No 2023-22 – November

Working Paper

CEPII

Containing Tariff Evasion*

Clément Anne, Cyril Chalendard, Ana Fernandes,
Bob Rijkers & Vincent Vicard

Highlights

- To identify transactions at risk of tariff evasion, this paper matches export transaction data from France with import transaction data from Madagascar using container identifiers.
- Reporting discrepancies between exporters and importers are prevalent but small, with over two-fifths of importers reporting in a way that increases their tariff liability.
- Aggregate tariff revenues are 24 percent lower due to discrepancies.
- Tariff revenue losses are highly concentrated: the top five evaders account for three-quarters of all tariff revenue losses and larger shipments are more at risk of evasion.
- Tariff enforcement in Madagascar is ineffective and only marginally mitigates revenue losses.

* We thank French and Madagascar Customs and GasyNet for sharing their data, feedback, and willingness to collaborate. The support of Dannick Fiononana, Antsa Rakotoarisoa, Cédric Catheline, Stéphane Manouvrier, Prisca Michea, Tolotra Hasinantenaina Ramarosandratana, and Tahina Rabesalama, was critical to the success of this project. We thank Nicolas Santos and Yewon Choi for excellent research assistance. Aaditya Mattoo and Gael Raballand provided extremely useful comments and guidance, especially in the early stages of the project. We thank seminar/conference participants at the 2017 World Customs Organization PICARD Conference, for useful comments and discussions. This paper has been partly supported by the Umbrella Facility for Trade trust fund (financed by the governments of the Netherlands, Norway, Sweden, Switzerland and the United Kingdom). We also acknowledge the generous financial support from the World Bank research support budget and the Knowledge for Change Program (KCP), a trust funded partnership in support of research and data collection on poverty reduction and sustainable development housed in the office of the Chief Economist of the World Bank. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank of Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the countries they represent. The views expressed herein are those of the authors and do not necessarily reflect the views of the International Trade Centre and the United Nations. All errors are our responsibility. The authors can be contacted at canne@worldbank.org, cchalendard@intracen.org, afernandes@worlbank.org,brijkers@worldbank.org, and vincent.vicard@cepii.fr.

Abstract

To identify transactions at risk of tariff evasion, this paper matches export transaction data from France with import transaction data from Madagascar using container identifiers. Reporting discrepancies between exporters and importers are prevalent but small, with over two-fifths of importers reporting in a way that increases their tariff liability. Yet, aggregate tariff revenues are 24 percent lower due to discrepancies. These revenue losses are highly concentrated: the top five evaders account for three-quarters of all tariff revenue losses and larger shipments are more at risk of evasion. Tariff enforcement in Madagascar is ineffective and only marginally mitigates revenue losses.

Keywords

Tax Evasion, Mirror Statistics, Trade, Corruption, Exporters, Importers, Tariffs.

JEL

F14, H26.

Working Paper

CEPII

Centre d'études prospectives
et d'informations internationales
20, avenue de Ségur
TSA 10726
75334 Paris Cedex 07

contact@cepii.fr
www.cepii.fr – @CEPII_Paris
Press contact: presse@cepii.fr

CEPII Working Paper
Contributing to research in international economics

CEPII (Centre d'Études Prospectives et d'Informations Internationales) is a French institute dedicated to producing independent, policy-oriented economic research helpful to understand the international economic environment and challenges in the areas of trade policy, competitiveness, macroeconomics, international finance and growth.

Editorial Director:
Antoine Bouët

Visual design and production:
Laure Boivin

ISSN 1293-2574

November 2023

To subscribe to
The CEPII Newsletter:
www.cepii.fr/KeepInformed

RESEARCH AND EXPERTISE
ON THE WORLD ECONOMY

1. Introduction

Tax collection is a critical development challenge. Low-income countries raise less revenue as a share of GDP than rich ones and derive a larger share of their revenues from trade taxes (Baunsgaard and Keen, 2010; Cagé and Gadenne, 2018; Besley and Persson, 2014). Paradoxically, they also suffer higher levels of evasion at the border (Jean and Mitaritonna, 2010). Enhancing tariff enforcement is thus especially important for low-income countries.

Detecting tariff evasion is difficult. It requires verifying which products are imported from what country and accurately assessing their value(s) to determine the correct tariff liability. Existing studies typically use product-level discrepancies between export values reported by the source country and import values recorded in the destination country as a proxy for evasion (following Fisman and Wei (2004)). While such discrepancies are a useful diagnostic tool, their interpretation is complicated by transport costs, which are typically included in importers' records but not in exporters' reports. Moreover, such discrepancies offer only limited help in identifying which transactions are most at risk of being fraudulent. This in turn undermines their usefulness for pinpointing perpetrators of tariff evasion, examining which evasion methods are most important, and assessing the effectiveness of customs administrations in curbing fraud. These questions, which are central to the design of remedial policies, are addressed in this paper.

To identify which transactions are most at risk of tariff evasion we match export transaction data from a high-income country, France, with import transaction data from a low-income importing country, Madagascar, using identifiers for the containers in which goods are transported. To the best of our knowledge, the resulting dataset is the first of its kind, and permits novel and uniquely rich container-specific measurement of tariff evasion net of discrepancies due to freight and insurance costs (designated in what follows as 'transport costs' for which we have accurate

measures). Direct comparison of the value and products declared in exporters' and importers' customs declarations allows us to quantify the importance of undervaluation of the declared value and misclassification to less taxed products.[2] Information on the customs duties paid by the importer makes it possible to assess the loss of tariff revenue due to tariff evasion.[3] Finally, information on revisions to value and tax liability made by inspectors during the import customs clearance process allows us to assess the efficiency of customs in curbing tariff evasion.

Containerized trade between France and Madagascar provides an ideal context to examine tariff evasion. Taxes levied at the border account for 48 percent of Madagascar's government total tax revenue in spite of rampant corruption and widespread tariff evasion (Chalendard, Duhaut, Fernandes, Mattoo, Raballand, and Rijkers, 2020; Chalendard, Fernandes, Raballand, and Rijkers, 2023). France is the second most important source country for Malagasy imports (after China) and is known to have strong tax enforcement capacity.[4] Its high-quality customs transaction data have been extensively used in the trade literature (e.g. Eaton, Kortum, and Kramarz (2004); Berman, Martin, and Mayer (2012); Mayer, Melitz, and Ottaviano (2014)) because of their reliability. We focus on containers because containerized trade is the dominant form of maritime trade, accounting for 46.7% of global trade in 2020 (Coşar and Demir, 2018) and because container identifiers enable us to unambiguously match the reports of exporters and importers on the same transaction.

Our main findings can be summarized as follows. First, discrepancies between exports and imports recorded - or trade value gaps (Fisman and Wei, 2004) - are small on average but highly prevalent and only in part motivated by a desire to evade tariffs. For the average container, the

[2] By definition, our focus on matched exporter-importer transactions from France to Madagascar prevents us from investigating misreporting of country of origin.
[3] Firm-level information additionally permits the identification of firms benefiting from tariff-free imports (free zone enterprises, FZE), which have no incentive to misreport for tariff evasion purposes.
[4] France has the 2nd highest tax to GDP ratio (45.4%) of all OECD countries.

export value recorded in France exceeds the import value recorded in Madagascar by 6 percent. But over two-fifths of containers exhibit gaps that increase their tariff liability relative to what would have been paid had importers reported the values in the same way they were registered by French exporters, suggesting the firms importing them are not operating strategically to minimize their tax burden. The same holds for product misclassification: for only half of all containers is the same set of products recorded in both France and Madagascar, but only a subset of them (19 percent) features misclassification towards less taxed products. In fact, most misclassifications are tariff rate-neutral.

Second, reporting discrepancies nonetheless correlate with tariff rates, tariff evasion risk scores, and proxies for corruption risk (Chalendard, Fernandes, Raballand, and Rijkers, 2023), consistent with the premise that those discrepancies are informative about evasion of import tariffs. Value gaps and the prevalence of tariff-decreasing product misclassification both increase especially rapidly with tariffs for larger shipments, for which more tariff revenue is at stake.

Third, the aggregate costs of evasion are high. If imports in Madagascar were reported in the same way they were recorded in France, tariff revenue would be 24 percent higher relative to the status quo. A decomposition exercise reveals that misclassification accounts for roughly one-third of all tariff revenue losses associated with misreporting, while undervaluation accounts for 67 percent. These findings align with evidence from a nationwide survey of customs inspectors in Madagascar, according to whom undervaluation is the dominant form of evasion.

Fourth, evasion is highly concentrated among importers. The top 5 importers in terms of tariff revenue losses together account for three-quarters of aggregate tariff revenue losses associated with reporting discrepancies. Tax revenue losses are even more concentrated than trade values, which are notoriously concentrated (Freund and Pierola, 2015). Interestingly, we find that top

evaders are not among top importers.

Fifth and related, the identities of the importer and exporter firms are the best predictors of reporting discrepancies and dominate the predictive power of product characteristics. Unregistered importers are more likely to engage in evasion, whereas state-owned enterprises (SOE) which may face softer budget constraints, are less likely to. By contrast, the identity of inspectors and the customs office have limited explanatory power.

Last but not least, customs inspections in Madagascar appear poorly targeted and ineffective. Value adjustments by the customs reduce tariff losses by 1.8 percent and are only weakly correlated with value gaps. This may be because evasion is difficult to predict but could also be consistent with corruption. These findings dovetail with prior literature showing that taxation is harder to implement in developing countries (Gordon and Li, 2009).

This paper contributes to and builds on several strands of literature. To start with, we complement the literature on tax and tariff evasion (Andreoni, Erard, and Feinstein, 1998; Fisman and Wei, 2004; Slemrod, 2019; Sequeira and Djankov, 2014; Sequeira, 2016) by quantifying the extent of tariff evasion, documenting its channels, and highlighting the granularity of evasion and the predictive power of shipment characteristics. A handful of importers account for the bulk of aggregate tariff revenue losses and the identity of importing and exporting firms is the best predictor of reporting discrepancies. Evasion is concentrated in high value shipments. The granularity of evasion suggests that altering the behavior of a handful top evaders could have sizable macroeconomic effects, much in the same way that the volatility of aggregate output and trade has been shown to be driven by the performance of a small set of large firms (Gabaix, 2011; Di Giovanni and Levchenko, 2012; Di Giovanni, Levchenko, and Mejean, 2014).

Second, misreporting is rife and a substantial share of firms are not acting as "strategic opti-

mizers" but rather make costly mistakes by reporting in a way that increases their tax liability, thus contravening assumptions of standard models of firm responses to taxation (Allingham and Sandmo, 1972; Kleven, Knudsen, Kreiner, Pedersen, and Saez, 2011). These findings complement the work of Almunia, Hjort, Knebelmann, and Tian (2022) who use value-added tax (VAT) transaction data from Uganda to show that a non-negligible proportion of firms (about a quarter) consistently make costly reporting errors, even though the majority of firms underreport their tax liability as conventional models assume.[5]

Third, our findings have implications for the measurement of illicit financial flows and product-level mirror statistics in particular. Product-level mirror statistics studies show such discrepancies varying with tariff rates and other taxes (Ferrantino, Liu, and Wang, 2012) and their enforcement (Mishra, Subramanian, and Topalova, 2008), product characteristics (Javorcik and Narciso, 2008), non-tariff measures (Kee and Nicita, 2022), and country characteristics such as institutional quality or the level of corruption (Carrère and Grigoriou, 2015; Jean, Mitaritonna, Vatan, et al., 2018; Kellenberg and Levinson, 2019). While discrepancies in mirror statistics are often interpreted as evidence of evasion, we show that caution is warranted because a substantial share of reporting discrepancies are hard to reconcile with strategic optimization and, moreover, could arise from recording differences associated with the contractual terms under which trade takes place. Failing to appropriately account for transport and insurance costs would cause one to underestimate trade value gaps by 12.7 percent for the average shipment.

The remainder of the paper is organized as follows; the next section describes the data and the container matching procedure. Section 3 presents the measures of reporting discrepancies. Section 4 characterizes reporting discrepancies and documents how concentrated evasion is

[5]Methodologically, our paper is also linked to a recent set of papers that use VAT transaction data to examine evasion including Carrillo, Pomeranz, and Singhal (2017) and Naritomi (2019).

among importers. Section 5 estimates the determinants of trade value gaps, misclassification, and transport costs. Section 6 examines the efficacy of Madagascan customs in addressing evasion. A final section concludes.

2. Data and measurement of tariff evasion

Our analysis uses French customs export declarations data matched with Madagascan customs import declarations data at the container level over the 2014-2016 period. This section briefly describes the French and Madagascan customs databases, the matching process, and the measures of export and import values and tariffs. It is useful to first define three relevant concepts: customs declaration, item, and freight/shipping container. A customs declaration is an official document which records all details of an import or an export transaction, covering information on the goods traded, the parties (generally firms) involved, as well as the customs regime(s) used. A customs declaration can contain different products, referred to as items. For each item, the relevant HS 6-digit (HS6) product code, value and weight are recorded. Key customs variables such as values or weights are available at the declaration-item level. A shipping container, hereafter referred to as 'container', is a standardized steel box used to transport goods across locations or countries, with a unique container identifier. Our main unit of analysis is the container, which may include several HS6 products exported by one or more French firms to a Madagascan importer at a specific date of arrival.

2.1. Data

We use customs data on the universe of French export transactions to Madagascar over the period 2013-2016, complemented by additional information from the French Statistical Database on External Trade on container identifiers (Foreign Trade Statistics of the General Directorate of Customs and Indirect Excises). We combine this with data provided by the Madagascan

customs administration on import declarations from France (labeled as country of provenance or origin) over the period 2014-2016. Unique features of the Madagascan data include: (i) declaration-item level measures of detailed transportation costs that we use to convert reported import Cost Freight and Insurance (CIF) values to import Free on Board (FOB) values as described in Section 2.3; (ii) tariffs paid by declaration-item; information on the identity of both (iii) the importer, (iv) the inspector assessing the declaration and the broker submitting it on behalf of the importer. For each declaration, the customs database is complemented by data from GasyNet, a private-public-partnership that assists Madagascan customs with risk analysis, on a proprietary risk score, an inspection clearance channel recommendation, and the presence of a scan (see Chalendard, Fernandes, Raballand, and Rijkers (2023) for details). Details on these data are provided in Appendix A1.2.

