	0.097	0.160	0.175	0.301
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓

* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Tariffs on similar products and its own are calculated as simple average from transaction-level data by origin country, year and 6-digit product. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Table 14: Regression results including minimum tariff rate on similar products. 2004.

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	2.1195***	1.8545***	1.1887***	1.2286***
	(0.3957)	(0.2104)	(0.3755)	(0.3497)
Similar	-0.3265	-0.5244**	-0.4299	-0.4528
	(0.4099)	(0.2073)	(0.3497)	(0.3314)
Num.Obs.	32516	33214	29472	30403
R2 Adj.	0.102	0.168	0.162	0.292
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓

* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Tariffs on similar products and its own are calculated as simple average from transaction-level data by origin country, year and 6-digit product. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Tables 11-14 show that for every year we overestimate the ECDE in relation to our panel data. This reinforces our empirical strategy that explores the exogenous introduction of PIS/COFINS on imports that affected products with different intensities using 2003-2004 data.

Next, since the literature has found a nonlinear effect of tax rates on evasion (Fisman and Wei 2004; Dowd, Landefeld, and Moore 2017; Fuest, Hugger, and Neumeier 2022), we also check whether tariffs have a nonlinear effect on evasion responses of customs duties, i.e., the marginal effect on evasion of a tariff increase would depend on the tariff rate itself.[38] We introduce a squared tariff exposure term to our regression specified in equation (8), where Table 15 shows our results.

Table 15: Inclusion of squared tarrif rates

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	1.9260***	1.7352***	0.2622	0.1906
	(0.6631)	(0.4954)	(0.8149)	(0.7495)
Tariff^2	0.1029	-0.3393	1.2262	1.1918
	(0.9701)	(0.7009)	(1.2741)	(1.1766)
Similar	-0.4934**	-0.3976***	-0.2521	-0.2219
	(0.2116)	(0.1476)	(0.3200)	(0.2895)
Num.Obs.	63226	64481	57267	59023
R2 Adj.	0.538	0.570	0.546	0.619
FE: cp	✓	✓	✓	✓
FE: ct	✓	✓	✓	✓
FE: pt	✓	✓	✓	✓

* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Tariffs on similar products and its own are calculated as simple average from transaction-level data by origin country, year and 6-digit product. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Although we do not find statistically significant coefficients associated with squared tariff exposure, it is interesting to interpret the negative coefficients for the evasion in values. They indicate larger marginal effects on low-taxed goods, a similar result in Dowd, Landefeld, and Moore (2017) and Fuest, Hugger, and Neumeier (2022). Importers of low-taxed products are more sensitive to tariff hikes. At the same time, quantity evasion appears to be more extreme the higher the tariff rate. This may be because truly high-taxed products are not easily misclassified, leaving importers with only two evasion channels, under-reporting the value and smuggling. The first was already being explored, while the latter, a riskier practice, happens mostly after some tariff rate threshold has been surpassed.

[38]Dowd, Landefeld, and Moore (2017) and Fuest, Hugger, and Neumeier (2022) are studies about profit shifting and, although it is not technically evasion, we are alluding to any form of tax planning as evasion.

5.1 The modulating effect of NTB

Before moving to our estimation results concerning NTBs, we check whether the proportion of transactions involving nontariff barriers has significantly changed from the pre to postshock period. Figure 3 shows the distribution of NTBs by tariff quintile for 2003 and 2004. The chart shows that NTB-related imports did not change their distribution across tariff quintiles.

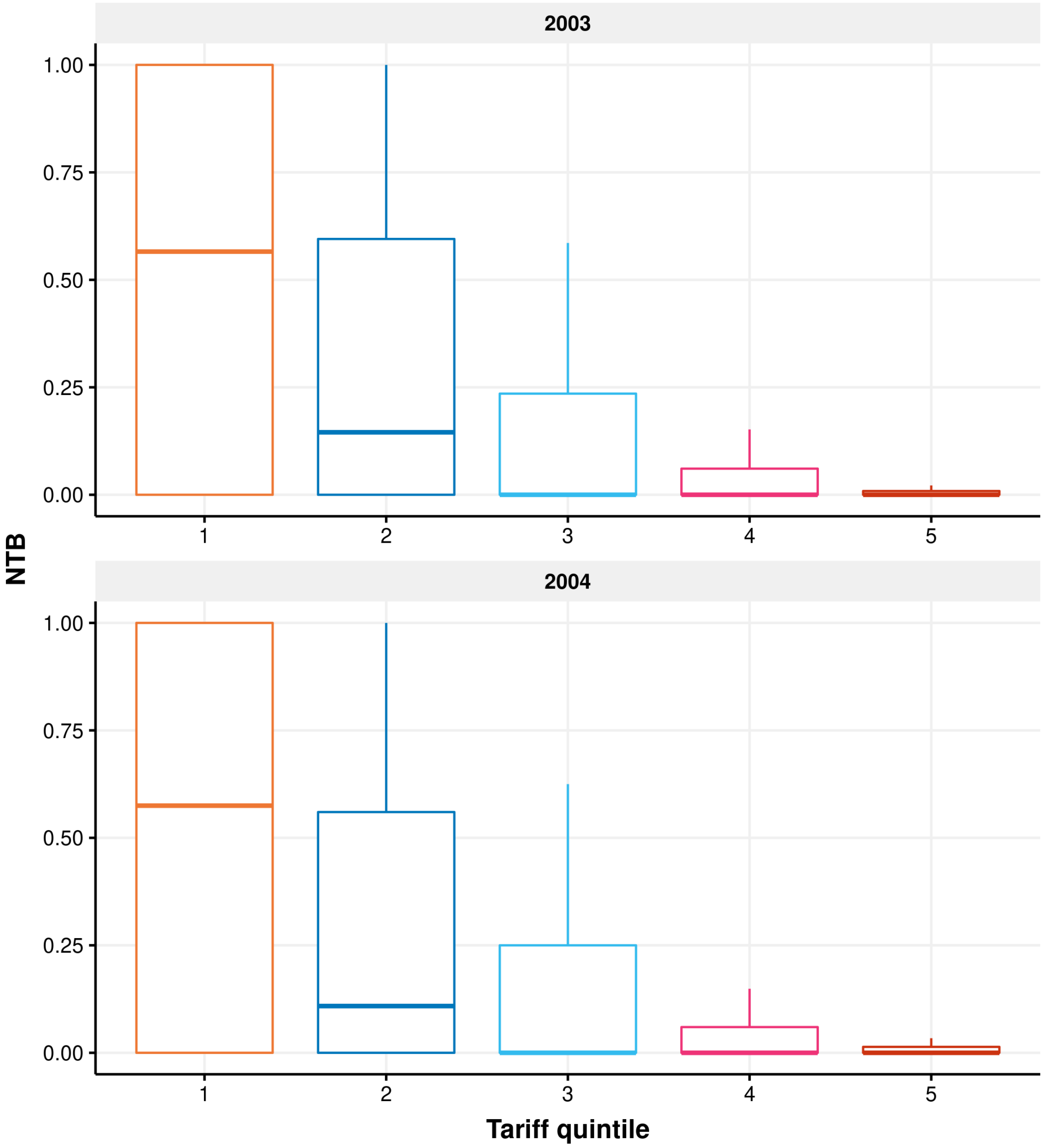

Figure 3: Boxplots representing the distribution of NTB imports by tariff quintiles.

The question we address in this section is how the imposition of one type of NTB, the import licensing restriction affects the elasticity of customs duties evasion. Table 16 presents our results. The imposition of an import license does have a modulating effect on the elasticity of evasion, even after controlling for tariff rates and misclassification on similar products, consonant with our machine learning model's suggestion in Appendix D. This relevance of NTBs for elasticity heterogeneity always figures among the top three most important features. For every evasion measure in Table 16, we find that the higher the proportion of imports subject to the NTB, the lower the elasticity of evasion.[39] For instance, consider the evasion of value model in column (1). The marginal effect of tariffs on evasion at the mean point of NTB is 2.13 (2.62 -1.86 * 0.265), slightly higher than the evasion elasticity in Table 9 (1.98). Note, however, that this estimated elasticity response goes to 0.76 (2.62-1.86*1) if all imported goods are subject to the NTB. For quantity evasion (column (3), Table 9), the corresponding estimated elasticity would be negative (-0.09 = 1.90-1.99*1).

Table 16: Effect of non-automatic import licensing on evasion

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	2.6194***	2.1622***	1.8979***	1.8751***
	(0.2849)	(0.1995)	(0.3824)	(0.3573)
Tariff X NTB	-1.8592***	-1.4200***	-1.9916***	-1.8288***
	(0.3943)	(0.2722)	(0.4874)	(0.4607)
Similar	-0.6280***	-0.6048***	-0.7158**	-0.6927**
	(0.2281)	(0.1577)	(0.2837)	(0.2723)
NTB	0.0250	-0.0180	0.2338*	0.1697
	(0.0933)	(0.0652)	(0.1382)	(0.1261)
Num.Obs.	63226	64481	57267	59023
Avg. Tariff	0.1908	0.1921	0.1894	0.1909
Avg. NTB	0.2647	0.2646	0.2696	0.2684
R2 Adj.	0.121	0.187	0.185	0.311
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓
FE: t	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product from the administrative dataset. NTB is the proportion of imports subject to non-automatic licensing for a given triplet, country, product, and year. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

We interpret that the nontariff barrier plays a substitution role with tariff reduction in

[39] As a further robustness check, we computed a permutation test on the interaction term $Tariff \times NTB$. The results are shown in Figure B.4 in Appendix B.

curbing customs duties evasion.[40] Given the level of tariffs, a product subject to the NTB has a lower elasticity of evasion. This effect is not due to an increased proportion of audits on NTB-subject imports,[41] but rather, associated with an increase in enforcement quality through detailed information access.[42] Indeed, the whole process of obtaining the import license by providing the government with detailed information may be faced by the importers as a higher level of enforcement. The license filing creates a verifiable paper trail by importers and can be later used by auditors, (Pomeranz 2015). A typical import license requirement conveys the following information: importer ID, country of origin, clearance port, exporter information, and goods information such as NCM classification, administrative treatments, taxation regime, payment method, and unit value of the imported good.[43] This type of information, accessible by tax authorities, plays a major role in tax enforcement (Kopczuk and Slemrod 2006; Gordon and Li 2009; Kleven et al. 2011; Arbex and Mattos 2015; Slemrod et al. 2017; Naritomi 2019) since authorities can cross-reference the prior information with those being presented in the import declaration and detect any discrepancies.

Furthermore, the negative sign associated with the interaction of tariff rate and proportion of import licenses is in accordance with our theoretical framework where the quality of enforcement reduces the elasticity of evasion.[44]

Finally, one last concern with the NTB's interaction specification is the possibility that the NTB is capturing a dosage effect of tariffs on evasion. Figure 3 shows that NTBs are mainly concentrated at low tariff products. We further investigate a possible tariff dosage effect by setting a dummy variable for low tariff products (i.e., zero tariff rate for both II and IPI) and implementing the same regression specification as Equation 9. However, we switch the proportion of NTBs by a low tariff indicator (Low Tariff). The results provided in Table 17 are reassuring that our NTB modulator effect on evasion does not come from its correlation with low tariffs.

[40]Kee and Nicita (2016) point out the substitution characteristic of tariff and NTMs *in relation to trade flows*, while we show the same aspect of the two instruments when evasion is investigated.

[41]See Table B.4 for a balance table on imports subject to NTB.

[42]Other sources of heterogeneity have been previously studied. Mishra, Subramanian, and Topalova 2008 and B. S. Javorcik and Narciso 2008 for differentiated products, while Anson, Cadot, and Olarreaga 2006 and Yang 2008 investigate pre-shipment inspections. Those are direct measures of enforcement. We argue that the mitigating effect of evasion from nontariff barriers is an unintended consequence since NTBs are not aimed at curbing evasive behavior.