2.2. Matching export and import declarations using container IDs

The crucial feature of both customs databases is the presence of container identifiers (hereafter 'container IDs') which are used to match them. As common international practice, container IDs follow an ISO 6346:1995 norm (hereafter 'ISO 6346') defined by the International Organization for Standardization (ISO) and managed by the Bureau of International Containers. Each container ID is unique and ISO 6346 is recognized by French and Madagascan customs in the process of identifying containers.[6] Incentives to misreport container IDs are very limited since they are included in the bill of lading and the shipping manifest used by shipping lines. Any modification to the container ID could complicate its departure from (and/or processing in) the (arrival) port.

Our study is to our knowledge the first to implement an accurate matching between customs

[6]Container IDs following ISO 6346 consist of 3 string characters (owner/operator code), a letter (equipment identifier), 6 digits (serial number), and a 7^{th} digit (check digit computed thanks to the previous characters), as is shown in Appendix Figure A1.

declarations recorded in different countries at the container level instead of comparing yearly mirror trade statistics aggregated at the HS6 product level. Using the container as the unit of analysis enables us to directly compare values or weights to construct trade gaps at a more finely disaggregated level than in previous studies and to compare recorded products in the export and import declarations. It thus enables us to distinguish product misclassification from undervaluation (for similarly reported products).[7]

Two limitations of the use of container IDs for the matching between French and Madagascan declarations are that (i) including a container ID is not mandatory when lodging the customs declaration which reduces the sample size and (ii) the analysis is restricted to containerized trade which has specific features including higher quantities and lower average unit prices (1.93 euros(€) per kilogram (kg) for containerized trade versus 5.27€/kg for overall trade as shown in Appendix Table A1).[8]

The procedure to match the customs databases from France and Madagascar comprises three steps discussed in Appendix A1.3: cleaning procedures to ensure compliance with ISO 6346 of container identification, aggregation of each customs database at the container-product-day level, then matching based on container ID provided the time lag between customs registration dates in France and Madagascar does not exceed 90 days to avoid incorrect matches due to the use of the same container for a different voyage.[9]

[7]We did not use exporter/importer names to match across French and Madagascan customs data because a given French declaration may include shipments from multiple exporters/importers, while Madagascan declarations include shipments from a single importer. Moreover, exporter/importer names recorded can differ from the effective exporter/importer (HQ/subsidiary names, acronyms, intermediaries, care-of. . .). A study that has used such names (and addresses) to match across United States import shipment records and Colombian export shipment records by Krizan, Tybout, Wang, and Zhao (2020) shows that for a given importing firm-exporting firm pair the match rate for specific transactions is low (less than 30 percent of transactions in both countries).

[8]Similar evidence of a higher unit price per weight for trade transported by air relative to lower average unit prices for containerized trade that travels mostly by sea is shown by Hummels (2007) for the United States.

[9]The gap between registration dates should account for travel time and potential transit times at each customs office. The estimated travel time between Metropolitan France and Madagascar is 21 days and 12 hours (based on searates.com as detailed in Appendix A1.3). However, the distribution of the registration day gap in our data

Appendix Table A1 keeps track of sample representativeness at various matching stages, taking French export data as the baseline (Sample 0). French declarations featuring a container ID represent 76 percent of the total export weight and 28 percent of the total export value (the smaller value share relative to the weight share is due to containerized products featuring lower unit prices) (Sample 1). We remove the 2 percent of containers whose IDs do not comply with ISO 6346 (Sample 2). Our matching of French and Madagascan customs databases at container-day level reduces the number of container IDs covered by 8.5% (Sample 3). Next, we drop observations for which multiple Madagascan declarations are linked to a unique container ID as inspector assignment, risk scores, and inspection channel recommendations are made at the declaration, not the container level.[10] We keep in the final sample observations for multiple French declarations linked to a unique container ID. This step leads to a reduction in the number of container IDs of 16.3% of the overall total (Sample 4) but because these are mostly small declarations, the reduction in export value is proportionately smaller. Limiting the time that passes between the dates of registration in France and in Madagascar to 90 days induces a very small additional reduction in the number of container IDs (Sample 5).[11] To limit excessive swings in value gaps originating from small value transactions we remove containers whose export or import FOB values are below 1,000€ and Madagascan declarations benefitting from tax exemptions (Sample 6). Since a key objective of this paper is to assess the relationship between trade gaps and tariffs, we also remove containers for which hypothetical tariffs defined in Section 2.4 cannot be computed (Sample 7). Our matching procedure results in a final sample of 7,461 container observations corresponding to 602 importers. Each container observation is

(in Appendix Figure A2) is concentrated around a median of 41 days since it also includes the transit time at each customs office and other transit time. The selection of 90 days enables the matching of relevant observations that should concern the same shipment while excluding only a small share of observations (3 percent) that likely concern unrelated shipments that use the same container for a different voyage.

[10] Keeping multiple declarations per container would thus create identification and inference challenges.

[11] The final sample includes only observations for which all containers listed in a Madagascan or French declaration have been matched successfully.

matched to a unique Madagascan declaration. The sample covers 55 percent of export value and 68 percent of export weight of declarations with non-missing container IDs.[12]

2.3. Making export and import values comparable

Our main variable of interest, the trade value gap defined in Section 4.1, compares French export values to Madagascan import values at the container level. Since our analysis focuses on the magnitudes and the role of tariff evasion in generating such gaps, it is essential to use comparable value measures across both countries so as to exclude gaps due to measurement error. Export and import flows are usually reported by customs agencies using different valuation methods: exports are reported Free on Board (FOB), i.e., excluding international transportation costs, while imports are reported on a Cost, Insurance and Freight (CIF) basis, i.e., including transportation costs up to the importing country's border. To accurately measure reporting discrepancies, we need to compare export FOB value to import FOB value.

A key feature of our Madagascan customs data is its inclusion of detailed transport costs – external freight costs, insurance costs, and other charges – which enable us to compute import FOB value at the container level. The import FOB value for container c is obtained as the sum, over all HS6 products m declared in container c, of the initial import CIF value[13] in Ariary $M^{cif}_{c,m}(Ar)$ minus the sum of transport costs in Ariary $TrCosts_{c,m}(Ar)$, then converted into euros.[14] The resulting import FOB value at the container level is comparable to the statistical

[12]Values and weights are recorded at the declaration-item level. If a declaration is linked to multiple containers, we obtain measures of values and weights at container level by dividing for each item its total values or weights by the number of containers.

[13]The initial import value is the value entered in the customs registration system by the broker in charge of the declaration, before the inspector examines the declaration and assesses whether value adjustments are required. Taxes and duties are paid based on the final value retained by customs, which is potentially adjusted by the inspector.

[14]We use as exchange rate the average between the monthly €/Ar exchange rates at the time of the French declaration registration er_x and the time of the Madagascan declaration registration er_m, from the Madagascan Central Bank. The import FOB value in euros $M^{fob}_{c,m}$ is given by: $M^{fob}_{c,m} = \frac{M\_f^{cif}_{c,m}(Ar) - TrCosts_{c,m}(Ar)}{0.5(er_m + er_x) * ValueAdjRatio_{c,m}}$. Transport costs recorded by customs are based on final import value $M\_f^{cif}_{c,m}(Ar)$, so we adjust the numerator by a final-to-initial value ratio $ValueAdjRatio_{c,m}$. Monthly rather than daily exchange rates are chosen to limit excessive

export FOB value provided by French customs since we remove transport costs. This improves upon past studies analyzing value discrepancies in international trade statistics, which typically have not been able to make this correction. Transport costs account for 12.7% of the import CIF value on average, as is shown in Table 1, and, moreover, vary significantly across HS6 products, as is shown in Appendix Figure A3. Correcting our measures of reporting discrepancies for their presence is thus important.

2.4. Tariff measures

Our analysis uses three tariff-related measures. First, we define a *hypothetical tariff rate* at the container level ($HypTariffRate_c$) as the weighted average tariff rate that should have applied as per the French export declaration, i.e., using HS6 product codes reported in the French customs declaration for container c. To account for preferential tariffs and their utilization, tariff rates by HS6 product-year are computed as the median of the tariff rate paid by HS6 product-year over all import declarations recorded in Madagascar with French origin or provenance (including those not in our matched sample) with the exception of declarations imported using special procedures (transshipment) or subject to tax exemptions.[15] $HypTariffRate_c$ is a weighted average of HS6 product-level tariff rates using as weights the French declared products' shares in total FOB exports of the container.[16] The variable is set to zero for Free Zone Enterprise (FZE) importers that benefit from multiple fiscal advantages among which tariff-free imports.

Second, we define a tariff revenue gap variable (TRG_c) to measure the difference between hypothetical tariff liability ($HypTariffLiability_c$) and paid tariffs on import FOB value ($PaidTariffs_c$)

day-to-day volatility. Import FOB value at container level is the sum of $M_{c,m}^{fob}$ across all N_m products in Madagascan declarations.

[15]Note that the Madagascar tariff schedule features predominantly tariff rates at 0, 5, 10, 15 and 20 percent, the highest tariff at 20 percent being applied to 20 percent of tariff lines in 2016.

[16]The formula used is $HypTariffRate_c = \sum_{f=1}^{N_f} shx_{c,f} * TariffRates_f$ where $shx_{c,f}$ is product f's share in total FOB exports of the container. $TariffRates_f$ are measured at the HS6-year level but the year subscript is excluded for simplicity.

at the container level:

$$TRG_c = HypTariffLiability_c - PaidTariffs_c \quad . \quad (1)$$

Both $HypTariffLiab_c$ and $PaidTariffs_c$ are computed taking into consideration which HS6 products are listed in both the French and the Madagascan declarations for the same container.[17] For matched products in a container, that is products declared both in France and Madagascar, we rely on the applied tariff rate by HS6 product (from the importer declaration) multiplied by the export or import FOB values, respectively. [18] For unmatched products in a container, declared only in France or only in Madagascar, we use the tariff rates at HS6-year level defined above multiplied by the export or import FOB values, respectively. Then we sum across all products in a container to obtain the hypothetical tariff liability and the estimated paid tariff at the container level, while setting both to zero for FZE importers.[19]

Third, to decompose the role of different channels of tariff evasion, we construct two components of the tariff revenue gap as follows:

$$TRG_c^{underval} = \left[X_c^{fob} - M_c^{fob}\right] * HypTariffRate_c \quad (2)$$

and

$$TRG_c^{misclass} = TRG_c - TRG_c^{underval} \quad . \quad (3)$$

[17] Paid tariffs rely on the final import CIF value after customs potentially makes tax adjustments. To obtain a number that is comparable to the hypothetical tariff liability, we use import FOB value on which we apply tariff rates to assess the estimated paid tariffs on FOB value.

[18] Relying on applied tariffs for matched products allows us to take into account the application of preferential tariff rates (in the instances when that happens).

[19] We also obtain a measure of the tariff revenue gap scaled by the average between the hypothetical tariff liability and the estimated paid tariff, setting the value at 0 percent in case both aggregates are zero. This measure follows Haltiwanger, Jarmin, and Miranda (2013) and is bounded between -2 and 2.

The undervaluation component ($TRG_c^{underval}$) is the container-level difference between declared export and import FOB values multiplied by the container-level hypothetical tariff rate. This component measures how much tariff revenues would have been altered without any changes to the product mix included in the container – more specifically it holds constant not only the composition but also the value shares of all products included in the container – thus effectively asking how much lower (or higher) would tariff revenues be if values of all imported products included in the container changed proportionately. The misclassification component ($TRG_c^{misclass}$) is the residual between the tariff revenue gap and the container-level undervaluation component.

3. Trade value gaps, product misclassification, and their determinants

Following Fisman and Wei (2004), the literature on tariff evasion has focused on trade value gaps defined as the log difference between the export FOB value declared by the exporting country and the mirror import CIF value declared by the importing country at the product level. Such trade value gaps reflect the net total effect of undervaluation of the import FOB value, product misclassification, and the lack of proper accounting for the presence of transport costs (in the import value). Disentangling the relative importance of each of these channels is not feasible with product-level data.[20] Our container-level matched data and our access to transport costs data allow us to i) compute value gap measures that are purged from transport costs, ii) directly measure each potential channel of tariff evasion (undervaluation and misclassification), and iii) assess their determinants using characteristics of the importing and exporting firms, the shipment, and contractual arrangements. Below we describe our main outcome variables and potential determinants.

[20] The literature has provided indirect evidence for product misclassification using the tariff rate of similar products and comparing gaps in value versus gaps in weights (Fisman and Wei, 2004). While informative qualitatively, such approach does not allow quantifying the importance of different channels of tariff evasion.

3.1. Trade value gaps and product misclassification

We focus on outcome variables that measure different dimensions of tariff evasion: the under-valuation of the customs value declared to the customs agency in the importing country and (different types of) product misclassification First, we define a trade value gap at the container level (VG_c) as the log difference between the export FOB value X_c^{fob} and the constructed import FOB value:[21]

$$VG_c = \ln\left[X_c^{fob}\right] - \ln\left[M_c^{fob}\right] \ . \quad (4)$$

VG_c measures pure misreporting of the shipment value at the border. By comparing exports and imports on an FOB basis, it does not include transport costs that differ importantly across products (see Section 2.3). Hence in our analysis the differential recording of transport costs is muted as a channel for tariff evasion. Formulas similar to that in Equation (4) are used to define weight gaps (WG) or unit price gaps (UPG) at the container level using, respectively, weights or unit prices instead of values, with unit prices being measured by the value divided by the weight.

Note that the interpretation of the value gap defined at the container level VG_c differs from that of a value gap defined at the product level as in Fisman and Wei (2004). In the latter case, any product misclassification increases the value gap. On the contrary, at the container level, there can be product misclassification but the value gap would be nonetheless zero if the total value in the export and import declarations in the matched container is the same.

Value gaps are a popular proxy for evasion because importers have an incentive to underreport their imports which exporters do not. Yet, it is important to bear in mind some limitations of

[21] The main value gap measure compares initial values in Madagascar to export values in France. Table 1 also provides summary statistics for value gap measures based on final values in Madagascar or defined without any logarithms.

this measure. First, value gaps do not detect smuggling (where neither the importer nor the exporter reports the transaction). Second, value gaps do not detect collusive evasion; if the importer and exporter both underreport, then the value gap may be zero even in the presence of substantial evasion. Third, reporting discrepancies may arise for a multitude of reasons other than evasion of import tariffs. In addition to resulting from genuine measurement error, they may also stem from deliberate misreporting of exports and imports to minimize corporate profit tax and VAT payments (Bussy, 2023), maximize VAT rebates, or to circumvent the application of non-tariff barriers (Kee and Nicita, 2022), capital or exchange rate controls (Ferrantino, Liu, and Wang, 2012). Finally, the use of minimal/reference prices may induce importers that know they cannot declare unit prices below a certain minimum price threshold to inflate import values, generating negative value gaps. In our context, capital controls and exchange rate controls are not relevant but we will explore the predictive power of non-tariff barriers and minimum values in robustness tests. In addition, we assess how predictive tariffs are of value gaps to understand the importance of evasion as a driver of reporting discrepancies relative to these other explanations.

Second, our matched data enables the identification of product misclassification at the container level. We first define a product misclassification indicator for containers whose set of HS6 products in the Madagascan declaration does not match exactly the set in the French export declarations.[22] Given our emphasis on tariff evasion, we focus on misclassification from high-tariff to low-tariff products. Our main variable of interest is thus an indicator for containers with tariff-decreasing misclassification when products are misclassified and the misclassification component of the tariff revenue gap ($TRG_c^{underval}$ defined in Section 2.4) is positive and

[22]The 6-digit level is the highest level of product disaggregation harmonized at the international level on which we can therefore match products from French and Madagascan declarations. Tariffs are however defined at the 8-digit level in Madagascar, so that unmeasured product misclassification could occur within HS6 categories. Only a subset of HS6 categories are however concerned: in 2014, 4472 out of 5203 HS6 categories had only one 8-digit product; and of the remaining 731 HS6 categories featuring multiple 8-digit products, 609 had 8-digit products all having similar tariffs, so that only 122 HS6 had variance in tariff rates across HS8 products within the HS6.

above 1€. Tariff-decreasing misclassification can be viewed as strategic product misclassification aiming at reducing the tariff liability on the import shipment. We also consider an indicator for containers with non-strategic misclassification when products are misclassified and the misclassification is tariff-neutral or tariff-increasing. These cases are referred to as non-strategic misclassification.[23]

3.2. Determinants of value gaps and product misclassification

We define here three sets of potential determinants of value gaps and product misclassification: importing and exporting firm characteristics, shipment characteristics, and contractual terms.[24] The importer and exporter characteristics are the following: (i) an unregistered importer indicator identifies the importers with no tax identifier in the Madagascan customs data; (ii) an SOE indicator identifies importers with some government participation (based on a list issued by the Treasury of Madagascar in 2016); (iii) an FZE indicator identifies importers located in free zones and benefitting from tax advantages among which tariff-free imports (the categorization of such firms is based on information on customs regimes); (iv) a trusted trader time-varying indicator identifies importers that are members of a trade facilitation program called "Procédure Accélérée de Dédouanement" (henceforth PAD) and subject to minimal inspections and reduced clearance time (based on a list provided by Madagascar customs); and (v) a non-commercial exporter indicator identifies exporters with no EORI (Economic Operator Registration and Identification) identifier number in France.[25]

The shipment characteristics are as follows: (i) a multiproduct shipment indicator for containers

[23] Tariff-neutral misclassification involves an absolute misclassification component of the tariff revenue gap within 1€ while tariff-increasing misclassification involves a misclassification component below -1€.
[24] Summary statistics for these variables are presented in Appendix Table A5.
[25] In robustness checks, we use other importer characteristics: size given by importer total import value across our sample, global value chain participation given by the record of at least one Madagascan export declaration for the importer, an indicator for new importers defined as those who start to import in the year of the matched container and were not importing in the previous 7 years, and an indicator for young importers defined as those that started to import in any of the two years prior to the year of the matched container.

declaring multiple HS6 products in France; (ii) a larger export value indicator for containers with export value above the median (16,794€); and (iii) risk characteristics including the risk score (ranging from 1 to 9), an indicator for the container being physically inspected (red channel) and another for being scanned.

The contractual terms relate to the possibility that customs agencies do not fully correct the declared invoice values to generate the customs statistical values which are used in our value gap measures. We consider two groups of incoterms (contractual terms) depending on the inclusion of main transport costs (sea freight) in the invoice: the incoterm group After shipment incoterms (ASI) (including such costs) and the incoterm group Before shipment incoterms (BSI) (excluding such costs). Then we define three indicator variables, one for containers whose incoterms are similar across France and Madagascar, one for containers BSI in France but ASI in Madagascar, and one for containers ASI in France but BSI in Madagascar.[26]

4. Characterizing reporting discrepancies

4.1. Value gaps and product misclassification

Descriptive statistics in Table 1 reveal that discrepancies in recorded import and export FOB values are highly prevalent but small on average. The average initial value gap is 1,692€ per container for an average export FOB value per container of 24,708€. Value gaps average 6.3 percent with a large standard deviation of 44.8 percent. If we had not accounted for transport costs, averaging 1.975€ per container, or 12.7 percent of import CIF value, we would have erroneously concluded that value gaps were on average negative, as the average CIF value of imports is 24,986€ (whereas the FOB value is 23,016€). Accounting for transport costs is thus crucial for accurate measurement of value gaps.

[26] The details on these incoterm definitions are provided in Appendix A1.2.

These fairly low average value gaps mask a high prevalence of discrepancies shown in Figure 1, which depicts the distribution of value, weight, unit price, and tariff revenue gaps across containers, as well as Appendix Table A2 which presents additional statistics. While just over half (53 percent) of all containers feature value gaps within a 5 percent absolute range, reporting discrepancies are widespread, with 21 percent of containers having a negative gap under 5 percent and the remaining 26 percent having a positive gap exceeding 5 percent. The distribution of value gaps is right-skewed (Figure 1A) and so is that of tariff revenue gaps (Figure 1D). Value gaps are predominantly driven by misreporting of unit prices (Figure 1B) as weight gaps are much smaller in absolute terms than unit price gaps (Figure 1C), presumably because underreporting of quantities is much easier to detect than underreporting of values.

Table 1 shows that customs inspections reduce the average value gap by 66 euros only, in part because only 3.7 percent of declarations are subject to value adjustments.[27] When made, adjustments represent on average 22.5 percent of the initial import value. Their low prevalence limits their effectiveness in curbing evasion, as will also be shown in Section 6.

Product misclassification is also pervasive (see Table 1 and Appendix Tables A2 and A3). Containers with correct product classification account for half of all containers and 39 percent of total imports. However, most instances of misclassification do not appear strategic, as they do not result in a lower tariff burden: only 19 percent of containers feature tariff-decreasing misclassification while 31 percent feature neutral or tariff-increasing misclassification.

An interesting aspect of misclassification is that Madagascan importers tend to report many more distinct HS6 products than French exporters for a given container, as is shown in Appendix Figure A4.[28] On average importers in Madagascar report 6 more products than French exporters.

[27] Customs inspections may result in the final import CIF/FOB value being larger than the initial CIF/FOB value.
[28] Note that 60 percent of containers declare the same number of distinct HS6 product codes on the export and the import side (even if the HS6 codes themselves may differ).

This pattern is not unique to our matched sample of containers. Based on a global sample of 156 countries reporting trade statistics both as importer and exporter to UN COMTRADE in 2014-2016, we compute the percentage of exporter country-importer country-year observations with more HS6 products reported by the importer than the exporter country. Around 70 percent of observations are characterized by more HS6 products reported by the importing than the exporting country (see Appendix Table A4).[29]

The high prevalence of negative tariff revenue gaps in Figure 1D suggests that a substantial share of the discrepancies is not driven by a desire to minimize tariff revenue but is consistent with innocuous misreporting in the importing country: for these transactions importers' total tariff liability increases relative to the amount they would have had to pay if they had reported in the same way as their matched French exporter. Consistent with this explanation, value gaps are much more prevalent for containers for which incoterms conflict (see Appendix Figure A5).[30]

4.2. Are reporting discrepancies driven by tariff evasion?

The high prevalence of discrepancies, and in particular the fact that many of them are associated with over-valuation begs the question to what extent reporting discrepancies reflect evasion of import tariffs. This section presents several stylized facts suggesting evasion is an important determinant of the documented discrepancies. As already discussed in Section 3.1, reporting discrepancies may arise for several reasons, including reporting errors, evasion of NTMs, minimum values, and/or minimization of VAT and corporate profit tax payments. While all these sources of discrepancies may generate value gaps and product misclassification, they should a priori be orthogonal to tariff rates and should not systematically induce a positive average tariff

[29] Considering trade between all partners and Madagascar that prevalence increases to 80-88 percent.
[30] Among consistent incoterms, containers featuring an ASI have higher average value gaps than containers featuring a BSI. ASI feature main transport costs in the invoice value, so this could be related to inconsistent removal of those transport costs in the export FOB value, driving upwards the value gap.

revenue gap as shown in Table 1.

Moreover, value gaps are systematically correlated with tariffs, as is shown in panel A of Figure 2, which plots these value gaps against tariffs separately for containers with smaller versus larger export value. Incentives to evade are stronger for the latter group, because more money is at stake. Consistent with this conjecture, we observe a strong monotonic positive association between tariff rates and value gaps for larger export value shipments. By contrast, for containers with a smaller export value, value gaps are negative at low levels of tariffs and exhibit a roughly U-shaped relationship with tariffs, being only marginally positive for containers subject to the highest tariff rates. Value gaps are thus largest for containers with most at stake, i.e., subject to higher tariffs and for which the export value is larger. The same pattern holds for product misclassification resulting in reduced tariff payments: panel B of Figure 3 shows that the propensity to misclassify products towards lesser taxed products is correlated with tariffs and this is especially the case for containers with larger export values, suggesting that misclassification is at least in part motivated by a desire to evade tariffs.

Value gaps are also systematically positively correlated with risk scores as shown in Figure 4a. The risk scores are given to each import declaration by the third party GasyNet based on a proprietary risk model that relates the characteristics of the declaration to the probability of non-compliance with customs regulations (including the payment of adequate tariffs). As an additional test of the conjecture that value gaps are informative about evasion risk, we consider a proxy for corruption risk proposed by Chalendard, Fernandes, Raballand, and Rijkers (2023) based on excess interaction between the customs inspector assessing the import declaration in a container and the Madagascan broker who registers it indicating potential manipulation of the inspector assignment.[31] Figure 4b shows that the positive correlation between value gaps and

[31] This proxy for corruption was validated by subsequent audits implemented by Madagascar customs that resulted

tariffs is stronger for containers that are suspect of being the object of corruption according to that proxy.

In short, value gaps are positively correlated with tariffs and corruption risk proxies, suggesting evasion is at play. At the same time, the high prevalence of discrepancies that increase the tariff burden relative to what importers pay serves as a reminder that gaps are by no means exclusively driven by evasion.

4.3. Tariff revenue losses

Our data allow the quantification of the impact of reporting discrepancies on aggregate tariff revenue. Small median tariff revenue gaps per container – 0€ as per Table 1 – mask large aggregate tariff revenue losses, because value gaps and product misclassification are concentrated in high tariff, large export value containers. The average hypothetical tariff liability per container is 2,266€ but the paid tariffs on import FOB value per container are only 1,828€. The average tariff revenue gap (i.e., the amount that is plausibly evaded) is thus 438€ per container. Put differently, the discrepancy between reported imports and exports for a given container induces a tariff revenue loss representing 24 percent of average paid tariffs. The difference between median and aggregate numbers attests to the granularity of tariff evasion, a theme to which we will return below.

We further disentangle in Table 2 the undervaluation effect from the misclassification effect on lost tariff revenues using the two components of the tariff revenue gaps described in Section 2. Containers with correctly classified products only account for a limited share of total tariff revenue losses. Containers with tariff-decreasing product misclassification are the main source of lost tariff revenues, and account for 82 percent of lost revenues (see Table 2 and Appendix Table

in the sanctioning of several inspectors suspected of corruption. See Chalendard, Fernandes, Raballand, and Rijkers (2023) for more details.

A3). Even for containers with misclassified products, a significant share of the tariff revenue losses is due to undervaluation (as per columns 2 to 4 in panel C of Table 2). In total the decomposition of the tariff revenue gaps in Table 2 shows that import undervaluation accounts for 67 percent while misclassification is responsible for the remaining 33 percent.

To summarize, discrepancies between recorded imports and exports are prevalent but small at the median. Yet, the average tariff revenue losses associated with such discrepancies are large because value gaps tend to increase with tariffs, and especially so for large shipments. Product misclassification is also prevalent, goes hand in hand with value gaps, and generates significant tariff revenue losses.

4.4. Explaining value gaps: Preliminary evidence

As an initial step in assessing the determinants of value gaps, we document the explanatory power of different fixed effects – Madagascan broker, inspector, and customs office, importer, exporter, incoterm, HS2 chapter (for simplicity designated as product), month-year – and hypothetical tariff rates for value gaps in Table 3. Panel A presents the R^2 and adjusted R^2 for regressions including each set of fixed effects or tariffs separately. Part B shows results for a Shapley decomposition of a regression that includes all variables following Shorrocks (2013). Each row reports a variable's average marginal contribution to the overall R^2 across all permutations of the order in which variables enter. Including all variables at once explains 71.2 percent of the variance in value gaps.