[43]In Annex E, we provide an example of an import license document. Fields that could potentially identify the importer have been redacted.

[44]We consider that the tariff equivalent of NTB on imports is not their tariff itself. The imposition of an NTB would raise the tariff equivalent. We simulated the result of a 20% increase in tariffs for NTB subject imports and present the results in Table B.6 in Appendix B.

Table 17: Dosage effect of tariff on evasion

Dep. Var.:	Value	Ext. Value	Quantity	Ext. Quantity
	(1)	(2)	(3)	(4)
Tariff	2.9777***	2.4275***	2.2922***	2.2751***
	(0.3061)	(0.2370)	(0.3743)	(0.3549)
Tariff X Low Tariff	1.3848	0.6596	2.0266	1.7954
	(1.0447)	(0.6832)	(1.3460)	(1.2099)
Similar	-0.5302**	-0.5394***	-0.5973**	-0.5886**
	(0.2320)	(0.1589)	(0.2882)	(0.2753)
Low Tariff	0.7881***	0.5781***	0.9576***	0.8735***
	(0.1777)	(0.1348)	(0.1973)	(0.1754)
Num.Obs.	63226	64481	57267	59023
Avg. Tariff	0.1908	0.1921	0.1894	0.1909
Avg. Low Tariff	0.04518	0.04486	0.04506	0.04492
R2 Adj.	0.126	0.191	0.190	0.315
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓
FE: t	✓	✓	✓	✓

* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product from the administrative dataset. Low tariff is a binary variable indicating that both, II and IPI have zero tax rate. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Although the Tariff and Similar coefficients are still statistically significant and comparable in magnitude with the NTB specification, the interaction of Tariff X Low Tariff is not significant in any model, with positive point estimates.

Validity Check

A concern for our difference-in-differences specification is about nontariff barrier status and whether it increases the likelihood of being audited by customs authorities. To investigate this issue, we analyzed the audit rates of products subject to NTBs and those that are not. We used data from 2001 to 2005 and plotted the monthly trends of audit rates for both groups of products.

As we can see from Figure 4, there is no significant difference in audit rates between NTB and non-NTB products over time. In February 2003, there was a peak in audit rates on NTB products that was reversed thereafter. By the end of 2004, both trends converged to very similar levels of audit rate. This suggests that NTBs do not have a systematic impact on the probability of being audited by customs authorities.

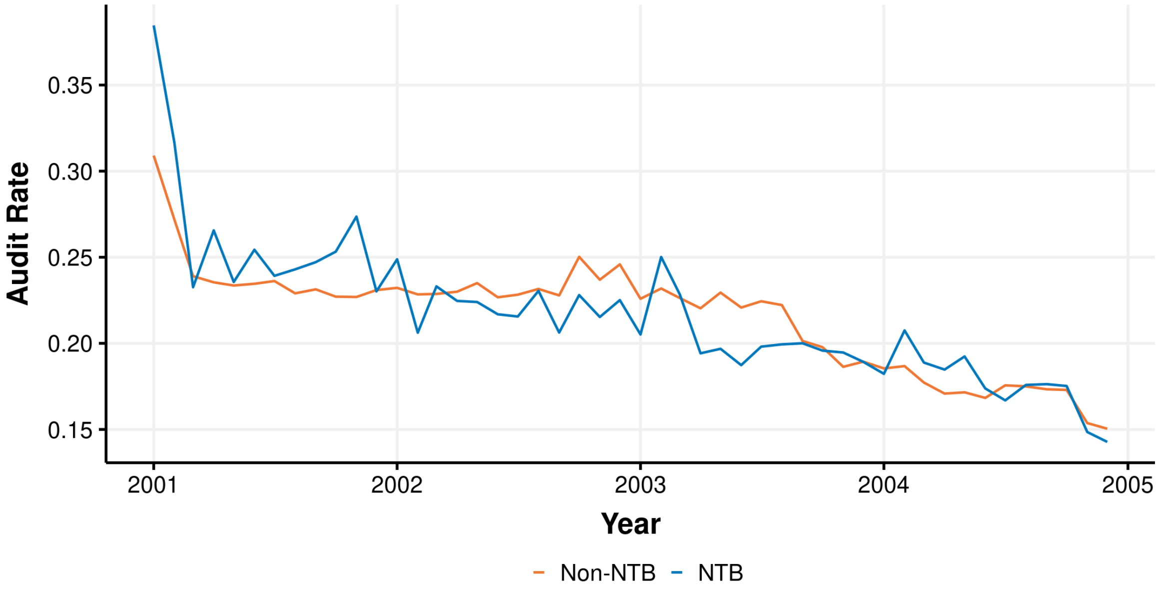

Figure 4: Monthly audit rates by NTB status.

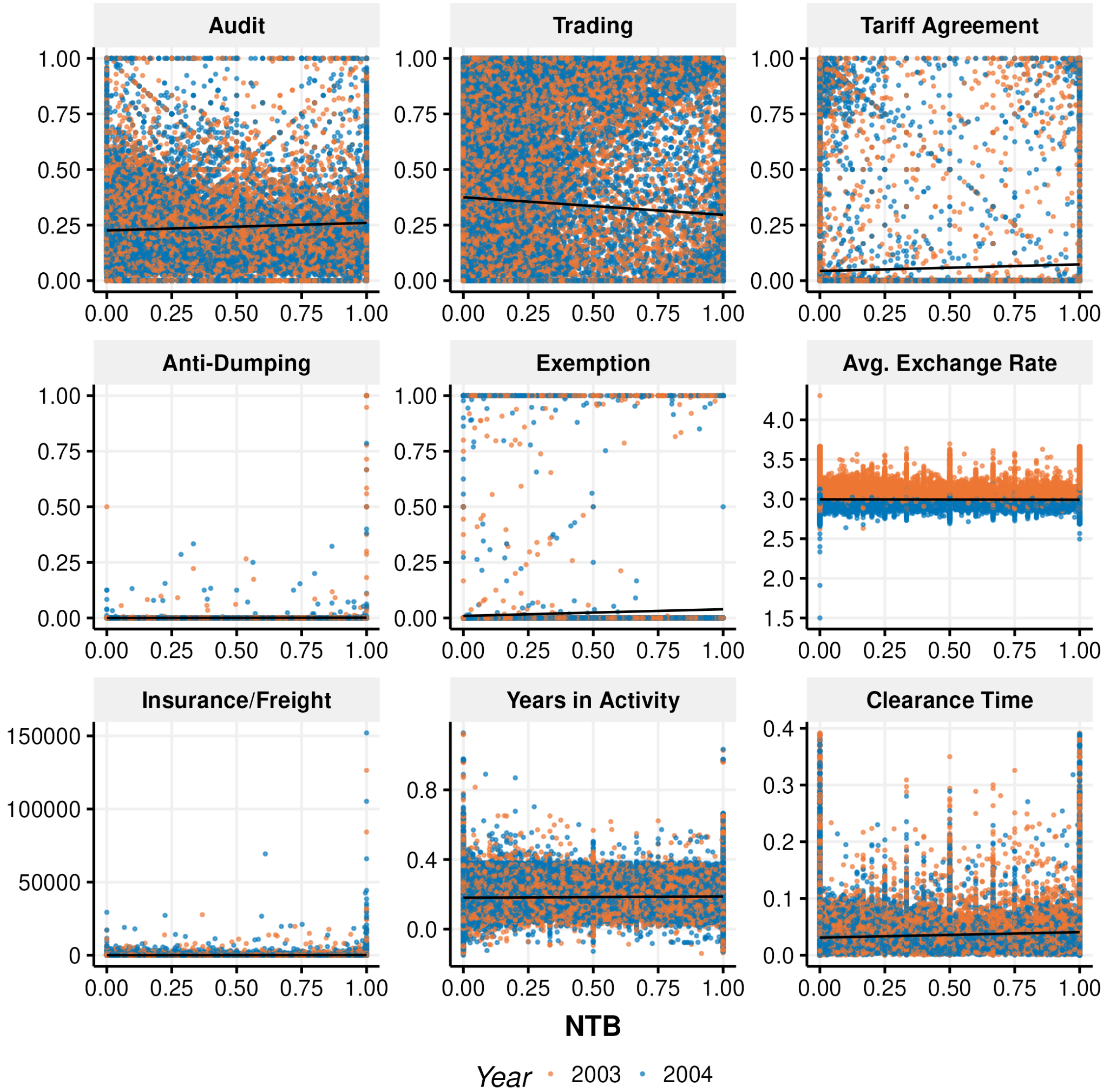

Figure 5: NTB correlation to possible evasion confounders.

Cross-section Models Last, Tables 19-20 show that, using cross-section data, the interaction between our NTB variable and the Tariffs produces a much larger estimated effect than our methodology. The overestimation magnitude can be as high as three times larger.

Table 18: Effect of non-automatic import licensing on evasion. 2001

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	3.2784***	2.6649***	3.0758***	3.2383***
	(0.4212)	(0.3114)	(0.6182)	(0.6307)
Tariff X NTB	-3.1906***	-2.3060***	-4.6268***	-4.5625***
	(0.6089)	(0.4373)	(1.1175)	(1.0939)
Similar	-0.3220	-0.3373*	-0.4591	-0.5205
	(0.2694)	(0.1840)	(0.3819)	(0.4072)
NTB	0.3697*	0.2259*	0.7955***	0.7032***
	(0.1883)	(0.1300)	(0.2696)	(0.2505)
Num.Obs.	30603	31211	28042	28666
Avg. Tariff	0.1847	0.1847	0.1847	0.1847
Avg. NTB	0.2179	0.2179	0.2179	0.2179
R2 Adj.	0.115	0.171	0.138	0.261
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product from the administrative dataset. NTB is the proportion of imports subject to non-automatic licensing for a given triplet, country, product, and year. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Table 19: Effect of non-automatic import licensing on evasion. 2002

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	4.0549***	3.1261***	3.8256***	3.7812***
	(0.4732)	(0.3508)	(0.5511)	(0.5333)
Tariff X NTB	-2.8745***	-2.3828***	-4.2186***	-4.2165***
	(0.8565)	(0.5873)	(1.1576)	(1.1115)
Similar	-0.5716	-0.4665*	-1.2490***	-1.2322***
	(0.3675)	(0.2417)	(0.3143)	(0.3364)
NTB	0.3015	0.1805	0.7763***	0.7128***
	(0.2017)	(0.1563)	(0.2626)	(0.2366)
Num.Obs.	30191	30784	27561	28208
Avg. Tariff	0.1767	0.1767	0.1767	0.1767
Avg. NTB	0.2354	0.2354	0.2354	0.2354
R2 Adj.	0.114	0.180	0.120	0.265
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓

* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product from the administrative dataset. NTB is the proportion of imports subject to non-automatic licensing for a given triplet, country, product, and year. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Table 20: Effect of non-automatic import licensing on evasion. 2003

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	3.1855***	2.5827***	2.6183***	2.5854***
	(0.3775)	(0.3190)	(0.5717)	(0.5155)
Tariff X NTB	-2.5012***	-1.8469***	-2.9681***	-2.5884***
	(0.5392)	(0.3964)	(0.6593)	(0.6299)
Similar	-1.0996***	-0.7971***	-1.1033***	-1.0465***
	(0.2849)	(0.2221)	(0.4006)	(0.3776)
NTB	-0.0808	-0.1180	0.2063	0.1433
	(0.1263)	(0.0941)	(0.1474)	(0.1356)
Num.Obs.	30710	31267	27795	28620
Avg. Tariff	0.1671	0.1671	0.1671	0.1671
Avg. NTB	0.2522	0.2522	0.2522	0.2522
R2 Adj.	0.099	0.162	0.176	0.302
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product from the administrative dataset. NTB is the proportion of imports subject to non-automatic licensing for a given triplet, country, product, and year. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Table 21: Effect of non-automatic import licensing on evasion. 2004

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	2.5420***	2.1810***	1.7035***	1.7214***
	(0.4363)	(0.2427)	(0.4488)	(0.4207)
Tariff X NTB	-2.0581***	-1.6583***	-2.1646**	-2.1983***
	(0.6617)	(0.4530)	(0.8527)	(0.7812)
Similar	-0.2699	-0.4771**	-0.3702	-0.3923
	(0.3972)	(0.1995)	(0.3404)	(0.3214)
NTB	0.1847	0.1287	0.3268	0.2847
	(0.1826)	(0.1333)	(0.2660)	(0.2361)
Num.Obs.	32516	33214	29472	30403
Avg. Tariff	NA	NA	NA	NA
Avg. NTB	0.2493	0.2493	0.2493	0.2493
R2 Adj.	0.103	0.169	0.163	0.293
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓

* p < 0.1, ** p < 0.05, *** p < 0.01

Notes: Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product from the administrative dataset. NTB is the proportion of imports subject to non-automatic licensing for a given triplet, country, product, and year. All standard errors are two-way clustered at country and product. All regressions include additional control variables.