The most striking finding in Table 3 is that hypothetical tariff rates explain only a very small share of the variance in value gaps whereas importer and exporter fixed effects have the most explanatory power. More than half of the variance in value gaps can be explained by importer fixed effects alone, as is evidenced by the high adjusted R^2 of 0.524 when importer fixed effects

are accounted for. The Shapley value of importer fixed effects is 0.231. Exporter fixed effects have similarly high explanatory power but since most importers buy from one exporter only it is not clear to what extent this reflects the importance of importer rather than exporter identity. Broker fixed effects have substantial explanatory power, which is not surprising given that brokers fill in forms on behalf of importers. Contractual terms, proxied by incoterm fixed effects, also explain some of the variance in value gaps, especially incoterms declared in France. The types of products included in the container, proxied by HS2 product fixed effects, also have considerable explanatory power. Madagascan product fixed effects account for 19 percent of the variance in value gaps but their Shapley value is rather low (0.052).[32] Value gaps do not vary systematically with the customs office nor the inspector. To summarize, the identities of the importer and the exporter associated with the transaction are the best predictors of value gaps.

4.5. Concentration of value gaps

Value gaps are concentrated among a select few Madagascan importers, as is shown in Figure 5, which plots the centiles of the distributions of firm-level total value gaps and firm-level total tariff revenue gaps.[33] Panels (a) and (b) provide unscaled gaps in euros while panels (c) and (d) display gaps scaled by trade (the average between total FOB export values and total import FOB values of the importers within the centile). Despite the higher prevalence of importers with negative value gaps, the aggregate value gap (across all importers) is positive, at 12.6 million €, as detailed in Table 4, due to the highly asymmetric nature of the distribution of value gaps, with the aggregate driven by a few importers at the top centiles of the distribution. The top 5 importers in the distribution of firm-level total value gaps account for more than 80 percent of the aggregate value gap. Tariff revenue gaps are similarly concentrated, with the

[32] Finer fixed effects for HS4 products explain 42 percent of the variance in value gaps. This high explanatory power may be related to the fact that we have only 7,481 observations and 660 HS4 product fixed effects.
[33] Firm-level totals are obtained summing across the 2014-2016 period for all of the firm's containers. These are constructed for a universe of 602 importers with a tax ID.

top 5 importers accounting for 75 percent of the aggregate tariff revenue gap. This finding underscores the granularity of evasion with those few importers accounting for nearly three-quarters of aggregate tariff revenue losses.

One may be concerned that the concentration of evasion gaps is simply an artefact of the concentration of trade (Freund and Pierola, 2015). The lower panels of Figure 5 normalize value gaps and tariff revenue gaps by trade values to account for differences in importer size. These plots show that importers vary in their propensity to misreport trade values and in their tariff liabilities: most importers have very low reporting discrepancies in aggregate while a select group is characterized by a high propensity to misreport. Interestingly, evasion does not appear concentrated in the largest importers in terms of import value: panel C of Table 4 shows that the top 10 firms in terms of total tariff revenue gap account for only 15 percent of total imports.[34] We also find the pairwise correlation between the year-to-year rank of importers in terms of average value gaps to be 0.63. This suggests that discrepancies are part of systematic reporting biases by importers rather than being idiosyncratic to a few transactions.

5. Which transactions are most likely to be misreported?

In this section we examine which transactions are most at risk of being misreported and in particular whether reporting discrepancies are related to tariff evasion. Given the central role importers and exporters seem to play and the importance of the size of the transaction documented in Section 4.1, we also test whether certain types of importers and exporters are more prone to exhibiting discrepancies and what role shipment characteristics play. Our most general estimating equation takes the form:

$$Y_c = \beta_T HypTariffRate_c + \beta_F F_c + \beta_{FT} F_c * HypTariffRate_c$$

[34] Likewise, the 5 largest importers in terms of import value accounting for 26 percent of total imports account for only 0.8 percent of total tariff revenue losses (panel A of Table 4).

$$+\beta_X X_c + \beta_{XT} X_c * HypTariffRate_c$$

$$+P_c + O_c \ + I_c + B_d + \tau_d + \varepsilon_c \quad (5)$$

where Y_c reflect an outcome variable, either the value gap (VG_c) or an indicator variable for misclassification. $HypTariffRate_c$ is the hypothetical tariff rate defined in Section 2.4. F_c is a vector of importer and exporter characteristics, X_c is a vector of declaration characteristics, P_c is a vector of French and Madagascan HS 2-digit product fixed effects, O_c is a vector of Madagascar customs office fixed effects, I_c is a vector of inspector fixed effects, B_d is a vector of broker fixed effects, and τ_d is a vector of month-year fixed effects.[35] Ordinary least squares estimation is used with standard errors clustered by importer.

5.1. Determinants of value gaps

Table 5 presents the results of regressions in which the value gap serves as the dependent variable, using specifications that progressively add (groups of) explanatory variables. Column 1 includes only the hypothetical tariff rate and month-year fixed effects as controls. Value gaps are positively correlated with tariffs with an elasticity of 0.87. Adding controls for contractual terms, as is done in column 2, reveals that after shipment incoterms (i.e., the exporter is responsible for payment of transport costs) are linked to significantly higher value gaps.

Column 3 adds controls for shipment characteristics, notably indicators for larger export value (i.e. above the median) and multiproduct shipments, while controlling for HS2 product, Madagascar customs office, inspector, and broker fixed effects. Larger export value and multiproduct shipments are significantly more undervalued. Adding interactions between shipment characteristics and hypothetical tariff rates, as is done in column 4, shows that larger export value shipments are especially more likely to be undervalued when the tariff rate is high. Evasion thus

[35] The set of product fixed effects covers both exporter declared, and importer declared HS2 products and is not mutually exclusive in the sense that a given container may include multiple HS2 products.

appears concentrated in larger value shipments, presumably because the resulting tax savings are larger for such shipments.

We test the robustness of this result to the inclusion of importer and exporter characteristics in column 5 while interacting those with hypothetical tariff rates in column 6.[36] The significantly higher sensitivity of value gaps to tariffs for larger value shipments documented in column 4 is robust. Value gaps are on average significantly lower for SOE importers but not for PAD importers (i.e. trusted traders) though these can import with minimal controls and might have incentives to evade.[37] Non-commercial exporters (characterized by small average export values) are associated with significantly lower value gaps and those are not correlated with the tariff rate. In contrast, unregistered importers exhibit slightly higher value gaps on average and a significantly positive tariff elasticity, suggesting they are more likely to evade. The last column in Table 5 adds controls for the risk score, red channel inspection assignment, and whether containers were scanned, variables which are potentially endogenous to the value gap. While gaps are higher for containers with higher risk scores and those scanned, including these controls does not affect the other patterns documented.

We perform various robustness checks to our findings on the determinants of value gaps and present them in Appendix Section A2. To start with, Appendix Table A6 examines their robustness to adding further controls. It presents specifications that start from column 6 of Table 5 (reproduced in column 1) and control for the container's products being subject to non-tariff measures (column 2) or minimum values (column 3), or being used goods (column 4). We then also add French customs office fixed effects (column 5), additional firm characteristics (column 6), and interactions between tariffs and incoterms (column 7). None of these meaningfully

[36] Hypothetical tariff rates are not interacted with FZE importers since they face 0 tariff rates.
[37] The fact that SOEs have lower value gaps could be related to them facing softer budget constraints and having weaker incentives to evade. It could also reflect different managerial practices.

impact the patterns of results documented above, though it is worth noting that gaps are lower for containers including used goods and imported by firms that both export and import and are relatively young.

Appendix Table A7 presents results from estimating Equation (5) for our preferred specification (again replicated in column 1) but using different dependent variables; column 2 uses value gaps computed using an alternative exchange rate to convert import values in euros, notably the one that applied on the date the declaration was registered by Madagascar customs (instead of the average of the exchange rates used at the day of departure and the day of registration). This does not meaningfully impact the estimates. Column 3 uses value gaps in which no adjustment for transport costs are made. While results are qualitatively similar to those obtained using our baseline measure of value gaps, the elasticity of value gaps with respect to tariffs for larger export value shipments is of smaller magnitude. Finally, columns 4 and 5 use the weight gap and unit price gap, respectively as dependent variables. Weight gaps are harder to predict, and, moreover, do not correlate as strongly with tariffs as unit price gaps, consistent with undervaluation being the primary driver of value gaps.[38]

Last but not least Appendix Table A8 shows that the results are robust to restricting the sample to observations with 1 French declaration per container (column 2), dropping casual traders i.e., the non-commercial exporters in France and the unregistered importers in Madagascar (column 3), dropping containers with conflicting incoterms (column 4), and limiting the sample to containers with a negligible weight gap (of less than 1%) (column 5).

[38]As an additional test of the hypothesis that value gaps indeed partially reflect evasion, we include the excess interaction proxy for potential corruption from Chalendard, Fernandes, Raballand, and Rijkers (2023) (discussed above) in our value gaps regressions along with the rest of the determinants in Table 5. The results presented in Appendix Table A9 show that value gaps for containers cleared by customs inspectors who interact excessively with the brokers registering them exhibit an especially strong positive elasticity to tariffs. The containers identified as potentially subject to corruption are those that also feature larger value gaps for larger tariffs.

5.2. Determinants of product misclassification

To examine the determinants of strategic product misclassification, we estimate Equation (5) using as dependent variable the tariff-decreasing product misclassification indicator defined in Section 3.[39] Results are presented in Table 6. Tariff-decreasing misclassification is positively correlated with tariffs as is shown in column 1 and this association remains after controlling for contractual terms in column 2, and shipment characteristics in column 3, which shows that containers with larger export value and multiproduct shipments are more susceptible to strategic misclassification. Such misclassification is especially prevalent among containers with larger export value that are subject to high tariffs. The introduction of importer and exporter characteristics and their interactions with hypothetical tariff rates in columns 5 and 6 does not alter the previous findings. Containers with shipments sent by non-commercial exporters exhibit a significantly higher likelihood of tariff-decreasing misclassification with a strong significantly positive tariff elasticity. According to our preferred estimates in column 6, larger export value containers subject to a 20 percent tariff rate (the maximum rate in our sample) are 42 percent (-0.037+0.200*2.287=0.420) more likely to exhibit tariff-decreasing misclassification than containers with lower export value.

Column 7 shows results for the same specification as in column 6 but using the non-strategic product misclassification indicator as dependent variable. Larger export value containers are significantly more likely to feature non-strategic product misclassification, but this effect decreases with tariffs. Only tariff-decreasing misclassification appears to be driven by tariff evasion motives whereas tariff-increasing misclassification is unrelated to tariff rates.

[39] Note that we do not control for HS 2-digit product fixed effects in these specifications.

6. Does Customs effectively fight evasion?

To end, we examine to what extent customs inspectors in the importing country reduce tariff evasion, and whether they effectively target declarations with an elevated risk of tariff evasion, i.e., those with high value gaps, by examining the difference between the log of the final import value retained by customs and the log of the value initially declared by the importer per container $\Delta Value_c$ (defined as a difference in logarithms). The difference between these values can plausibly be attributed to customs inspectors' actions. To assess customs' effectiveness we estimate regressions of the form:

$$\Delta Value_c = \beta_G VG_c + \beta_T HypTariff_c + \beta_X X_c + P_c + O_c + I_c + B_d + \tau_d + \varepsilon_c \quad (6)$$

where variables are defined as in Equation (5). The coefficient of interest is β_G and if value gaps were fully attributable to evasion and the customs agency operated efficiently we would expect β_G=1. This benchmark is admittedly extreme but helps anchor the discussion about the effectiveness of customs. Results from estimating Equation (6) are presented in Table 7. Column 1 presents the raw correlation between customs value adjustment and value gaps and shows that they are not correlated. The coefficient on the value gap is close to zero (-0.003) and this remains true even after we control for tariffs and contractual terms in column 2. Column 3 adds an interaction term between tariffs and value gaps to test whether customs are more likely to close such gaps when tariffs are high (and hence evasion more costly). The interaction between the gap and the tariff rate is statistically insignificantly negative, suggesting this is not the case. Column 4 adds importer and exporter characteristics. This results in a positive correlation between value gaps and customs value adjustment, but the correlation remains weak. Customs inspectors are significantly less likely to redress value gaps in declarations lodged by SOE, FZE, PAD and unregistered importers. By contrast, they are more likely to redress

the value of declarations sent by exporters who are not registered. Controlling for declaration characteristics, as is done in column 5, strengthens slightly the positive correlation between value adjustment and value gaps that becomes significant at the 10 percent level. Intriguingly, larger export value shipments are less likely to be revised by customs inspectors, ceteris paribus. Adding product, customs office, broker, and inspector fixed effects in column 6 renders the importer and exporter characteristics insignificant. The elasticity of value adjustment to value gaps becomes significant at the 5 percent even though it remains small in magnitude, with only 0.8 percent of value adjustment being realized to reduce the value gap.

In sum, value adjustment by customs inspectors is at best weakly correlated with value gaps and inspectors appear to treat different types of firms differently: they are most likely to insist on value adjustments for importers that are not SOE, FZE, or PAD, while least likely to change the values declared by unregistered importers. Customs reduce the costs associated with reporting discrepancies but this impact remains limited by the low prevalence of adjustments and the poor targeting as revisions to customs value are only very weakly correlated with reporting discrepancies.

7. Conclusion

Identifying which firms cheat and which transactions are most at risk of evasion is central to our understanding of tariff incidence and the design of policies to remedy tariff evasion. This paper matches customs transaction data on French exports with customs transaction data on Madagascan imports using container IDs to construct container-specific measures of reporting discrepancies. Differences due to freight and insurance costs are carefully accounted for. Without them average tariff evasion would be underestimated by 12.7% on average. Reporting discrepancies between exporters and their matched importers are highly prevalent but small

on average. Over two-fifths of all importers report in a way that increases their tariff liability relative to what it would have been had they declared their imports in the same way as the French exporters they were buying from recorded their shipments, suggesting these firms are not strategically minimizing tariffs. Yet, the aggregate costs of evasion are sizable, amounting to approximately 24% of current tariff revenues, because evasion is concentrated in large shipments subject to high tariffs.

Tariff evasion is also highly concentrated among importers, with the top five evaders accounting for three-quarters of all revenue losses. The identity of the importer is the best predictor of evasion and dominates the predictive power of product characteristics. Unregistered importers are more likely to engage in evasion, while FZE and SOE importers are less likely to do so.