Therefore, we focus on our panel data results to calculate the optimal tariff rate.

5.2 Optimal tariff rate

To establish the economic implication of our estimated evasion elasticity to the optimal tariff rate, consider differentiating equation (8) with relation to the tariff exposure, T_{cpt}. Rearranging in terms of trade flow elasticities, we have:

$$\eta_M = \eta_X - \beta \cdot T \tag{10}$$

where η_M and η_X are reported imports and exports elasticities with relation to tariffs, respectively. As we cannot access an estimate of η_X, and tariffs are part of the trade costs in a gravity model, the trade literature has found it to be negative (De Sousa, Mayer, and Zignago (2012) and, Head and Mayer (2014)). Consider the extreme scenario where $\eta_X = 0$ and our result of $\beta \approx 1.98\%$ (model 1 from Table 9), then, a tariff exposure greater than 50.5% would make imports tariff-elastic. That is, for a given product, an increase in the total tariff rate would reduce reported imports in more than one-to-one proportion. Data from the most recent year available, 2021, show that on average, tariff exposure was 27.00% with the median at 26.72%. A substantial number of 3,060,660 imports, about 22.03% for that year, had tariff exposure higher than 50.5% and would present an elasticity larger than one. Figure 6 below represents in a density plot the distribution of total tariff exposure for all Brazilian imports during 2021 alongside the median value and the tariff from where imports become elastic. We can clearly see there is substantial density mass to the right of the elastic imports level.

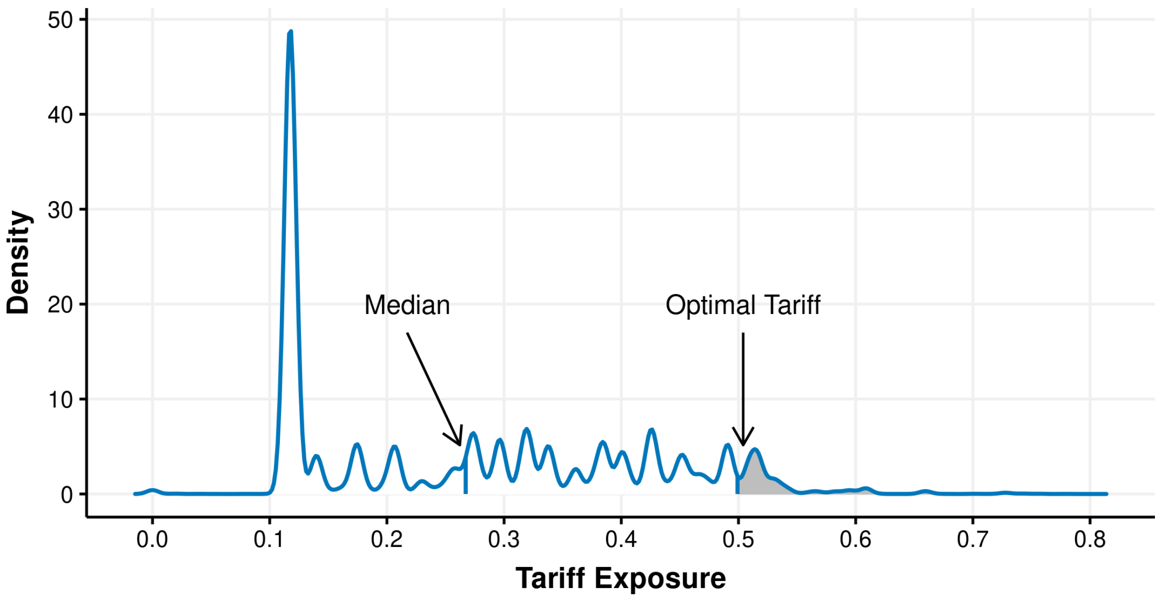

Figure 6: Distribution of tariff exposures for 2021 imports in Brazil.

6 Conclusion

Promoting voluntary tax compliance requires developing a risk-based audit system in which the Tax Authority should focus on high-risk importers to provide the most cost-effective outcome. The recent literature on causal machine learning helps to determine the potential candidates of risk factors that can modulate heterogeneous evasion elasticity of customs duties. Our results contribute to this goal by providing a robust identification of customs duties evasion elasticity heterogeneity stemming from three important characteristics that can be used for risk assessment: NTB, trading companies, and the existence of similar products with lower tariff rates.

We use a novel, confidential administrative dataset containing all Brazilian imports at the transaction level which permits us to identify all imports subject to NTBs to exploit the impact of an exogenous tariff shock. We follow Fisman and Wei (2004) and use the trade gap as our measure of evasion. Since our tariff shock is sharp and exogenous, we make a causal claim: higher tariffs cause an increase in customs duties evasion. Our results show that Brazilian importers also use the quantity margin (i.e., smuggling) to evade tariffs, an unusual result. We argue that a possible explanation is that the combination of tight regulation of prices and high tariffs triggers quantity evasion by importers. We find that importers also exploit the misclassification of high-taxed goods as lower-taxed varieties. Our (value) elasticity of evasion estimated at 1.98 suggests that products taxed at rates higher than 50.5% have elastic imports, implying that lowering rates would increase both trade volume and revenue collection. For the year 2021, about 22% of Brazilian imports had been subject to such high tariff rates.

We also show that a nontariff barrier, import licensing, lowers the elasticity of evasion. The NTB results are not due to a dosage effect, where NTB tends to be more prevalent among low-taxed goods. Importers seem to interpret the NTB as a higher level of enforcement, which is in accordance with our theoretical framework.

Last, we find that trading companies do not display evasive behavior different than ordinary importers, a result not related to greater enforcement on trading companies or postshock endogenous movement into low-taxed products.

References

Ahn, JaeBin, Amit K Khandelwal, and Shang-Jin Wei (2011). "The role of intermediaries in facilitating trade". In: *Journal of International Economics* 84.1, pp. 73–85.

Allingham, Michael G and Agnar Sandmo (1972). "Income tax evasion: A theoretical analysis". In: *Journal of public economics* 1.3-4, pp. 323–338.

Anson, Jose, Olivier Cadot, and Marcelo Olarreaga (2006). "Tariff evasion and customs corruption: Does pre-shipment inspection help?" In: *The BE Journal of Economic Analysis & Policy* 5.1.

Antras, Pol and Elhanan Helpman (2004). "Global sourcing". In: *Journal of political Economy* 112.3, pp. 552–580.

Arbex, Marcelo and Enlinson Mattos (2015). "Optimal sales tax rebates and tax enforcement consumers". In: *Oxford Economic Papers* 67.2, pp. 479–493.

Athey, Susan and Guido Imbens (2016). "Recursive partitioning for heterogeneous causal effects". In: *Proceedings of the National Academy of Sciences* 113.27. Publisher: National Academy of Sciences, pp. 7353–7360.

Athey, Susan, Julie Tibshirani, and Stefan Wager (2019). "Generalized random forests". In: *The Annals of Statistics* 47.2. Publisher: Institute of Mathematical Statistics, pp. 1148–1178.

Bao, Xiaohua and Larry D. Qiu (2010). "Do Technical Barriers to Trade Promote or Restrict Trade? Evidence from China". In: *Asia-Pacific Journal of Accounting & Economics* 17.3, pp. 253–278. DOI: 10.1080/16081625.2010.9720865. eprint: https://doi.or g/10.1080/16081625.2010.9720865. URL: https://doi.org/10.1080/16081625.2010.9720865.

Bouet, Antoine and Devesh Roy (2012). “Trade protection and tax evasion: Evidence from Kenya, Mauritius, and Nigeria”. In: *The Journal of International Trade & Economic Development* 21.2, pp. 287–320.

Breiman, Leo (2001). “Random forests”. In: *Machine learning* 45.1. Publisher: Springer, pp. 5–32.

Chetty, Raj (2009). “Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance”. In: *American Economic Journal: Economic Policy* 1.2, pp. 31–52.

Cowell, F. A. (1990). “Tax Sheltering and the Cost of Evasion”. In: *Oxford Economic Papers* 42.1, pp. 231–243. ISSN: 00307653, 14643812. URL: http://www.jstor.org/stable/2663356.

De Sousa, José, Thierry Mayer, and Soledad Zignago (2012). “Market access in global and regional trade”. In: *Regional Science and Urban Economics* 42.6, pp. 1037–1052.

Demir, Banu and Beata Javorcik (2020). “Trade policy changes, tax evasion and Benford’s law”. In: *Journal of Development Economics*, p. 102456.

Dowd, Tim, Paul Landefeld, and Anne Moore (2017). “Profit shifting of US multinationals”. In: *Journal of Public Economics* 148, pp. 1–13.

Egger, Peter et al. (2015). “Non-tariff barriers, integration and the transatlantic economy”. In: *Economic Policy* 30.83, pp. 539–584.

Ferrantino, Michael J (2006). “Quantifying the trade and economic effects of non-tariff measures”. In:

Fisman, Raymond, Peter Moustakerski, and Shang-Jin Wei (2008). “Outsourcing tariff evasion: A new explanation for entrepôt trade”. In: *The Review of Economics and Statistics* 90.3, pp. 587–592.

Fisman, Raymond and Shang-Jin Wei (2004). “Tax rates and tax evasion: evidence from “missing imports” in China”. In: *Journal of political Economy* 112.2, pp. 471–496.

Fuest, Clemens, Felix Hugger, and Florian Neumeier (2022). “Corporate profit shifting and the role of tax havens: Evidence from German country-by-country reporting data”. In: *Journal of Economic Behavior & Organization* 194, pp. 454–477.

Ghodsi, Mahdi et al. (2017). *The evolution of non-tariff measures and their diverse effects on trade*. Tech. rep. wiiw Research Report.

Gordon, Roger and Wei Li (2009). “Tax structures in developing countries: Many puzzles and a possible explanation”. In: *Journal of public Economics* 93.7-8, pp. 855–866.

Grübler, Julia and Oliver Reiter (2021). “Characterising non-tariff trade policy”. In: *Economic Analysis and Policy* 71, pp. 138–163.

Head, Keith and Thierry Mayer (2014). “Gravity equations: Workhorse, toolkit, and cookbook”. In: *Handbook of international economics*. Vol. 4. Elsevier, pp. 131–195.

Hines James R., Jr. and Eric M. Rice (Feb. 1994). “Fiscal Paradise: Foreign Tax Havens and American Business*”. In: *The Quarterly Journal of Economics* 109.1, pp. 149–182. ISSN: 0033-5533. DOI: 10.2307/2118431. eprint: https://academic.oup.com/qje/article-pdf/109/1/149/5233163/109-1-149.pdf. URL: https://doi.org/10.2307/2118431.

Imbens, Guido W and Donald B Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Imbens, Guido W. and Jeffrey M. Wooldridge (Mar. 2009). “Recent Developments in the Econometrics of Program Evaluation”. In: *Journal of Economic Literature* 47.1, pp. 5–86. DOI: 10.1257/jel.47.1.5. URL: https://www.aeaweb.org/articles?id=10.1257/jel.47.1.5.

Javorcik, Beata S and Gaia Narciso (2008). “Differentiated products and evasion of import tariffs”. In: *Journal of International Economics* 76.2, pp. 208–222.

– (2017). “WTO accession and tariff evasion”. In: *Journal of Development Economics* 125, pp. 59–71.

Kaplow, Louis (1990). “Optimal taxation with costly enforcement and evasion”. In: *Journal of Public Economics* 43.2, pp. 221–236. ISSN: 0047-2727. DOI: https://doi.org/10.1016/0047-2727(90)90031-C. URL: https://www.sciencedirect.com/science/article/pii/004727279090031C.

Kee, Hiau Looi and Alessandro Nicita (2016). “Trade frauds, trade elasticities and non-tariff measures”. In: *5th IMF-World Bank-WTO Trade Research Workshop, Washington, DC, November*. Vol. 30.

Kinzius, Luisa, Alexander Sandkamp, and Erdal Yalcin (2019). “Trade protection and the role of non-tariff barriers”. In: *Review of World Economics* 155.4, pp. 603–643.

Kleven, Henrik Jacobsen et al. (2011). “Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark”. In: *Econometrica* 79.3, pp. 651–692.

Kopczuk, Wojciech and Joel Slemrod (2006). “Putting firms into optimal tax theory”. In: *American Economic Review* 96.2, pp. 130–134.

Kume, Honorio, Guida Piani, and Pedro Miranda (2011). “Tarifas de importação e evasão fiscal no Brasil”. In: *Economia Aplicada* 15, pp. 65–82.

Mayshar, Joram (1990). “On measures of excess burden and their application”. In: *Journal of Public Economics* 43.3, pp. 263–289.

Mishra, Prachi, Arvind Subramanian, and Petia Topalova (2008). “Tariffs, enforcement, and customs evasion: Evidence from India”. In: *Journal of public Economics* 92.10-11, pp. 1907–1925.

Naritomi, Joana (2019). “Consumers as tax auditors”. In: *American Economic Review* 109.9, pp. 3031–72.

Pomeranz, Dina (2015). “No taxation without information: Deterrence and self-enforcement in the value added tax”. In: *American Economic Review* 105.8, pp. 2539–69.

Rauch, James E (1999). “Networks versus markets in international trade”. In: *Journal of international Economics* 48.1, pp. 7–35.

Slemrod, Joel (2001). "A general model of the behavioral response to taxation". In: *International Tax and Public Finance* 8.2, pp. 119–128.

Slemrod, Joel et al. (2017). "Does credit-card information reporting improve small-business tax compliance?" In: *Journal of Public Economics* 149, pp. 1–19.

Sousa, Maria da Conceição Sampaio de, Maria Eduarda Tannuri-Pianto, and Carlos Antônio Silva dos Santos (2008). "Imposto de importação e evasão fiscal: uma investigação do caso brasileiro". In: *Revista Brasileira de Economia* 62, pp. 77–93.

Wager, Stefan and Susan Athey (2018). "Estimation and inference of heterogeneous treatment effects using random forests". In: *Journal of the American Statistical Association* 113.523. Publisher: Taylor & Francis, pp. 1228–1242.

Werneck, Rogério LF (2006). "An evaluation of the 2003 tax reform effort in Brazil". In: *Brazilian Journal of Political Economy* 26, pp. 75–94.

WTO, World Trade Organization (2012). *Trade and public policies: A closer look at non-tariff measures in the 21 st century*. World Trade Report.

Yang, Dean (2008). "Can enforcement backfire? Crime displacement in the context of customs reform in the Philippines". In: *The Review of Economics and Statistics* 90.1, pp. 1–14.

Yitzhaki, Shlomo (Jan. 1974). "A note on income tax evasion". In: *Public Finance Quarterly* 15, pp. 123–137.

A Appendix

A.1 Elastic imports

Consider the case where the amount M responds to the enforcement level, $M(E)$, with $M, E > 0$. The cost of evasion is dependent on the fraction smuggled and enforcement, $C(\gamma, E) > 0$, and the same assumptions in the main text hold. The firm's objective function is the net benefit of evasion and is given by: $B = \gamma M(E) \cdot T - C(\gamma, E)$. The first-order condition (FOC) is the usual expression and defines implicitly the optimal level of evasion, $M(E) \cdot T = C_1(\gamma^*, E)$. Moreover, the semi-elasticity of evasion with relation to the tariff rate is unchanged, $\partial\gamma^*/\partial T = M(E)/C_{11}(\gamma, E) > 0$.

Next, if we differentiate the FOC with relation to enforcement we will have a slightly different expression

$$\frac{\partial\gamma^*}{\partial E} = \frac{T \cdot M_1 - C_{12}}{C_{11}} \tag{11}$$

and $C_{12} > 0$ by assumption (i.e. marginal cost of evasion is increasing in enforcement). Let's define the reported import elasticity w.r.t. enforcement as $\sigma_M^E \equiv \frac{dM}{dE} \cdot \frac{E}{M}$. Substitute this definition and the FOC into equation (11) and we have the expression for how evasion changes with enforcement.

$$\frac{\partial \gamma^*}{\partial E} = \frac{\sigma_M^E C_1/E - C_{12}}{C_{11}} \tag{12}$$

Equation (12) will be lower than zero whenever $\sigma_M^E < E \cdot \frac{C_{12}}{C_1}$. The right-hand side of this inequality is greater than zero by our assumptions and, if we consider the case where enforcement is proxied by a nontariff barrier, we expect $\sigma_M^E < 0$ according to the literature, (Ghodsi et al. 2017; Kinzius, Sandkamp, and Yalcin 2019). Hence, the inequality holds as was the case with inelastic imports. To analyze how enforcement changes the elasticity of evasion, $\partial^2\gamma^*/\partial E \partial T$, we differentiate the FOC with relation to tariff and enforcement to arrive at the following:

$$\frac{\partial^2 \gamma^*}{\partial E \partial T} = \frac{M}{E \cdot C_{11}} \left[\sigma_M^E - \sigma_C^E\right] \tag{13}$$

where $\sigma_C^E \equiv \frac{\partial C_{11}}{\partial E}\frac{E}{C_{11}}$. Consider the case for an NTB. Equation (13) is negative if $\sigma_C^E > 0$, since $\sigma_M^E < 0$, which is again the same result as in the main text, $\partial C_{11}/\partial E > 0$.

B Appendix

B.1 Descriptive plots

Figure B.1 shows the distribution of total tariff exposure by tariff quintile and year. Notice how 2004 has, on average, higher values of tariff exposure due to the introduction of PIS and COFINS on imports.

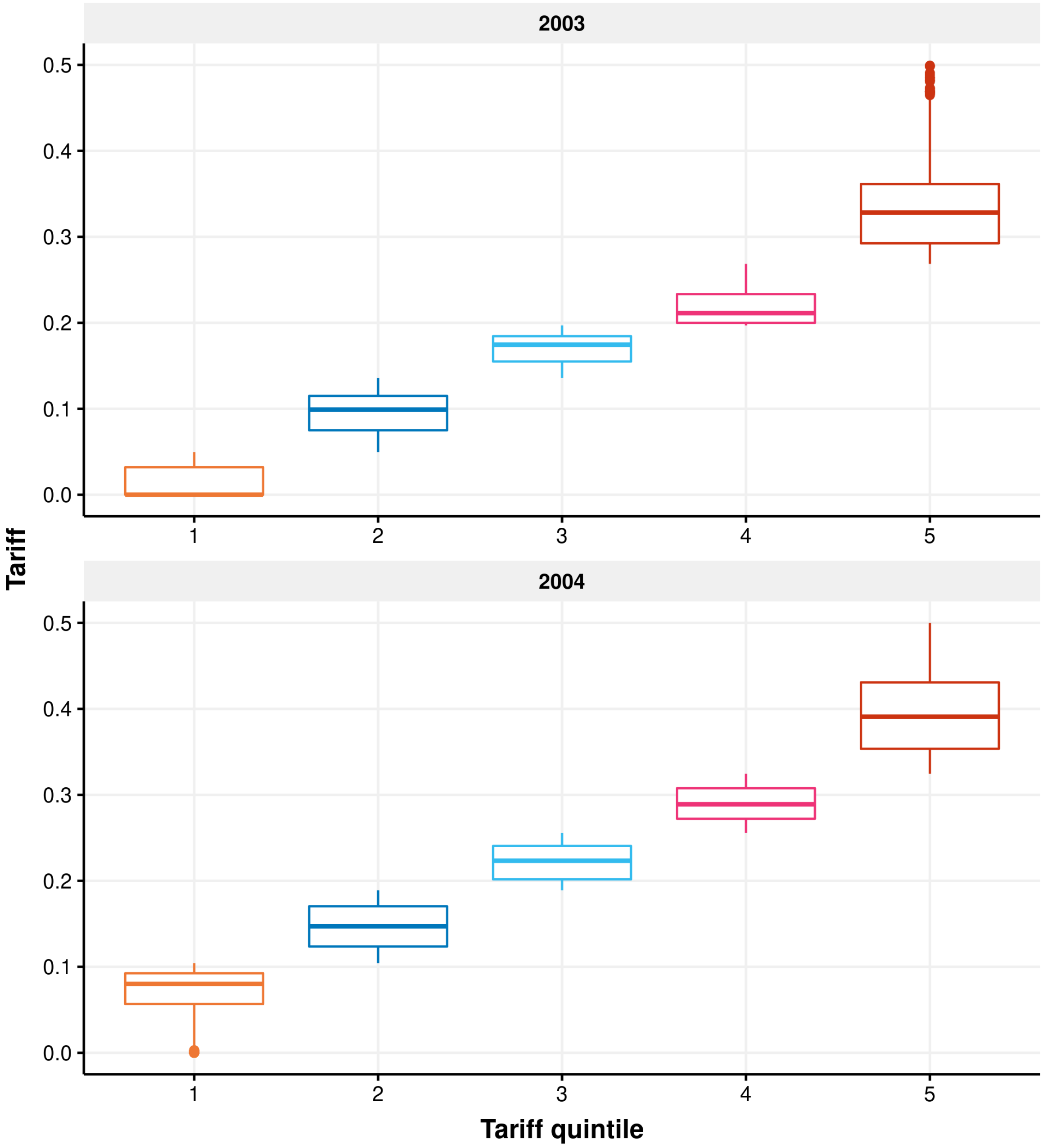

Figure B.1: Distribution of total tariff rates by year and tariff quintile

The mean trade gap by tariff quintile shows how the misclassification of products may occur. Low-taxed products that are similar to high-taxed ones lie mostly on quintiles one and two, exactly those that show negative trade gaps (i.e. more imports reported than exports), see Figure B.2.