Evasion goes hand in hand with product misclassification, which is highly prevalent and accounts for roughly a third of all tariff revenue losses associated with reporting discrepancies. For only half of all declarations is the set of products recorded in France and in Madagascar identical. The propensity to strategically misclassify is especially high for larger value shipments subject to high tariffs, presumably because incentives to misclassify are higher as more revenue is at stake. Yet, evasion is by no means the only determinant of product misclassification.

Customs inspectors reduce the costs associated with reporting discrepancies but the aggregate impact of their actions is limited by the very low prevalence of adjustments and the poor targeting since revisions to customs value are only very weakly correlated with reporting discrepancies.

The fact that evasion is so highly concentrated among a select few importers suggests that successfully combatting evasion by a handful of firms could have sizable macro-fiscal implications. At the same time, our findings are a reminder that reporting discrepancies in trade data (and the literature based on mirror trade statistics) must be interpreted with a great deal of caution,

since a sizable portion of them appear driven by genuine misreporting and differences in the way freight and insurance costs are accounted for.

References

Allingham, M. G., and A. Sandmo (1972): "Income tax evasion: A theoretical analysis," *Journal of public economics*, 1(3-4), 323–338.

Almunia, M., J. Hjort, J. Knebelmann, and L. Tian (2022): "Strategic or confused firms? evidence from "missing" transactions in Uganda," *Review of Economics and Statistics*, pp. 1–35.

Andreoni, J., B. Erard, and J. Feinstein (1998): "Tax compliance," *Journal of economic literature*, 36(2), 818–860.

Baunsgaard, T., and M. Keen (2010): "Tax revenue and (or?) trade liberalization," *Journal of Public Economics*, 94(9-10), 563–577.

Berman, N., P. Martin, and T. Mayer (2012): "How do different exporters react to exchange rate changes?," *The Quarterly Journal of Economics*, 127(1), 437–492.

Besley, T., and T. Persson (2014): "Why do developing countries tax so little?," *Journal of economic perspectives*, 28(4), 99–120.

Bussy, A. (2023): "Corporate tax evasion: Evidence from international trade," *Available at SSRN 3607016*.

Cagé, J., and L. Gadenne (2018): "Tax revenues and the fiscal cost of trade liberalization, 1792–2006," *Explorations in Economic History*, 70, 1–24.

Carrère, C., and C. Grigoriou (2015): "Can mirror data help to capture informal international trade?," Discussion paper, FERDI Working Paper.

Carrillo, P., D. Pomeranz, and M. Singhal (2017): "Dodging the taxman: Firm misreporting and

limits to tax enforcement," *American Economic Journal: Applied Economics*, 9(2), 144–164.

Chalendard, C., A. M. Fernandes, G. Raballand, and B. Rijkers (2023): "Corruption in customs," *The Quarterly Journal of Economics*, 138(1), 575–636.

Chalendard, C. R., A. Duhaut, A. M. Fernandes, A. Mattoo, G. J. Raballand, and B. Rijkers (2020): "Does Better Information Curb Customs Fraud?," World Bank Policy Research Working Paper 9254.

Coşar, A. K., and B. Demir (2018): "Shipping inside the box: Containerization and trade," *Journal of International Economics*, 114, 331–345.

Di Giovanni, J., and A. A. Levchenko (2012): "Country size, international trade, and aggregate fluctuations in granular economies," *Journal of Political Economy*, 120(6), 1083–1132.

Di Giovanni, J., A. A. Levchenko, and I. Mejean (2014): "Firms, destinations, and aggregate fluctuations," *Econometrica*, 82(4), 1303–1340.

Eaton, J., S. Kortum, and F. Kramarz (2004): "Dissecting trade: Firms, industries, and export destinations," *American Economic Review*, 94(2), 150–154.

Ferrantino, M. J., X. Liu, and Z. Wang (2012): "Evasion behaviors of exporters and importers: Evidence from the US–China trade data discrepancy," *Journal of international Economics*, 86(1), 141–157.

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

Freund, C., and M. D. Pierola (2015): "Export superstars," *Review of Economics and Statistics*, 97(5), 1023–1032.

Gabaix, X. (2011): "The granular origins of aggregate fluctuations," *Econometrica*, 79(3), 733–772.

Gordon, R., and W. Li (2009): "Tax structures in developing countries: Many puzzles and a possible explanation," *Journal of Public Economics*, 93(7-8), 855–866.

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013): "Who creates jobs? Small versus large versus young," *Review of Economics and Statistics*, 95(2), 347–361.

Hummels, D. (2007): "Transportation costs and international trade in the second era of globalization," *Journal of Economic perspectives*, 21(3), 131–154.

Javorcik, B. S., and G. Narciso (2008): "Differentiated products and evasion of import tariffs," *Journal of International Economics*, 76(2), 208–222.

Jean, S., and C. Mitaritonna (2010): "Determinants and pervasiveness of the evasion of customs duties," Discussion paper.

Jean, S., C. Mitaritonna, A. Vatan, et al. (2018): *Institutions and Customs Duty Evasion*. CEPII.

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

Kellenberg, D., and A. Levinson (2019): "Misreporting trade: Tariff evasion, corruption, and auditing standards," *Review of International Economics*, 27(1), 106–129.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011): "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica*, 79(3), 651–692.

Krizan, C. J., J. R. Tybout, Z. Wang, and Y. Zhao (2020): *Are Customs Records Consistent Across Countries?: Evidence from the US and Colombia*. US Census Bureau, Center for Economic Studies.

Mayer, T., M. J. Melitz, and G. I. Ottaviano (2014): "Market size, competition, and the product

mix of exporters," *American Economic Review*, 104(2), 495–536.

Mishra, P., A. Subramanian, and P. Topalova (2008): "Tariffs, enforcement, and customs evasion: Evidence from India," *Journal of public Economics*, 92(10-11), 1907–1925.

Naritomi, J. (2019): "Consumers as tax auditors," *American Economic Review*, 109(9), 3031–3072.

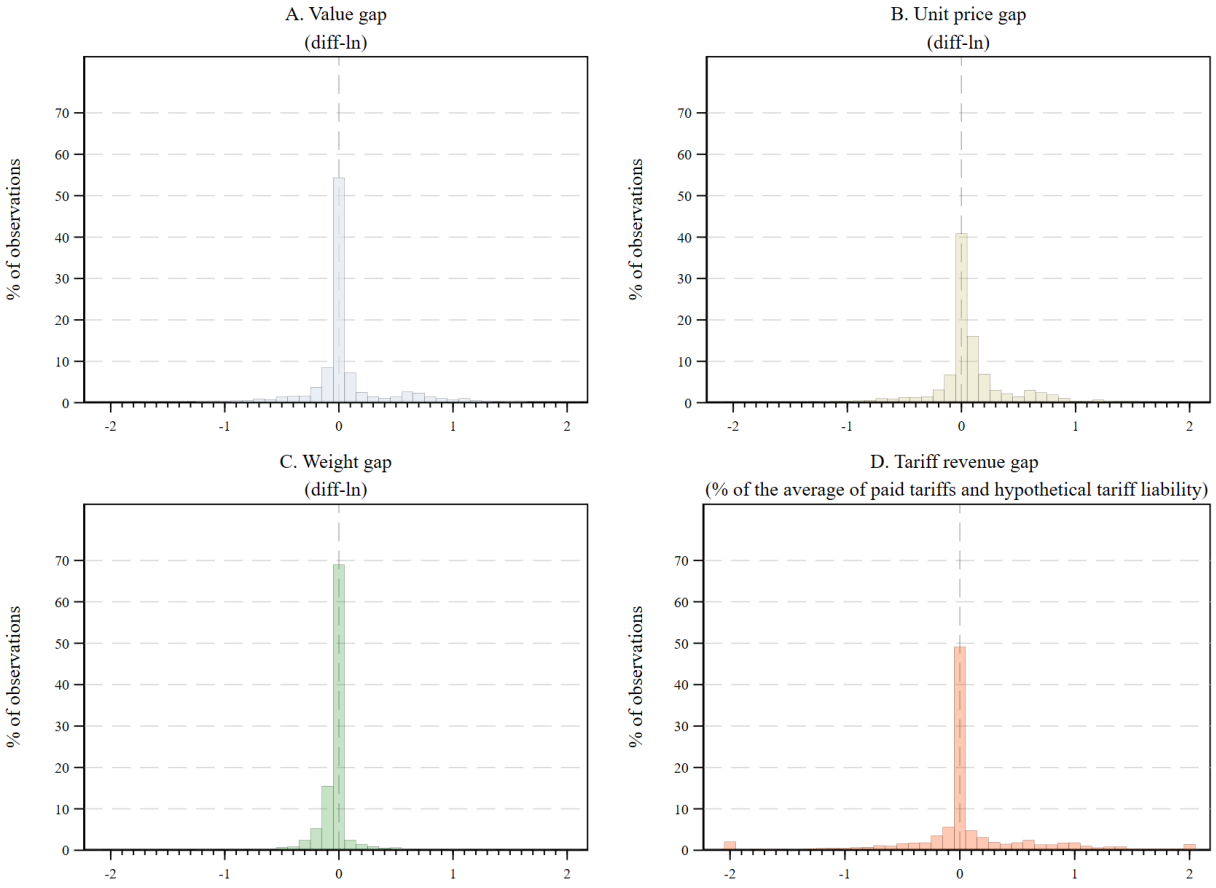
Sequeira, S. (2016): "Corruption, trade costs, and gains from tariff liberalization: Evidence from Southern Africa," *American Economic Review*, 106(10), 3029–63.

Sequeira, S., and S. Djankov (2014): "Corruption and firm behavior: Evidence from African ports," *Journal of International Economics*, 94(2), 277–294.

Shorrocks, A. F. (2013): "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value," *The Journal of Economic Inequality*, 11(1), 99–126.

Slemrod, J. (2019): "Tax compliance and enforcement," *Journal of Economic Literature*, 57(4), 904–54.

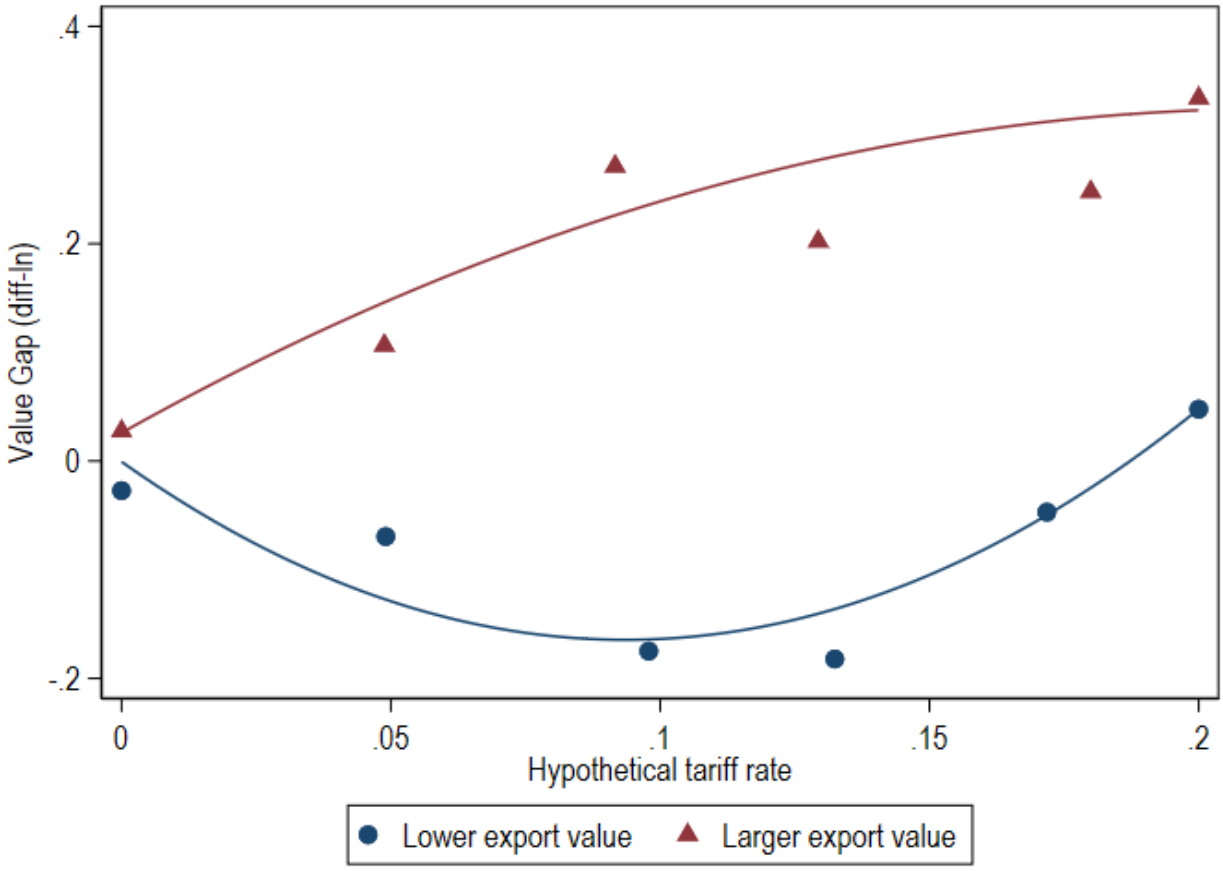
Figure 1 – Distribution of reporting discrepancies across containers



Note: the plots in Panels A, B, and C exclude observations (containers) in the first and last centile of distributions.