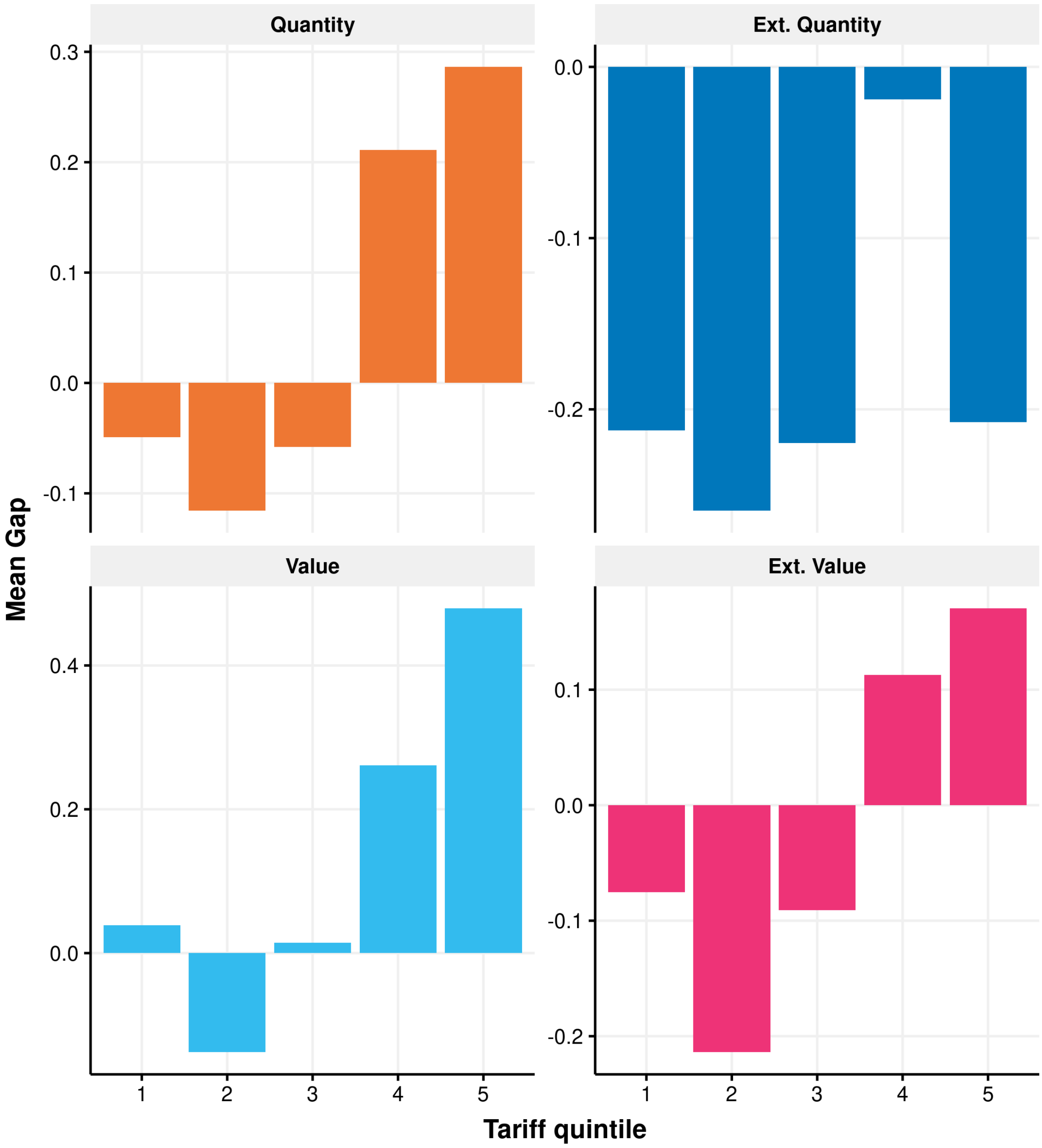

Figure B.2: Mean trade gap by tariff quintile.

Last, Table B.1 illustrates how misleading the MFN tariff from UNCTAD/TRAINS can be.

Table B.1: Averages of MFN and EAR tariff rates

Products	Description	FOB Value	Avg. MFN	Avg. EAR
270900	Petroleum oils and oils, crude	11,106,237	0.0000	0.0324
854221	Electronic integrated circuits	2,596,973	0.0157	0.0768
852990	Reception and transmission apparatus	2,222,516	0.1129	0.1255

Table B.1: Averages of MFN and EAR tariff rates *(continued)*

Products	Description	FOB Value	Avg. MFN	Avg. EAR
271019	Oils not containing biodiesel, not crude, not waste oils	2,093,829	0.0118	0.0654
271011	Light oils and preparations	1,857,186	0.0000	0.0568

Notes:
Products classified according to their 6-digit HS code. Brief description of products enclosed by the HS code is provided, not the legal definition. FOB values expressed in thousand dollars.

B.2 Robustness of main results

We test the robustness of misclassification results by modifying the tariff rate of the typical "similar" product, from the minimum to the average leave-one-out tariff within the same HS 4-digit code in Table B.2.

Next, we run a series of regressions starting from a baseline specification and incrementally including covariates to arrive at our full specification shown in Table 9. The baseline includes as regressors only the tariff exposure and the minimum tariff of similar products. We add tariff agreement, anti-dumping, drawback, PIS/COFINS exemption, audits, average exchange rate, proportion of trading companies, average insurance and freight costs, number of imports and importers, average importer's years in activity and average days until clearance, to come to the administrative specification. Lastly, the States specfication, additionaly include the proportion of imports made through each Brazilian State[45].

[45] The state of Tocantins did not record any imports in both periods. São Paulo is chosen as the reference State.

Table B.2: Regression results including average tariffs on similar products

	Minimum Similar Tariff				Average Similar Tariff			
Dep. Var.:	Value	Ext. Value	Quantity	Ext. Quantity	Value	Ext. Value	Quantity	Ext. Quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tariff	1.98403***	1.54441***	0.96569**	0.86792**	1.64516***	1.29039***	0.81096*	0.74732
	(0.30817)	(0.23818)	(0.44010)	(0.40590)	(0.33648)	(0.24710)	(0.47171)	(0.45494)
Similar	-0.49600**	-0.38939**	-0.28776	-0.25464	-0.22517	-0.21610	0.08774	0.11362
	(0.21309)	(0.14938)	(0.32077)	(0.28747)	(0.30372)	(0.23440)	(0.43545)	(0.42503)
Num.Obs.	63226	64481	57267	59023	53213	54312	48019	49502
R2 Adj.	0.538	0.570	0.546	0.619	0.529	0.568	0.529	0.612
FE: cp	✓	✓	✓	✓	✓	✓	✓	✓
FE: ct	✓	✓	✓	✓	✓	✓	✓	✓
FE: pt	✓	✓	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariffs on similar products and its own are calculated as a simple average from transaction-level data by origin country, year and 6-digit product. All standard errors are two-way clustered at country and product.

Table B.3: Robustness assessment to covariates inclusion

	Baseline				Administrative				States			
Dep. Var.:	Value	Ext. Value	Quantity	Ext. Quantity	Value	Ext. Value	Quantity	Ext. Quantity	Value	Ext. Value	Quantity	Ext. Quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Tariff	2.43***	1.92***	1.45***	1.31***	2.04***	1.54***	0.92**	0.79*	1.98***	1.54***	0.97**	0.87**
	(0.32)	(0.23)	(0.42)	(0.38)	(0.33)	(0.24)	(0.44)	(0.40)	(0.31)	(0.24)	(0.44)	(0.41)
Similar	-0.54**	-0.42***	-0.36	-0.35	-0.51**	-0.40***	-0.31	-0.29	-0.50**	-0.39**	-0.29	-0.25
	(0.22)	(0.16)	(0.33)	(0.29)	(0.21)	(0.15)	(0.32)	(0.29)	(0.21)	(0.15)	(0.32)	(0.29)
Num.Obs.	63458	64715	57457	59227	63226	64481	57267	59023	63226	64481	57267	59023
R2 Adj.	0.532	0.565	0.543	0.616	0.536	0.568	0.545	0.618	0.538	0.570	0.546	0.619
FE: cp	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
FE: ct	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
FE: pt	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The baseline specification include as regressors only the tariff exposure and the minimum tariff of similar products. The administrative specification is incremental, adding tariff agreement, anti-dumping, drawback, PIS/COFINS exemption, audits, exchange rate, tradings, insurance and freight cost, number of imports and importers and the average company's years in activity and days until the import is cleared. States specification additionally include the proportion of imports made through each Brazilian State. Tariffs on similar products and its own are calculated as a simple average from transaction-level data by origin country, year and 6-digit product. Similar products have the same first 4-digits. All standard errors are two-way clustered at country and product.

B.3 Robustness of NTB results

We present at Table B.4 a balance table of importers' characteristics by the presence of NTB on their imports during a given year. The averages of the exchange rate in the import date, freight value in thousand dollars, number of products (8-digit NCM) in an import process, and, number of import processes are presented. Besides, we also show the proportion of products subject to anti-dumping measures, drawback regimes, physical inspection, and tariff agreements. In that table, we provide two measures to assess the overlap of relevant covariates between the two importer groups, those without any import license (i.e. No NTB) associated with their transactions and importers that had at least one product subject to such license. The measures are the difference in means and the normalized difference as described in Guido W Imbens and Rubin (2015).

Table B.4: Importer characteristics by NTB presence

	NTB ($N_t = 22246$)		No NTB ($N_c = 23971$)		Overlap Measures				
Variable	Mean	Std.Dev	Mean	Std.Dev	t-stat	Norm. Diff.	Log Ratio	$\pi_c^{0.05}$	$\pi_t^{0.05}$
Avg. exchange rate	2.99	0.13	3.00	0.16	-3.90	-0.04	-0.20	0.09	0.03
Avg. FOB	1.30	1.85	0.58	1.83	41.78	0.39	0.01	0.07	0.07
Avg. freight	2.52	31.05	0.99	9.25	7.05	0.07	1.21	0.04	0.09
Avg. #products	2.36	3.44	2.14	2.58	7.98	0.07	0.29	0.46	0.38
Avg. #imports	97.32	527.62	12.54	33.90	23.92	0.23	2.75	0.26	0.31
Prop. anti-dumping	0.00	0.02	0.00	0.00	5.88	0.06	1.76	1.00	1.00
Prop. drawback	0.08	0.23	0.00	0.00	50.88	0.48	Inf	1.00	1.00
Prop. audit	0.32	0.33	0.32	0.38	-0.44	0.00	-0.15	0.36	0.17
Prop. tariff agreement	0.15	0.31	0.08	0.25	26.45	0.25	0.23	0.87	0.68

Notes:

Variables are either average or proportion values taken over importer-year. Avg. number of products represents the mean number of 8-digit products included in an import process. Average number of imports is the mean number of import processes. All other variables are mean values per 8-digit products. Monetary values (FOB and freight) are given in logarithm of thousand dollars.

Robustness assessments for our NTB regression results follow. We start from a baseline specification and incrementally including covariates to arrive at our full specification shown in Table 16. The covariates included in the baseline, administrative and states specifications are the same as described in Sub-section B.2.

The aforementioned controls' regression coefficients are not shown due to space and clarity considerations in Table B.5. Figure B.3 gives us a clear picture of coefficients estimation robustness.

Then, we present the permutation test for the interaction term $Tariff \times NTB$. The proportions of transactions made under import license were shuffled 10^4 times and for each permutation, the associated interaction coefficient is extracted and plotted in the histogram presented in Figure B.4. The dot-dashed line represents our estimated coefficient, indicating the results are not due to mere chance.

Ad valorem equivalent of imports subject to NTBs is simulated considering an increase of 20% in the tariff rate. Tariffs are increased because nontariff barriers reduce trade flows, an effect similar to a tariff increase. Our results are shown in Table B.6.

Table B.5: NTB effect on evasion. Robustness assessment to covariates inclusion

	Baseline				Administrative				States			
Dep. Var.:	Value	Ext. Value	Quantity	Ext. Quantity	Value	Ext. Value	Quantity	Ext. Quantity	Value	Ext. Value	Quantity	Ext. Quantity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Tariff	3.00***	2.45***	2.34***	2.33***	2.62***	2.15***	1.87***	1.82***	2.62***	2.16***	1.90***	1.88***
	(0.25)	(0.18)	(0.37)	(0.38)	(0.26)	(0.18)	(0.38)	(0.35)	(0.28)	(0.20)	(0.38)	(0.36)
Tariff X NTB	-1.68***	-1.28***	-1.76***	-1.62***	-1.80***	-1.38***	-1.92***	-1.77***	-1.86***	-1.42***	-1.99***	-1.83***
	(0.38)	(0.26)	(0.45)	(0.44)	(0.39)	(0.28)	(0.49)	(0.46)	(0.39)	(0.27)	(0.49)	(0.46)
Similar	-0.69***	-0.64***	-0.79***	-0.75***	-0.62***	-0.60***	-0.70**	-0.68**	-0.63***	-0.60***	-0.72**	-0.69**
	(0.23)	(0.16)	(0.29)	(0.28)	(0.23)	(0.16)	(0.28)	(0.27)	(0.23)	(0.16)	(0.28)	(0.27)
NTB	-0.19**	-0.17***	-0.06	-0.05	-0.02	-0.05	0.18	0.14	0.02	-0.02	0.23*	0.17
	(0.09)	(0.06)	(0.11)	(0.10)	(0.09)	(0.06)	(0.13)	(0.12)	(0.09)	(0.07)	(0.14)	(0.13)
Num.Obs.	63458	64715	57457	59227	63226	64481	57267	59023	63226	64481	57267	59023
Avg. Tariff	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
Avg. NTB	0.27	0.27	0.27	0.27	0.26	0.26	0.27	0.27	0.26	0.26	0.27	0.27
R2 Adj.	0.115	0.181	0.178	0.304	0.119	0.186	0.183	0.309	0.121	0.187	0.185	0.311
FE: c	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
FE: p	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
FE: t	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The baseline specification include as regressors only the tariff exposure and the minimum tariff of similar products. The administrative specification is incremental, adding tariff agreement, anti-dumping, drawback, PIS/COFINS exemption, audits, exchange rate, tradings, insurance and freight cost, number of imports and importers and the average company's years in activity and days until the import is cleared. States specification additionally include the proportion of imports made through each Brazilian State. Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product from the administrative dataset. NTB is the proportion of imports subject to non-automatic licensing for a given triplet, country, product, and year. All standard errors are two-way clustered at country and product.

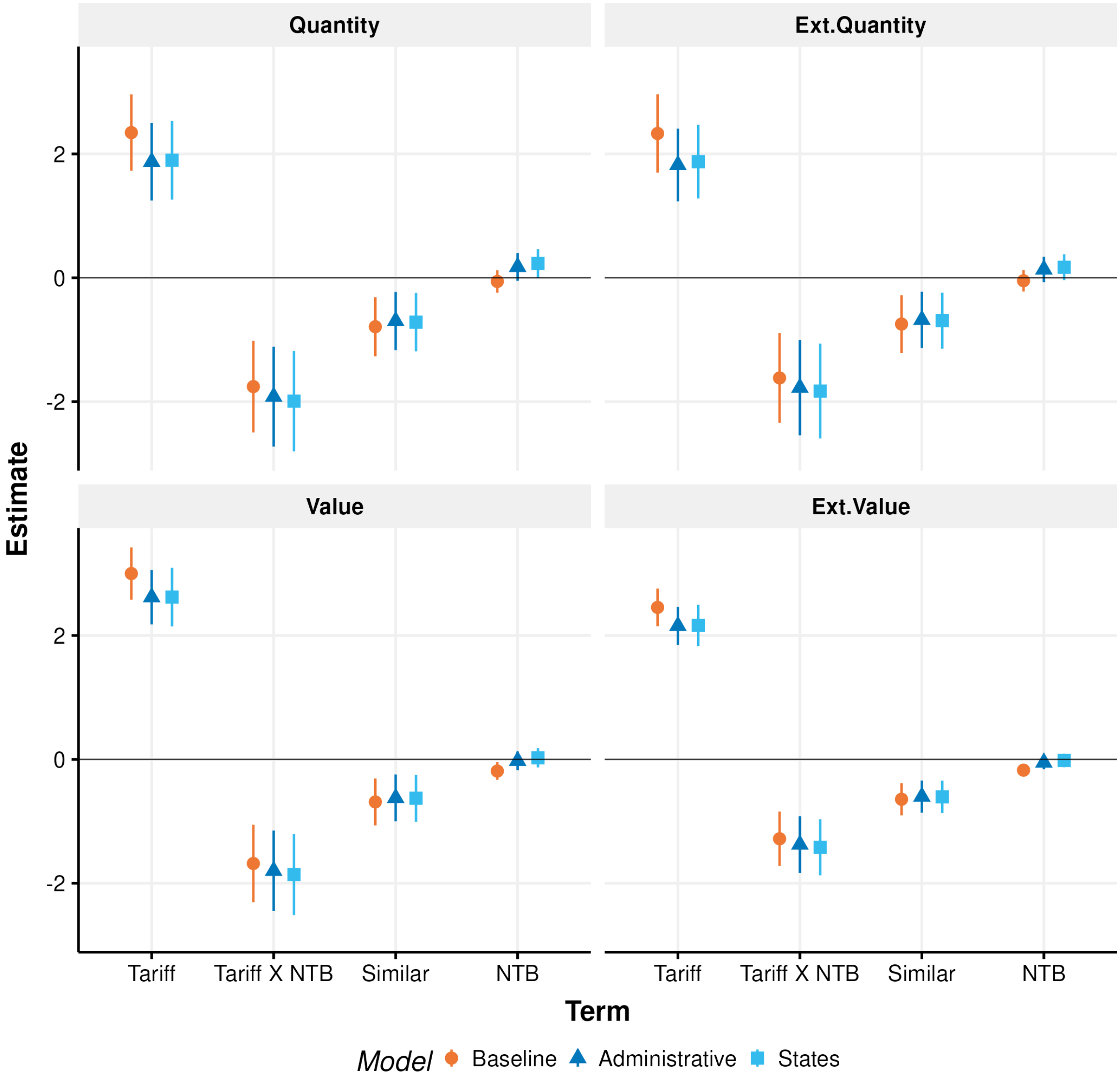

Figure B.3: Robustness assessment for different specifications of NTB effect on evasion elasticity. Model NTB does not include controls X_{cpt}.

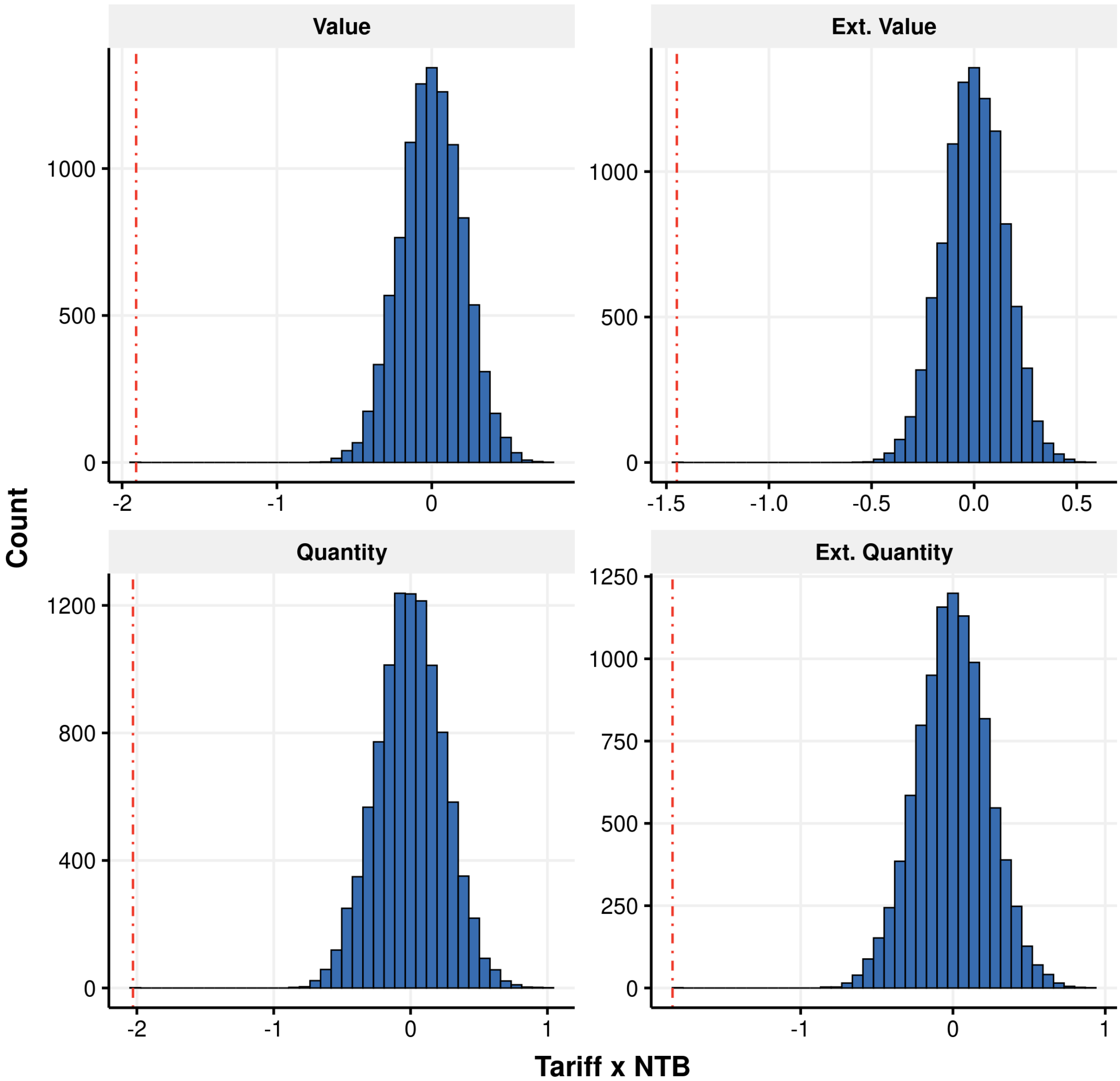

Figure B.4: Permutation test for the interaction coefficient, $Tariff \times NTB$, from Table 16. 10^4 permutations.

Table B.6: Results of ad valorem equivalent of non-automatic import licensing on evasion

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	2.5240***	2.0991***	1.7822***	1.7643***
	(0.2849)	(0.1986)	(0.3826)	(0.3547)
Tariff X NTB	-1.7413***	-1.3809***	-1.6957***	-1.5621***
	(0.3435)	(0.2387)	(0.4297)	(0.4044)
Similar	-0.6229***	-0.6016***	-0.7101**	-0.6873**
	(0.2282)	(0.1578)	(0.2840)	(0.2725)
NTB	-0.0231	-0.0506	0.1793	0.1165
	(0.0905)	(0.0628)	(0.1348)	(0.1234)
Num.Obs.	63226	64481	57267	59023
Avg. Tariff	0.1983	0.1996	0.1971	0.1986
Avg. NTB	0.2647	0.2646	0.2696	0.2684
R2 Adj.	0.120	0.187	0.185	0.311
FE: c	✓	✓	✓	✓
FE: p	✓	✓	✓	✓
FE: t	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Tariff is calculated as a simple average from transaction-level data by origin country, year and 6-digit product from the administrative dataset. NTB is the proportion of imports subject to non-automatic licensing for a given triplet, country, product, and year. All standard errors are two-way clustered at country and product. All regressions include additional control variables. The ad valorem equivalent is simulated by increasing the tariff rate of an import subject to the import license by 20%. Those imports are then aggregated by HS6 codes.

C Appendix

C.1 Trading companies

Inspired by Fisman, Moustakerski, and Wei (2008), which explores the role of expert knowledge in facilitating smuggling in China's indirect trade through Hong Kong's warehouses and our machine learning model in the Appendix D, which reveals that trading companies can influence evasion elasticity in our data, we control for proportion of trading companies in all of our regressions.

Trading companies specialize in export and import procedures and operate on behalf of their clients.In that case, the trading company is acting as an intermediary that charges a commission on trading operations. Another common activity for such companies is to represent different international brands or products locally which means that trading companies may also operate in their own capacity, importing products for further redistribution. In that case, trading companies can map suppliers in foreign countries, negotiate the terms of sale and deliver the imported goods, deal with customs officials and the processes of international trade. These companies know better than the average importer the inner workings of international trade, shipments, and customs legislation, which poses the question of whether these companies respond differently to tariff increases, either by the greater possibility of misclassifying the goods into a less taxed code or even exploring evasion alternatives unknown by an ordinary importer.

Our Table C.1 shows that trading companies do not evade different from the ordinary importer, either considering the regular definition of trade gap or the extreme one, as depicted by the lack of significance in the estimated coefficients associated with the proportion of imports made by those companies (i.e., *Trading*).