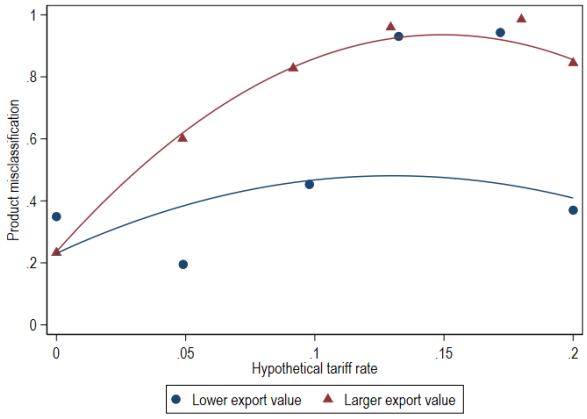
1.8

Figure 2 – Value gaps versus tariffs

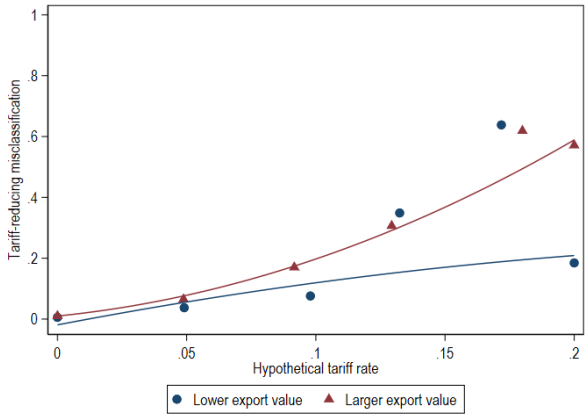


Note: The figure shows a binned scatterplot that groups observations (containers) in equal sized bins according to hypothetical tariff rate (the x-axis variable) and plots the averages of that x-axis variable against those of the value gap (diff-ln) (the y-axis variable). Observations are divided into two groups based on whether the export value reported in France is above or below the sample median. A quadratic fit estimated using OLS is plotted.

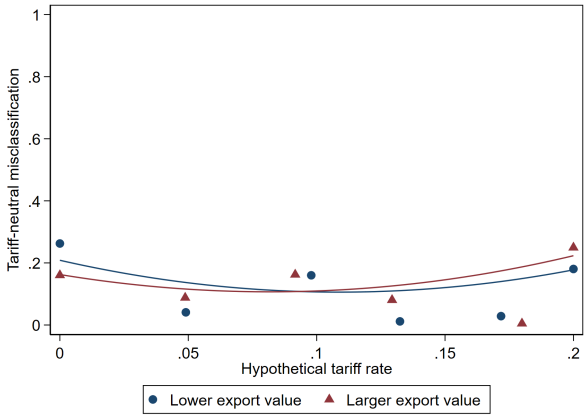
Figure 3 – Misclassification versus tariffs



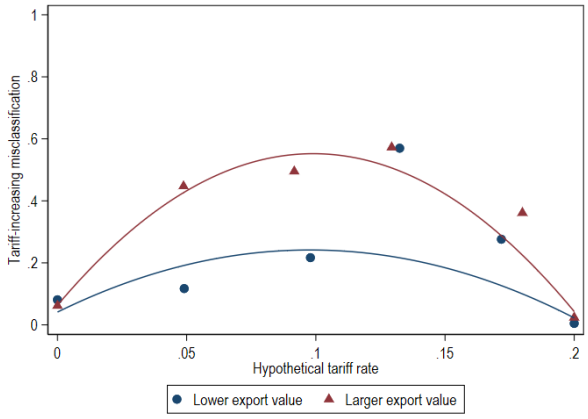
(a) Misclassification (any)



(b) Tariff-decreasing misclassification



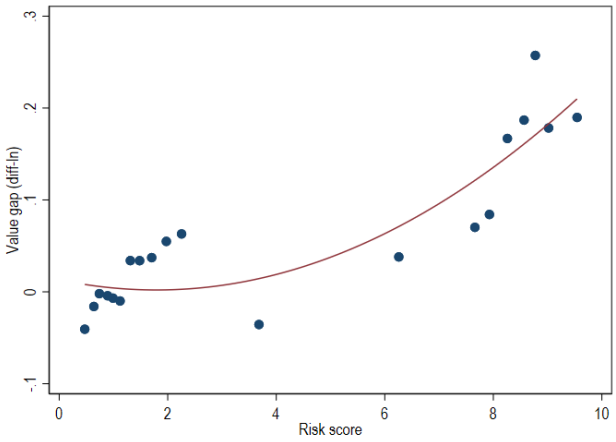
(c) Tariff-neutral misclassification



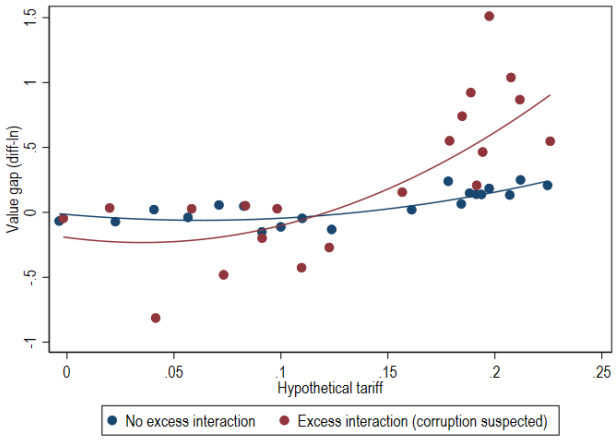
(d) Tariff-increasing misclassification

Note: all panels in the figure show a binned scatterplot that groups observations (containers) in equal sized bins according to the x-axis variable (hypothetical tariff rate) and plots the averages of that x-axis variable against those of the y-axis variable (product misclassification indicator) over each bin. Observations are divided into two groups based on whether the export value reported in France is above or below the sample median. A quadratic fit estimated using OLS is plotted.

Figure 4 – Value gaps versus corruption risk



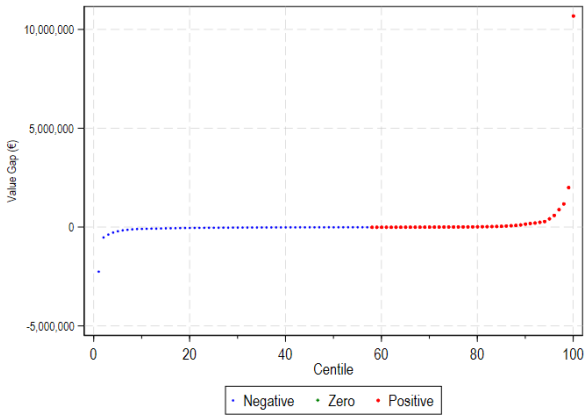
(a) Value gap versus risk scores



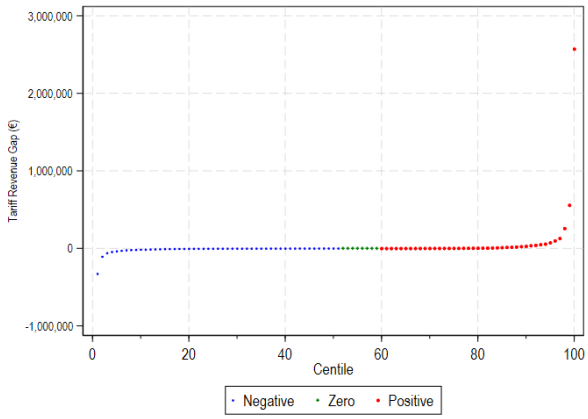
(b) Value gap versus tariffs - by corruption risk

Note: both panels show a binned scatterplot that groups observations (containers) in equal sized bins according to the x-axis variable (the risk score in panel a and the hypothetical tariff rate in panel b) and plots the averages of that x-axis variable against those of the y-axis variable (the value gap) over each bin. Observations in panel b are divided into two groups depending on whether or not they are handled by an inspector who interacted excessively frequently with the broker registering them (if this is the case, corruption is suspected). A quadratic fit estimated using OLS is plotted.

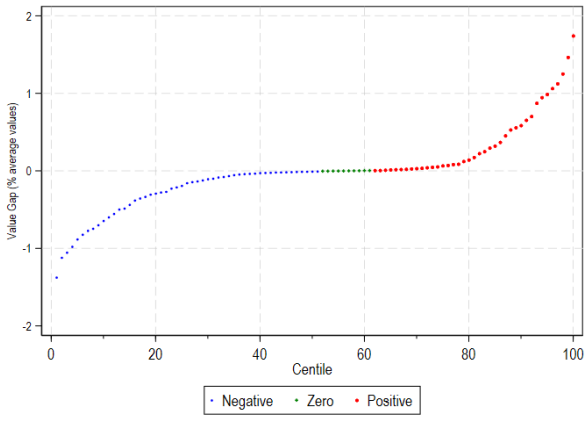
Figure 5 – Concentration of value gaps among importers



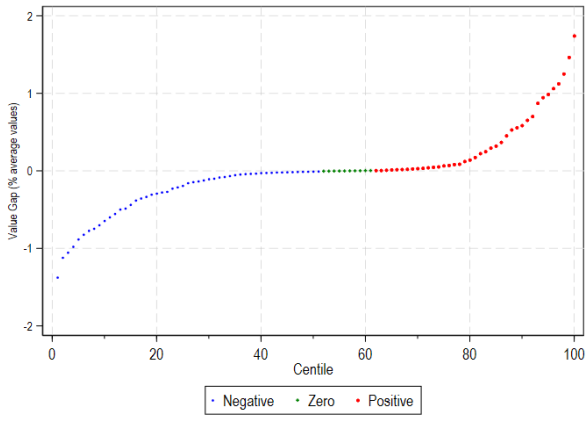
(a) Value gaps



(b) Tariff revenue gaps



(c) Value gaps - scaled by trade



(d) Tariff-Revenue gap - scaled by trade

Note: the sample is based on 602 Madagascan importers. Scaled gaps correspond to unscaled gaps divided by the average between export and import FOB values. Value and tariff revenue gaps are considered to fall into the zero category when they are lower than 1€, scaled value and tariff gaps are considered as zero when within they account for less than 0.5% of average trade flows (i.e., the average of exports and imports)

Table 1 – Descriptive statistics (N=7461)

	Mean	s.d.	Min	Median	Max
Trade and value gaps					
Export FOB value (€)	24 708	30 542	1 010	16 794	604 500
Initial import CIF value (€)	24 986	31 330	1 271	16 495	601 559
Initial import FOB value (€)	23 016	30 448	1 001	14 242	525 472
Transport costs (€)	1 975	1 855	0	1 714	76 087
Transport costs (% Initial import CIF)	0.127	0.095	0.000	0.102	0.759
Transport costs (% Export FOB)	0.158	0.175	0.000	0.106	2.424
Initial value gap (€)	1 692	11 542	-180 322	-1	395 038
Initial value gap (diff-ln)	0.063	0.448	-3.103	0.000	4.685
Weight and unit price gaps					
Weight gap (kg)	-385	2 007	-21 097	-10	20 589
Weight gap (diff-ln)	-0.030	0.252	-7.004	-0.001	5.096
Initial unit price gap (diff-ln)	0.093	0.467	-3.692	0.028	7.351
Customs value adjustment					
Value adjustment (frequency)	0.037	0.189	0.000	0.000	1.000
Value adjustment (€)	65.897	709	-6 624	0	42 604
Value adjustment (diff-ln)	0.008	0.065	-0.381	0.000	1.653
Tariff revenue adjustment (€)	8	74	-1 325	0	1 705
Value adj. cond. on adjusting (€)	1 769	3 245	-6 624	1 302	42 604
Value adj. cond. on adjusting (diff-ln)	0.225	0.256	-0.381	0.175	1.653
Tariff rev. adjustment (€) (if>0)	216	320	-1 325	163	1 705
Product misclassification					
Product misclassification	0.507	0.500	0.000	1.000	1.000
Tariff-decreasing misclassification	0.193	0.395	0.000	0.000	1.000
Tariff-neutral misclassification	0.160	0.366	0.000	0.000	1.000
Tariff-increasing misclassification	0.154	0.361	0.000	0.000	1.000
Tariffs and tariff losses					
Hypothetical tariff rate	0.109	0.080	0.000	0.100	0.200
Applied tariff rate	0.106	0.078	0.000	0.100	0.200
Hypothetical tariff liability (€)	2 266	3 209	0	929	44 324
Estimated paid tariffs on FOB value (€)	1 828	2 424	0	974	42 377
Tariff revenue gap (€)	438	2 297	-15 172	0	25 880

Table 2 – Decomposition of tariff revenue losses

	Correct Classification	Misclassification			Total
		Tariff-decreasing	Tariff-neutral	Tariff-increasing	
A. Prevalence					
Number of containers	3680	1441	1191	1149	7461
% of containers	49.3%	19.3%	16.0%	15.4%	100%
B. Average tariff revenue loss per container €					
Undervaluation	121.11	464.94	605.96	315.93	294.91
Misclassification (residual)	0	1388.64	0.01	-813.07	142.99
Total	121.11	1853.58	605.97	-497.14	437.90
C. Share of total tariff revenue losses (%)					
Undervaluation	13.6%	20.5%	22.1%	11.1%	67.4%
Misclassification (residual)	0.0%	61.2%	0.0%	-28.6%	32.6%
Total	13.6%	81.8%	22.1%	-17.5%	100%

Table 3 – Shapley Decomposition

Dependent variable: Value gap (diff-ln) - Container level						
Category of fixed effects	**A - Separate regressions**				**B - Shapley adjusted R^2**	
	R^2	Adjusted R^2	Df	N	Shapley val	% adjusted R^2
MDG brokers	0.252	0.241	91	7461	0.086	12.08%
MDG inspectors	0.036	0.025	80	7461	0.012	1.66%
MDG customs offices	0.013	0.012	10	7461	0.006	0.84%
MDG importers	0.562	0.524	343	7461	0.231	32.38%
French exporters	0.556	0.508	344	7461	0.246	34.60%
Importers + Exporters	0.693	0.642	484	7461		
Incoterm indicators[a]	0.037	0.037	3	7385		
Fr incoterm	0.086	0.085	10	7461	0.026	3.67%
MDG incoterm	0.022	0.02	9	7461	0.007	1.00%
Fr incoterm + MDG incoterm	0.114	0.111	19	7461		
HS2 chapter MDG	0.194	0.184	86	7461	0.052	7.30%
HS2 chapter Fr	0.141	0.132	72	7461	0.037	5.21%
HS2 chapter Fr or MDG	0.204	0.195	88	7461		
HS2 chapter MDG + HS2 chapter Fr	0.255	0.238	156	7461		
HS4 chapter MDG	0.421	0.353	660	7461		
HS4 chapter Fr	0.380	0.335	379	7461		
HS4 chapter Fr or MDG	0.451	0.386	680	7461		
HS4 chapter MDG + HS4 chapter Fr	0.571	0.485	943	7461		
Month-year	0.01	0.005	35	7461	0.004	0.50%
Hypothetical tariff rate	0.024	0.023	1	7461	0.005	0.76%
All Shapley			815	7461	0.712	

Notes: [a] Incoterm indicators consist of After Shipment Incoterm in both countries (ASI), BSI in France but ASI in Madagascar, and ASI in France but BSI in Madagascar.