The quantity evasion results confirm these findings. Trading companies do not appear to have any distinct behavior, either in level or elasticity responses, even though the pointwise level estimations (i.e., $Trading$ coefficient) are positive for quantity and have the opposite sign for evasion values.

Regarding customs duties, trading companies have the same level of compliance, and their evasion elasticity in relation to tariff rates is not distinguishable from the other companies operating in foreign trade. Although not being a source of elasticity evasion heterogeneity, the presence of trading companies is important in international trade and all our regressions in this paper include the proportion of trading companies as a control variable.

Table C.1: Trading companies and misclassification

Dep. Var.:	Value (1)	Ext. Value (2)	Quantity (3)	Ext. Quantity (4)
Tariff	2.4148***	1.9444***	1.4404***	1.3465***
	(0.3565)	(0.2491)	(0.4642)	(0.4250)
Tariff X Trading	0.0344	-0.0967	0.1812	0.0145
	(0.5301)	(0.3237)	(0.6059)	(0.5159)
Trading	-0.0146	-0.0235	0.1956	0.2007
	(0.1456)	(0.0954)	(0.1727)	(0.1580)
Similar	-0.5401**	-0.4157***	-0.3699	-0.3496
	(0.2180)	(0.1569)	(0.3234)	(0.2923)
Num.Obs.	63458	64715	57457	59227
Avg. Tariff	0.1907	0.1920	0.1893	0.1909
Avg. Trading	0.3485	0.3475	0.3374	0.3358
R2 Adj.	0.532	0.565	0.543	0.616
FE: cp	✓	✓	✓	✓
FE: ct	✓	✓	✓	✓
FE: pt	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes:
Tariff is the simple average from transaction-level data by origin country, year and 6-digit product. Trading is the proportion of imports made by trading companies for a given triplet, country, year and product. Tariff on similar products is the minimum tariff at the 4-digit level. All standard errors are two-way clustered at country and product.

D Appendix

D.1 Machine learning model

We use the new machine learning algorithm of generalized random forest (GRF), (Athey, Tibshirani, and Wager 2019), to estimate the heterogeneous treatment effects of tariffs on evasion using a causal forest (CF), (Wager and Athey 2018). The GRF algorithm extends the classical random forest of Breiman (2001) to allow the estimation of other statistical quantities and causal effects among them. Causal forests are ideal to assess the treatment heterogeneity found in this paper as it is first designed for that purpose (i.e., flexibly detect treatment heterogeneity), and second, it is a non-parametric method with minimal functional form assumptions.

Although a causal forest is a highly non-linear model and not directly comparable to linear regression, we want to make it close enough to our main linear regression specification.

Therefore, we choose to first demean the outcomes, tariff exposure, and other relevant variables by the interacted fixed-effects used in equation (8). Control variables included in the forest are our effect modulators, NTB and $Trading$, the $Similar$ products minimum tariff, administrative treatments, the average exchange rate, and the States' share in several imports. The GRF was fit with the following parameter settings: $\alpha = 0.15$, $min.node.size = 20$ and $2,000$ trees.

Table D.1 shows the average treatment effect (ATE) found in our training sample for each evasion measure. Those are the average evasion's semi-elasticity concerning tariffs and, albeit carrying the same economic meaning as previous linear regressions' coefficient on $Tariff$, we do not expect the ATE from the causal forest to be necessarily equal to the elasticity found in Table 9[46]. That is because the weighting scheme used in GRF is different than ordinary least squares (OLS). GRF creates a list of neighboring training examples weighted by how many times this example fell in the same leaf as the test example and, predicts the causal effect for this test example using the outcomes and treatment status of the neighboring examples. The ATE is finally computed by plugging causal forest predictions into a doubly robust average treatment effect estimator.

Table D.1: Average treatment effects

	(1)	(2)	(3)	(4)
ATE	1.7852	1.4173	0.9640	0.7858
	(0.2220)	(0.1721)	(0.2843)	(0.2700)
Num.Obs.	63457	64714	57462	59233
Dep. Var.	Value	Ext. Value	Quantity	Ext. Quantity
RMSE	0.7327	0.5766	0.8946	0.8776

Notes:
Doubly-robust estimates of the average partial effect, E[Cov[T, Y | X] / Var[T | X]], via augmented inverse-propensity weighting. Standard errors in parenthesis.

Table D.2 below shows an omnibus evaluation of the quality of the causal forest outcome predictions. It is intended to show if the forest is accurately predicting the outcome and, whether it was able to detect heterogeneity. Forest predictions appear to be accurate for all models, as shown by close to one $Prediction$ coefficient, although, heterogeneity was found only in value evasion.

[46] The appropriate model specification comparison of our CF is equation (8) once we have included the minimum tariff of similar products into the training features.

Table D.2: Causal forests goodness of fit

	(1)	(2)	(3)	(4)
Prediction	0.96347	1.01107	1.04030	0.99587
	(0.15483)	(0.12890)	(0.31368)	(0.35251)
Heterogeneity	1.67483	1.12217	0.82332	-0.20276
	(0.59721)	(0.61981)	(0.63012)	(0.74068)
Dep. Var.	Value	Ext. Value	Quantity	Ext. Quantity
Num.Obs.	63457	64714	57462	59233

Notes:
This is a test calibration of the forest. A coefficient of 1 for Prediction suggests that the mean forest prediction is correct. Also, if the Heterogeneity coefficient is significantly greater than 0, then we can reject the null of no heterogeneity. Robust standard-errors in parenthesis.

Causal forests are tailor-made to detect effect heterogeneity and, with this machinery, we can investigate which features (i.e., machine learning jargon for covariates) contribute to these heterogeneous effects. Based on the work of Athey and G. Imbens (2016), a CF individual tree is grown according to the goodness of a split, which is determined by how much it increases heterogeneity in the causal effect. Feature x importance is a weighted sum of how many times the feature was used to split a node at each depth in the forest and, since the split criterion is the heterogeneity, features that must appear at the top nodes are the ones contributing the most to the heterogeneous effect.

We need to be careful when analyzing feature importance for heterogeneity based on the number of splits the feature appears. It is important to realize that if a pair of features are highly correlated, the tree might split on any of those without materially affecting the final result. That is, it is possible just by chance, the tree algorithm splits always on one feature disregarding its correlated pair, and we would interpret the second feature as not important even though it carries the same information as the one used in the splits.

Figure D.1 shows a correlation matrix of all selected features we are using to grow our causal forest. Some positive correlations are worth noticing, like tariff agreement with the State RS and, NTB is lightly correlated to the drawback regime. On the negative correlations side, we have a tariff agreement and NTB correlating to the minimum similar tariff rate. In any case, the correlations are not close to $\{-1, 1\}$ and we should not worry about the tree splits choosing only one feature in the deterrence of another highly correlated one.

Figure D.2 brings the top eight most important features for each evasion measure. The maximum tree depth considered was four, with a decay factor of two. Features typically showing on the top rankings include the minimum tariff rate on *Similar* products, the

NTB, and the proportion of $Trading$ companies. In the linear regression setup we explicitly included those variables in the model, either as control (i.e., Similar) or as an effect modulator (i.e., NTB and Trading), in this machine learning setup on the other hand, these features naturally came up as a result of the data-driven algorithm.

In Figure D.3 we analyze how the three features, $Similar$, NTB, and $Trading$ change their importance with varying values for the split imbalance parameter α. The method does not provide confidence intervals, but qualitatively we could argue the $Similar$ feature is the most important driver of effect heterogeneity, while NTB comes in second when quantity evasion (and extreme quantity) is considered but, it is tied with $Trading$ for value and extreme value evasion.

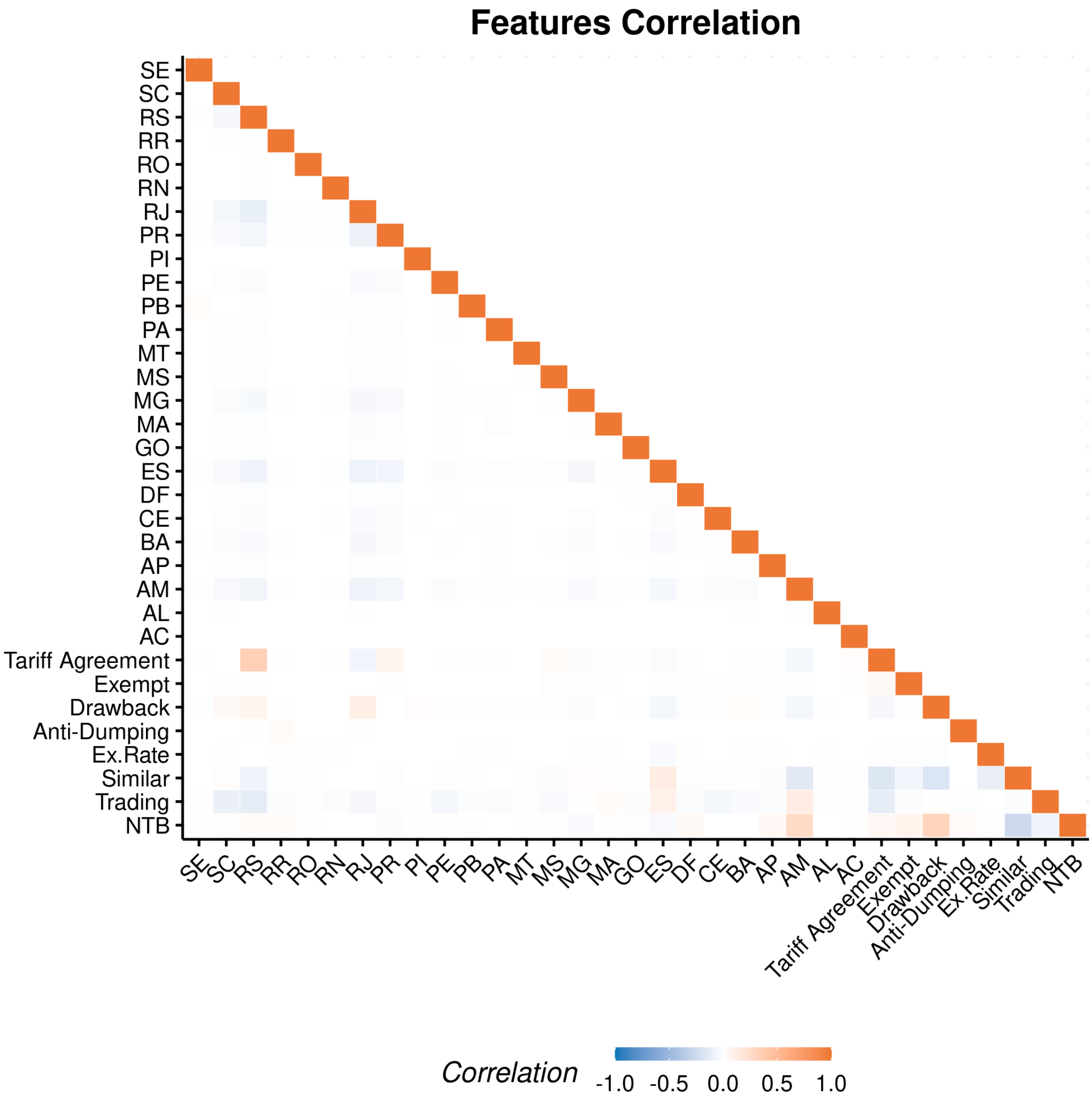

Figure D.1: Pearson's correlation matrix of features showing absence of strong correlations

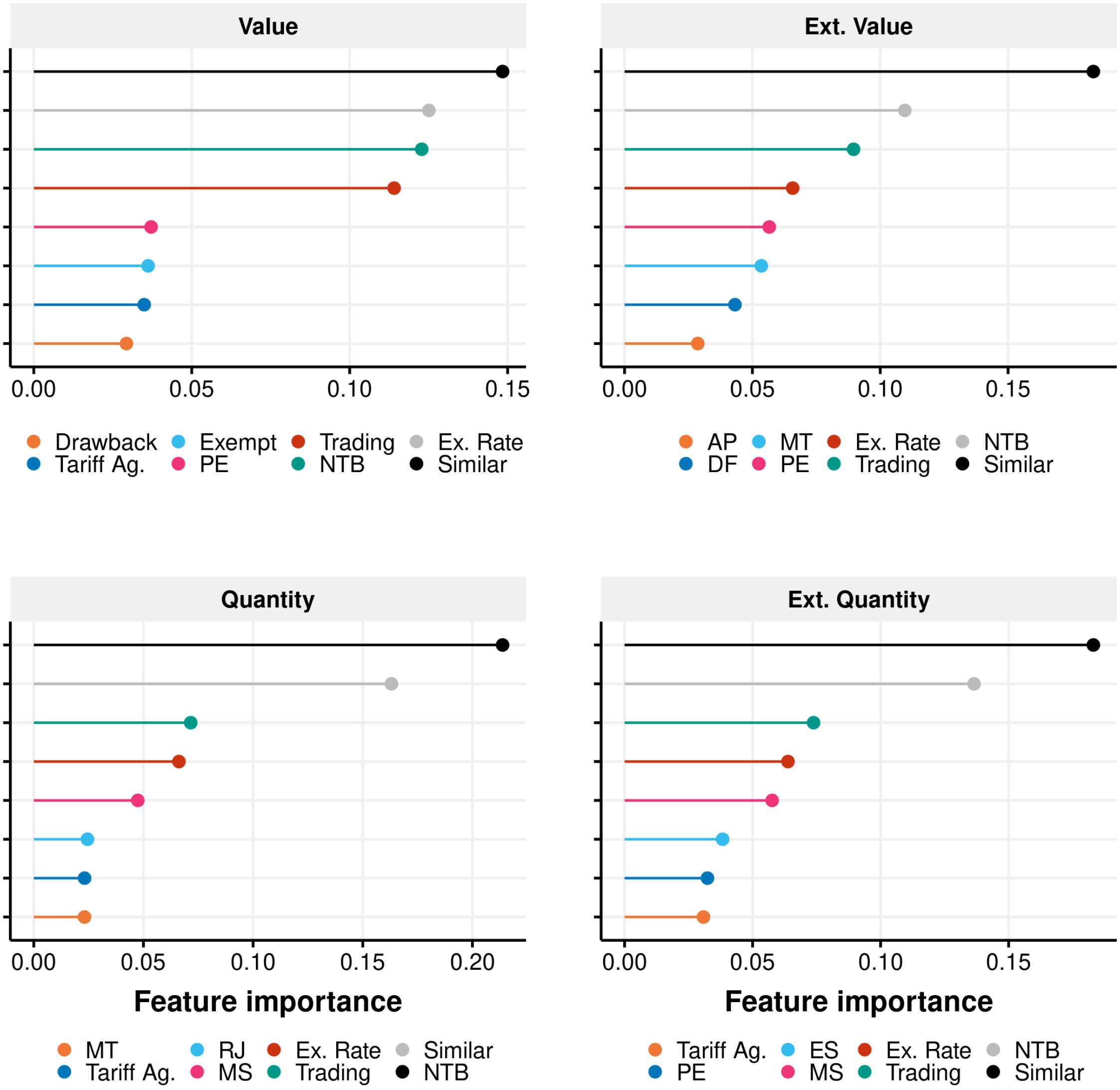

Figure D.2: Features relevant to determine heterogeneous effect

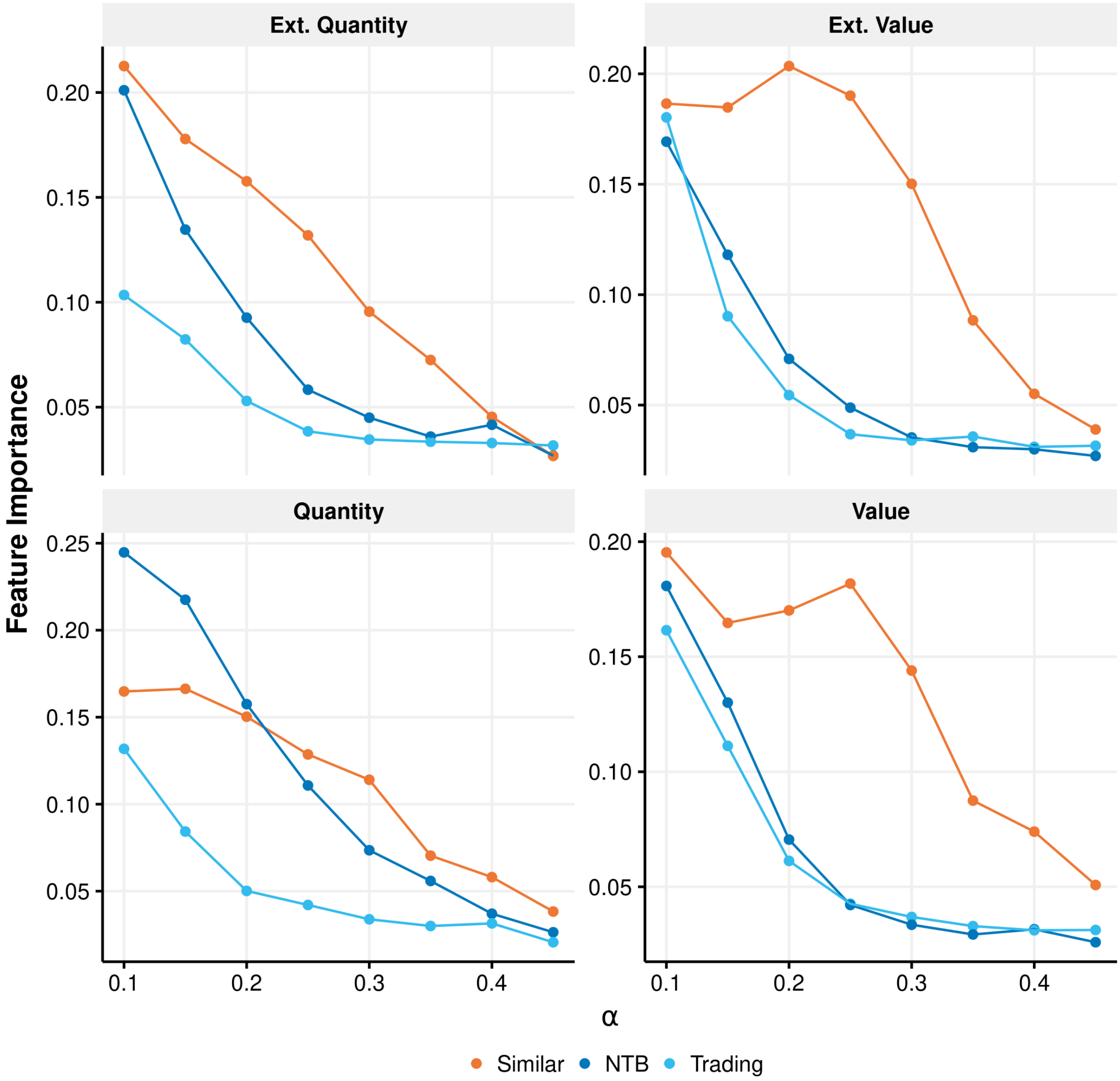

Figure D.3: Features importance for varying forest parameterization

a

E Annex

E.1 Import License information

Número: [redacted] **Data de Registro:** 23/12/2002 **Impresso em:** 10/11/2021 12:53
Situação: DEFERIDA VINCULADA À DI

SISCOMEX - Sistema Licenciamento de Importação

Extrato de Licença de Importação

Informações da LI

Licenciamento:	[redacted]
Data e Hora do Registro:	23/12/2002 [redacted]
Data e Hora da Situação:	02/01/2003 [redacted]
Situação:	DEFERIDA VINCULADA À DI
Validade da LI para Embarque:	31/01/2003
Validade da LI para Despacho:	01/05/2003
LI Substituída:	[redacted]

Básicas

Importador

Tipo do Importador:	Pessoa Jurídica
Nome do Importador:	
CNPJ:	[redacted]
Razão Social:	[redacted]
País:	
Atividade Econômica:	
Natureza Jurídica:	[redacted]
Logradouro:	[redacted]
Complemento:	
Número:	S/N
Bairro:	[redacted]
Cidade/Distrito:	[redacted]
CEP:	[redacted]
UF:	[redacted]
Telefone:	[redacted]

Outras Informações

País de Procedência:	ESTADOS UNIDOS
URF de Despacho:	AEROPORTO INTERNACIONAL DE BRASÍLIA
URF de Entrada:	AEROPORTO INTERNACIONAL DE BRASÍLIA

Informações Complementares

[redacted]

[redacted].

Fornecedor

Exportador/Fabricante/Produtor

Nome:	BEN MEADOWS COMPANY
E-Mail:	
Responsavel:	
País de Aquisição:	ESTADOS UNIDOS
País de Origem:	ESTADOS UNIDOS
Logradouro:	401 South Wright Road
Número:	005277
Complemento:	
Cidade:	Janesville
Estado:	WI

Mercadoria

Dados Gerais

NCM:	9604.00.00
Descrição da NCM:	Peneiras e crivos, manuais.
Destaque NCM:	
Unidade da Medida Estatística:	UNIDADE
NALADI/SH:	
Moeda Negociada:	DOLAR DOS EUA
INCOTERM:	CPT - CARRIAGE PAID TO

Condição da Mercadoria

Tipo da Condição da Mercadoria:	Nenhuma
Enquadramento Material Usado:	Nenhuma
Tipo de Operação:	Nenhuma

Processo Anuente

Número do Processo	Órgão Anuente
[redacted]	[redacted]

Detalhes da Mercadoria

Produto 1

Unidade Comercializada:	UNIDADE
Peso Líquido Kg:	0,00000
Qtde. na Unidade Comercializada:	1,00000
Qtde. na Medida Estatística:	0,00000
Valor do Produto no Local de Embarque:	0,0000000
Valor Unitário na Condição de Venda:	79,3300000
Valor do Produto na Condição de Venda:	79,3300000

Especificação:
PENEIRAS E CRIVOS, MANUAIS, SENDO;

PENEIRA DE 63 um (NO 230

TESTING SIEVE)

REF.: 220791

Totalizadores

Qtde.Total na Medida Estatística:	1,00000
Peso Líquido Total em Kg:	1,00000
Valor Total no Local do Embarque:	79,33
Valor Total na Condição de Venda:	79,3300000

Negociação

Modalidade Drawback:	
Acordo Tarifário:	
Acordo Aladi:	
Regime de Tributação:	ISENCAO
Fundamentação:	[redacted]
Cobertura Cambial:	COM COBERTURA CAMBIAL E PAGAMENTO FINAL A PRAZO DE ATE' 180
Modalidade de pagamento:	PAGAMENTO ANTECIPADO TOTAL OU PREPONDERANTE
Qtde. Dias Limite pagto:	
Instituição Financeira:	
Motivo:	

LI / Anuências

Informações da LI Vinculada a DI

Declaração Vinculada: [redacted]

Adição Vinculada: 002

Retificação:

Informações do Cancelamento/Vencimento da LI

Motivo:

CPF do Imp. que efetuou o cancelamento da LI:

Data do Cancelamento/Vencimento:

Hora do Cancelamento/Vencimento:

Andamento das Anuências

Anuência 1

Órgão Anuente:	[redacted]
Tratam. Administrativo:	REGIME TRIBUTARIO / FUNDAMENTO LEGAL DO REGIME
Situação:	DEFERIDA
Data da Situação:	23/12/2002
Hora da Situação:	00:00
Validade da Anuência para Embarque:	31/01/2003
Validade da Anuência para Despacho:	01/05/2003

Diagnóstico do Anuente:

#refere-se a cota 2002