Table 4 – Concentration of gaps among Madagascar's top importers

Rank	% total M	% total VG	% total TRG	% containers
	A. Top importers based on total import FOB value			
Top 5	25.7%	1.0%	0.8%	16.4%
Top 10	42.2%	-1.1%	28.7%	37.0%
	B. Top importers based on total value gap			
Top 5	6.1%	81.1%	49.8%	12.5%
Top 10	7.3%	97.1%	58.8%	14.1%
	C. Top importers based on total tariff revenue gap			
Top 5	8.6%	70.4%	75.0%	12.2%
Top 10	15.4%	83.4%	92.3%	24.6%
	D. All importers			
602 importers	Total M (€)	Total VG (€)	Total TRG (€)	N containers
	170 707 491	12 586 508	3 242 304	7 119

Notes: Importers without a tax identifier are excluded from the sample.

Table 5 – Determinants of value gaps

	Value gap (diff ln.)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hypothetical tariff rate	0.869***	1.105***	0.738***	0.000	-0.038	-0.087	-0.194
	(0.316)	(0.410)	(0.214)	(0.276)	(0.276)	(0.286)	(0.288)
After shipment incoterm in both countries		0.196**	0.092***	0.081**	0.088***	0.086***	0.077**
		(0.079)	(0.031)	(0.032)	(0.033)	(0.033)	(0.031)
BSI in France but ASI in MDG		0.054	0.034	0.029	0.031	0.029	0.039
		(0.086)	(0.038)	(0.039)	(0.037)	(0.038)	(0.042)
ASI in France but BSI in MDG		0.372**	0.162***	0.154**	0.155***	0.152***	0.149**
		(0.180)	(0.062)	(0.062)	(0.059)	(0.059)	(0.058)
Larger export value			0.272***	0.113**	0.109**	0.094**	0.097**
			(0.043)	(0.045)	(0.046)	(0.040)	(0.039)
Hypothetical tariff rate * Larger export value				1.332***	1.270***	1.419***	1.440***
				(0.368)	(0.382)	(0.373)	(0.370)
Multiproduct shipment (Fr)			0.062**	0.038	0.031	0.047	0.040
			(0.031)	(0.042)	(0.042)	(0.043)	(0.043)
Hypothetical tariff rate * Multiproduct shipment (Fr)				0.189	0.256	0.109	0.154
				(0.370)	(0.347)	(0.303)	(0.297)
SOE					-0.120***	-0.034	-0.047
					(0.031)	(0.049)	(0.046)
FZE					-0.109	-0.099	-0.097
					(0.070)	(0.070)	(0.071)
Unregistered importer					0.090*	-0.085	-0.090
					(0.051)	(0.085)	(0.085)
PAD					-0.009	-0.013	-0.006
					(0.028)	(0.043)	(0.042)
Hypothetical tariff rate * SOE						-1.107**	-0.995*
						(0.518)	(0.515)
Hypothetical tariff rate * Unregistered importer						1.278**	1.315**
						(0.580)	(0.582)
Hypothetical tariff rate * PAD						0.018	0.122
						(0.312)	(0.302)
Non-commercial exporter					-0.332***	-0.365***	-0.382***
					(0.057)	(0.115)	(0.117)
Hypothetical tariff rate * Non-commercial exporter						0.228	0.299
						(0.745)	(0.749)
Risk score							0.014***
							(0.003)
Red channel							-0.034
							(0.025)
Scanned							0.051*
							(0.027)
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
French and Madagascan HS 2-digit product FE	No	No	Yes	Yes	Yes	Yes	Yes
Madagascar customs office FE	No	No	Yes	Yes	Yes	Yes	Yes
Inspector FE	No	No	Yes	Yes	Yes	Yes	Yes
Broker FE	No	No	Yes	Yes	Yes	Yes	Yes
N. obs.	7461	7385	7385	7385	7385	7385	7380
R^2	0.033	0.085	0.470	0.476	0.487	0.489	0.495
Adjusted R^2	0.029	0.080	0.440	0.446	0.458	0.460	0.465

Notes: The table presents OLS estimates of the determinants of value gaps. Standard errors are clustered by importer and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. FE stands for fixed effects.

Table 6 – Determinants of product misclassification

Misclassification type	Tariff-decreasing						Non-strategic
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hypothetical tariff rate	1.924***	1.784***	2.020***	1.179**	1.061**	0.476	-0.782
	(0.499)	(0.455)	(0.369)	(0.426)	(0.386)	(0.339)	(0.410)
After shipment incoterm in both countries		-0.004	-0.030	-0.033	-0.029	-0.034	-0.023
		(0.062)	(0.047)	(0.044)	(0.044)	(0.039)	(0.039)
BSI in France but ASI in MDG		0.223	0.214*	0.195*	0.174**	0.145*	-0.043
		(0.140)	(0.089)	(0.076)	(0.066)	(0.060)	(0.042)
ASI in France but BSI in MDG		-0.000	-0.066	-0.064	-0.037	-0.038	0.001
		(0.069)	(0.051)	(0.054)	(0.049)	(0.048)	(0.063)
Larger export value			0.225***	0.011	0.013	-0.037	0.293***
			(0.046)	(0.051)	(0.051)	(0.042)	(0.069)
Hypothetical tariff rate * Larger export value				1.771**	1.868***	2.287***	-1.248**
				(0.565)	(0.558)	(0.472)	(0.482)
Multiproduct shipment (Fr)			0.084**	0.078	0.045	0.100*	0.240***
			(0.030)	(0.052)	(0.052)	(0.046)	(0.071)
Hypothetical tariff rate * Multiproduct shipment (Fr)				-0.050	-0.006	-0.445	0.017
				(0.475)	(0.452)	(0.394)	(0.667)
FZE					0.059	0.057	-0.283**
					(0.054)	(0.054)	(0.098)
SOE					0.031	0.145*	-0.265*
					(0.061)	(0.058)	(0.112)
Unregistered importer					0.002	-0.085	0.072
					(0.032)	(0.068)	(0.080)
PAD					0.075	-0.024	0.093
					(0.051)	(0.046)	(0.075)
Hypothetical tariff rate * SOE						-1.459**	1.473
						(0.518)	(0.860)
Hypothetical tariff rate * Unregistered importer						0.340	-0.117
						(0.483)	(0.580)
Hypothetical tariff rate * PAD						0.910	-0.532
						(0.496)	(0.510)
Non-commercial exporter					0.277***	-0.228*	0.444***
					(0.037)	(0.090)	(0.120)
Hypothetical tariff rate * Non-commercial exporter						4.111***	-3.350***
						(0.668)	(0.830)
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
French and Madagascan HS 2-digit product FE	No	No	No	No	No	No	No
Madagascar customs office FE	No	No	Yes	Yes	Yes	Yes	Yes
Inspector FE	No	No	Yes	Yes	Yes	Yes	Yes
Broker FE	No	No	Yes	Yes	Yes	Yes	Yes
N. obs.	7461	7385	7385	7385	7385	7385	7385
R^2	0.160	0.188	0.402	0.421	0.439	0.462	0.374
Adjusted R^2	0.156	0.184	0.383	0.402	0.421	0.444	0.353

Notes: The table presents OLS estimates of the determinants of tariff-decreasing misclassification in columns 1 through 6 and non-strategic misclassification in column 7. Standard errors are clustered by importer and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. FE stands for fixed effects. Containers with tariff-decreasing misclassification feature both product misclassification and a misclassification component of the tariff revenue gap above 1€.

Table 7 – Determinants of value adjustment by customs inspectors

	Customs value adjustment (diff-ln)					
	(1)	(2)	(3)	(4)	(5)	(6)
Value gap (diff-ln)	-0.003	-0.002	-0.007	0.004	0.005*	0.008**
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.003)
Hypothetical tariff rate		0.019	0.016	-0.012	-0.014	0.003
		(0.018)	(0.017)	(0.018)	(0.019)	(0.021)
After shipment incoterm in both countries		-0.007**	-0.007**	-0.000	0.001	0.001
		(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
BSI in France but ASI in MDG		-0.008**	-0.008*	-0.001	0.001	0.000
		(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
ASI in France but BSI in MDG		-0.005	-0.005	-0.001	-0.000	0.002
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Value gap (diff-ln) * Hypothetical tariff rate			0.038			
			(0.038)			
SOE				-0.006**	-0.006**	-0.001
				(0.002)	(0.002)	(0.003)
FZE				-0.009***	-0.006*	-0.003
				(0.003)	(0.003)	(0.007)
Unregistered importer				-0.036***	-0.037***	-0.034***
				(0.008)	(0.008)	(0.007)
PAD				-0.007***	-0.006***	-0.003
				(0.002)	(0.002)	(0.002)
Non-commercial exporter				0.051***	0.049***	0.018
				(0.010)	(0.010)	(0.012)
Multiproduct shipment (Fr)					0.002	-0.002
					(0.002)	(0.003)
Larger export value					-0.006**	-0.009***
					(0.002)	(0.003)
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes
French and Madagascan HS 2-digit product FE	No	No	No	No	No	Yes
Madagascar customs office FE	No	No	No	No	No	Yes
Inspector FE	No	No	No	No	No	Yes
Broker FE	No	No	No	No	No	Yes
N	7461	7385	7385	7385	7385	7385
R^2	0.017	0.020	0.021	0.058	0.059	0.162
Adjusted R^2	0.012	0.015	0.015	0.052	0.053	0.114

Notes: The table presents OLS estimates of the determinants of value adjustments made by customs inspectors during the customs clearance process. Standard errors are clustered by importer and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. FE stands for fixed effects.

Appendix

Supplemental Material for "Containing Tariff Evasion"

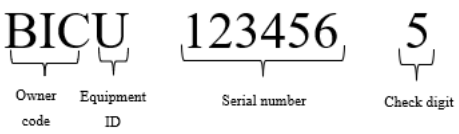
C. Anne, C. Chalendard, A.Fernandes, B. Rijkers, and V. Vicard

A1. Data construction

A1.1. Container identification

Container identification according to ISO 6346 is made of 3 string characters (owner or operator code), a letter (equipment identifier), 6 digits (serial number), and a 7th digit (check digit computed thanks to the previous characters as shown in Appendix Figure A1.

Figure A1 – Example of an ISO 6346 container ID



A1.2. Customs data sources and incoterms

Customs data To ensure the French customs data contain the key characteristics of interest, such as the container ID, we restrict attention to declarations registered following DELTA (Dédouanement en Ligne par Traitement Automatisé) procedures, an online procedure recording French customs clearances. French customs declarations not registered under DELTA correspond mainly to paper declarations of occasional exporters that represent a small share of declarations (5.5%) and exported value (1.7%). Every French customs declaration is recorded in a Single Administrative Document (SAD) defined by a unique identifier.

We use data provided by the Madagascan customs administration on import declarations from France (labeled as country of provenance or origin) over the period 2014-2016. This data is collected via the Automated System for Customs Data (ASYCUDA), an IT customs management system developed by the United Nations Conference on Trade and Development (UNCTAD) and adopted by more than 100 countries. For each declaration, the customs database is complemented by data from GasyNet, a private-public-partnership that assists Madagascan customs with risk analysis, on a proprietary risk score, an inspection clearance channel recommendation, and the presence of a scan.[40]

Incoterms Incoterms are predefined commercial terms included in international trade contracts that help to identify the division of risks and responsibilities between the buyer (importer) and the seller (exporter) as well as the place where the importer becomes responsible for the shipment. Definitions of incoterms are provided by the International Chamber of Commerce (ICC) and we rely on the 2010 incoterms which identify 11 incoterms in our sample.[41] We classify incoterms

[40] The clearance channel can be either the red channel in case of physical frontline inspection and documentary control, the yellow channel in case of only documentary control, or the blue channel without any inspection.

[41] Incoterms also define the responsibility for other charges such as insurance costs. However, since insurance costs represent on average under 1% of import value (when available), we define below incoterm groups depending on the entity in charge of the main transport costs only.

into two groups depending on who is in charge of main transport costs between France and Madagascar:

- Before Shipment Incoterms (BSI) consist of 4 incoterms where risks and responsibilities change *before* the main transportation so the importer is responsible for the main transportation (sea freight). Under such incoterms, main transport costs should not appear in the invoice value.[42]
- After Shipment Incoterms (ASI) consist of 7 incoterms where risks and responsibilities change *after* the main transportation so the exporter is responsible for the main transportation. Under such incoterms main transport costs should appear in the invoice value since the exporter charges them to the importer.[43]

In theory, customs offices should adjust invoice values for transport costs depending on incoterms to estimate the at the border statistical value and thus there should be no effect of incoterms on the value gap.[44]

In practice, customs may make inconsistent value adjustments for a variety of reasons ranging from time-efficiency, "reasonable approximations", or lower incentives to accurately assess the value on the export side. To address this possibility, we define an indicator for *After-shipment incoterm in both countries* which takes the value 1 when both countries declare an ASI, which means the exporter is responsible for main transport costs between France and Madagascar. In a scenario of correct reporting by both France and Madagascar, the same incoterms should be declared for any matched container. However, for 14% of containers in our sample, different incoterm groups (BSI versus ASI) are observed. We define an indicator for containers having conflicting incoterms: *BSI in France but ASI in Madagascar* which means neither party claims being responsible for main transport costs, *ASI in France but BSI in Madagascar* which means both entities claim being responsible for main transport costs.

A1.3. Matching methodology

We detail in this appendix the methodology followed to match French export customs data and Madagascan import customs data that can be broken down into the following 3 main steps:

1. Container ID cleaning and testing

[42] The most common BSI in our sample according to Madagascan declarations is "Free on Board" (FOB) with 41%, followed by "Ex Works" (EXW) with 13%, and "Free Carrier" (FCA) with 3%, while we also observe a few "Free Alongside Ship" (FAS).

[43] The most common ASI in our sample according to Madagascan declarations is "Cost Insurance and Freight" (CIF) with 24%, followed by "Cost and Freight" (CFR) with 15%, "Carriage and Insurance Paid to" (CIP) and "Carriage Paid to" (CPT) with 1% each, while we also observe a few declarations using "Delivered at Place" (DAP), "Delivered at Terminal" (DAT) or "Delivered Duty Paid" (DDP).

[44] French customs are also supposed to adjust the invoice value for other elements beyond transport costs such as commissions, royalty fees, or export subsidies. Madagascan customs are also supposed to adjust the invoice value for other elements such as administered prices. Statistical export values provided by French customs should reflect at the border FOB value after customs adjustment from the invoice value to remove transport costs depending on incoterms. The Madagascan statistical value provided by customs is the CIF after customs adjustment to add transport costs from the invoice value depending on incoterms.

2. Pre-matching data cleaning and aggregation
3. Matching of container-day level data

Step 1: Container ID cleaning and checks We rely on container identification to enable the matching between French and Madagascan customs data. Declarations without a container ID are either non-containerized trade or did not have the container ID variable filled since it is not mandatory to fill it but just common practice. Container IDs are handwritten so cleaning is necessary to correct small mistakes and to check whether they comply with international standards. According to the international system of container identification following the ISO 6346 norm, container IDs should be composed of 4 string characters followed by 7 digits. The first 3 letters refer to the container owner while the 4^{th} letter corresponds to the equipment ID and can only have 3 occurrences: "U" for containers, "J" for detachable freight container-related equipment, or "Z" for container-related trailers and chassis.

Container IDs are first cleaned separately in the French and the Madagascan customs databases. The cleaning tasks range from treating special characters (e.g., a "/" separator appearing sometimes between the 6^{th} and 7^{th} digit), to correcting container IDs with 10 or 12 characters for which we can identify the missing or extra character thanks to another container ID in the database when they are similar, correcting the 4 letters code in case of a container ID in the database sharing the same last 7 digits with a potentially misspelled string code, or eliminating irrelevant characters before or after the 11 characters code. We consider container IDs as similar when they share similar secondary characteristics (close registration dates, same office of registration, same intermediary, same exporter, same importer) while having minimal container ID discrepancies. Containers not following the ISO 6346 norm are mostly shipper-owned and composed of 6 or 7 digits with a dummy prefix added by customs agents such as "XXXX" or "NONE". Since we cannot certify the uniqueness of those container IDs we drop them from the analysis. We also remove containers whose 4^{th} letter was neither "U" nor "J" nor "Z" after the data cleaning described above. All container IDs in the final sample had "U" as the 4th letter, even though some unmatched instances had "J" or "Z". The share of non-missing container IDs not following the ISO 6346 format over 2014-2016 is only 2.08% in French customs data (0.76% do not follow the 4 letters + 7 digits structure while 1.31% have an invalid 4^{th} letter) and 0.42% in Madagascan customs data (0.17% do not follow the 4 letters + 7 digits structure while 0.24% have an invalid 4^{th} letter).

Step 2: Pre-matching data cleaning and aggregation We aim at aggregating French customs data at the container-day level to implement the matching with its Madagascan equivalent over the period 2014-2016. As such, we include French customs declarations starting in 2013 to enable the matching of some export data recorded in late 2013 in France to import data recorded in early 2014 in Madagascar given the time it takes for a container to travel from France to Madagascar. We consider Madagascan import declarations in which France is labeled either as the country of origin or the country of provenance. As we will detail below, we keep container-day level observations for which a single French or Madagascan customs declaration is related to one or multiple containers. However, we enable multiple French declarations re-

lated to a unique container-day, but not multiple Madagascan declarations related to a unique container-day. In case of multiple Madagascan declarations linked to a unique container ID with the same arrival date we suspect multiple Madagascan inspectors could be assigned to the same container and drop them.

Step 2.1 French customs data Aggregation at the container-declaration-HS 6-digit product level: French customs data in its most disaggregated form are defined at the declaration-item level with duplicate observations when the declaration-item level is split between multiple containers. Items are recorded under a TARIC 10-digit code whose first 6 digits are equivalent to the HS 6-digit code. We rely on the HS 6-digit level of aggregation for product codes because it is the level which enables consistent comparisons across countries. Values and weights are recorded at the declaration-item level. For a given declaration-item related to multiple containers, we divide values and weights by the number of containers and use those divided values and divided weights as the correct values and weights for our analysis. Thus, we aggregate items from the same HS 6-digit product which have the same container-declaration and sum their values and weights to obtain a French export database at the container-declaration-HS 6-digit product level.

Aggregation at the container-day-HS 6-digit product level: Multiple declarations are found at the container-day level in French customs data referring to the same shipment. In those cases, we aggregate declarations recorded on the same day regarding the same container ID and sum their values and weights for each product, so that multiple declarations linked to a unique container are not removed (unlike for Madagascan data). This results in a French export database at the container-day-HS 6-digit product level.

Treatment of recurrent container IDs: We aggregate container-day level observations for which the container ID is identical and the time gap between declaration registration dates is small, as it suggests declarations related to the same shipment. Thus, we sum values and weights related to those container-level observations for each product after some checks thanks to secondary observations. The registration date used for the matching with Madagascan data is the last of those ones. We keep observations with the same container ID as distinct when it is not possible to determine whether those observations refer to the same shipment or to distinct shipments traveling at different dates using the same container (which will be assessed using matched Madagascan observations as described below).

Removal of abnormal weights: We rely on the ISO 668:2013 norm (ISO 668 hereafter) which defines the maximum mass of a container under operation at 30,480 kg. It enables varying maximum net weights due to differences in materials used by container builders. Although the empty weight of a container cannot be zero, we apply a low-binding constraint and drop from the sample the two containers for which the container-day level weight is above this threshold of 30,480 kg, as it could signal inaccurate matching or mis-recorded weights. Some containers related to a unique declaration have a weight above 30,480 kg only before being divided by the number of containers related to this declaration, so we do not remove those.

Container-day level and container-day-product HS 6-digits level database: After the previ-

ous steps we end up with a French export database at container-day-product HS 6-digits level. We aggregate products for a unique container-day and sum their values and weights to obtain the French export database at container-day level.

Step 2.2 Madagascan customs data HS product code conversion to the 2012 revision: We identify in the Madagascan customs data a few HS product codes that do not belong to the list of products under the HS 2012 revision[45] (which should be the revision used by customs to record products over the period 2014-2016) but belong instead to the list of products under the HS 2017 revision. Since we have no way of directly disentangling HS 2012 from HS 2017 product codes, we rely on conversion tables from the United Nations Statistics Division (UNSTAT) applying the following strategy:

i. We take all product codes in the Madagascan customs data as if they all belonged to the HS 2012 revision and convert them into HS 1992 revision codes using UNSTAT conversion tables.

ii. For the product codes that were not converted (57 or 2.3% of codes) we try to convert them into HS 1992 revision codes but as if they belonged to the HS 2017 revision using the UNSTAT conversion tables.

iii. We assume that the 49 product codes which could not be converted from the HS 2012 revision to the HS 1992 revision but could be converted from the HS 2017 revision to the HS 1992 revision, belong to the HS 2017 revision.

iv. We convert those 49 product codes from the HS 2017 revision to the HS 2012 revision using the UNSTAT conversion tables.

v. We leave as unchanged the remaining 8 product codes which could neither be converted from the HS 2012 revision nor the HS 2017 revision to the HS 1992 revision.

Aggregation at the container-declaration-HS 6-digit product level: Madagascan customs data are defined at the declaration-item level while container-related information is only available at the declaration level. The item level of product identification is the HS 8-digit level, but we aim at aggregating to the HS 6-digit level. Unlike for French customs we first aggregate items from the same HS 6-digit product which have the same container-declaration and sum their values and weights. Then, we tackle the cases of multiple containers related to a unique declaration by dividing values and weights at the declaration-HS 6-digit product level by the number of containers related to this declaration and use those divided values and divided weights as the correct values and weights in a Madagascan import database at the container-declaration-HS 6-digit product level.

Removal of multiple declarations related to the same container: Unlike for French customs data, we are less comfortable in keeping containers related to multiple Madagascan declarations since they could be assigned to different inspectors. We use the arrival date of the container,

[45] All products declared in France belong to the HS 2012 revision, so no such conversion was needed.

which corresponds to the date at which the shipment arrives at the Madagascan customs office (but is different from the registration date which corresponds to the date when the customs declaration is registered by the broker (or importer) in the ASYCUDA system). We remove from the analysis containers for which multiple declarations are recorded concerning the same container arrival date. We also remove observations in a short time span related to those containers although having a different arrival date (e.g., a couple of declarations recorded in a first Madagascan customs office at a given arrival date, followed by another declaration later in another Madagascan customs office). However, we keep other cases of multiple declarations related to the same container ID.

Treatment of declarations under the transshipment regime: One of the reasons behind keeping multiple Madagascan declarations with the same container ID over a short time span relies on the presence of Madagascan declarations under a transshipment regime. Those declarations refer to goods under national transit which should be followed by other(s) declaration(s) under final import regimes (such as import for home use). To avoid duplicates in those cases, we remove the first declaration under the transshipment regime and keep the following ones with the same container ID while still using the registration date of the first declaration under transshipment regime for the matching with French customs data. For similar patterns involving more than 2 declarations in a short time span related to a unique container ID (e.g., a declaration under warehousing in-between), we keep the last declaration values and weights under final import regimes while using the date of the first declaration under the transshipment regime. We do so when the gap between declaration registration dates is below 90 days, since this threshold provides a trade-off between keeping late final import declarations after the transshipment one, and preventing the association of declarations related to different shipments using the same container ID. The remaining declarations under a transshipment regime are dropped from the sample as they should have been followed by another declaration if they were actual imports by Madagascar that should be considered in our analysis. Moreover, most of those declarations include a product whose HS 6-digit code "980700" refers to "Products under transit with a simplified declaration" which reinforces our choice of removing them from the sample.

Aggregation at the container-day-HS 6-digit product level: Unlike for French data we do not have multiple declarations related to the same container ID registered on the same day as most of those instances in Madagascan customs data have the same container arrival date, so they have already been dropped, as explained previously. Thus, the container-declaration-HS 6-digit product level is equivalent to the container-declaration-day-HS 6-digit product level at this stage.

Treatment of recurrent container IDs: Unlike for French data, we do not aggregate across declarations with the same container ID since we are unable to prove they are part of the same shipment after the removal of declarations under the transshipment regime. The remaining observations with recurrent container IDs would ultimately try to find one of many French customs container-day observations to be matched with.

Removal of abnormal weights: Like for French customs data, we remove the few containers for which the container-day level weight is above 30,480 kg, which is the maximum mass of a

container under operation defined by the ISO 668 norm.

Container-day level and container-day-HS 6-digit product level databases: After the previous steps we end up with a Madagascan import database at container-day-HS 6-digit product level. We aggregate HS 6-digit products for a unique container-day and sum their values and weights to obtain the Madagascan import database at the container-day level.

Step 3: Matching between container-day level databases Matching at the container level: We match French export and Madagascan import customs databases processed in step 2 at the container-day level using the container ID as matching variable and keeping only observations for which the French registration date is earlier than the Madagascan registration date.

Adding the product-level databases: We merge the matched observations with French and Madagascan container-day-HS 6-digit product databases including all products irrespective of whether they are similar across the matched containers or not. If we were matching container-day-HS 6-digit product databases, the resulting sample would exclude matched containers with no identical product codes. The inclusion of all products is crucial for our analysis of product misclassification across matched containers.

Treatment of recurrent container IDs: The remaining observations with recurrent container IDs (with different container-day instances) either from French or Madagascan customs data could not be identified previously as being either multiple declarations related to the same shipment (on the French side), or distinct declarations related to distinct shipments using the same container ID and travelling at different time periods (on both sides). For those observations, we apply the following strategy:

i. We investigate whether instances with multiple container-day level observations from French and Madagascan customs databases could correspond to distinct shipments or whether some of them should be aggregated on the French side. In case of data aggregation, we sum the values and weights while keeping the last French declaration registration date.

ii. We investigate potential matched data and select, when possible, the relevant observations to match and/or the ones to be aggregated on the French side thanks to secondary similar characteristics (exporter, importer, values, weights...) before matching. In case of data aggregation, we sum the values and weights while keeping the last French declaration registration date. The remaining unmatched observations are dropped from the sample.

iii. Regarding container IDs for which we could not determine whether to aggregate container-day observations from French customs data or to keep only one of them (the one that would match with a Madagascan equivalent while dropping the declaration which does not find its Madagascan equivalent), we choose the option which minimizes the weight gap between French and Madagascan customs data as it would decrease the probability of introducing biases in the matching. In case of data aggregation, we also sum the values and weights while keeping the last French declaration registration date.

iv. We remove from the sample the few container-day level observations for which the aggregated weight is above the ISO 668 norm of 30,480 kg either on the French or Madagascan side after aggregation.

Limit of 90 days between registration dates: We investigate the travel duration between France and Madagascar using estimates from searates.com since most of the containers in our matched database travelled by sea. The estimated travel time from Rouen to Toamasina is 21 days and 12 hours assuming the ship would travel through the Suez Canal with an average speed of 13 knots. Rouen is chosen as the baseline French seaport because sea exports originating from Paris mostly travel from Rouen on the Seine to the Channel (as confirmed by a standard Paris-Antananarivo trade route on searates.com). Likewise, Toamasina is chosen as the baseline Madagascan seaport because sea imports arriving at Antananarivo mostly travel from Toamasina to Antananarivo, and Toamasina is the main seaport hub for international trade in Madagascar. For the sake of clarity in the matching methodology, we do not consider alternative thresholds to the 90 days for shipments from the French overseas territories of La Reunion and Mayotte that are close to Madagascar even though shorter travel time is expected.

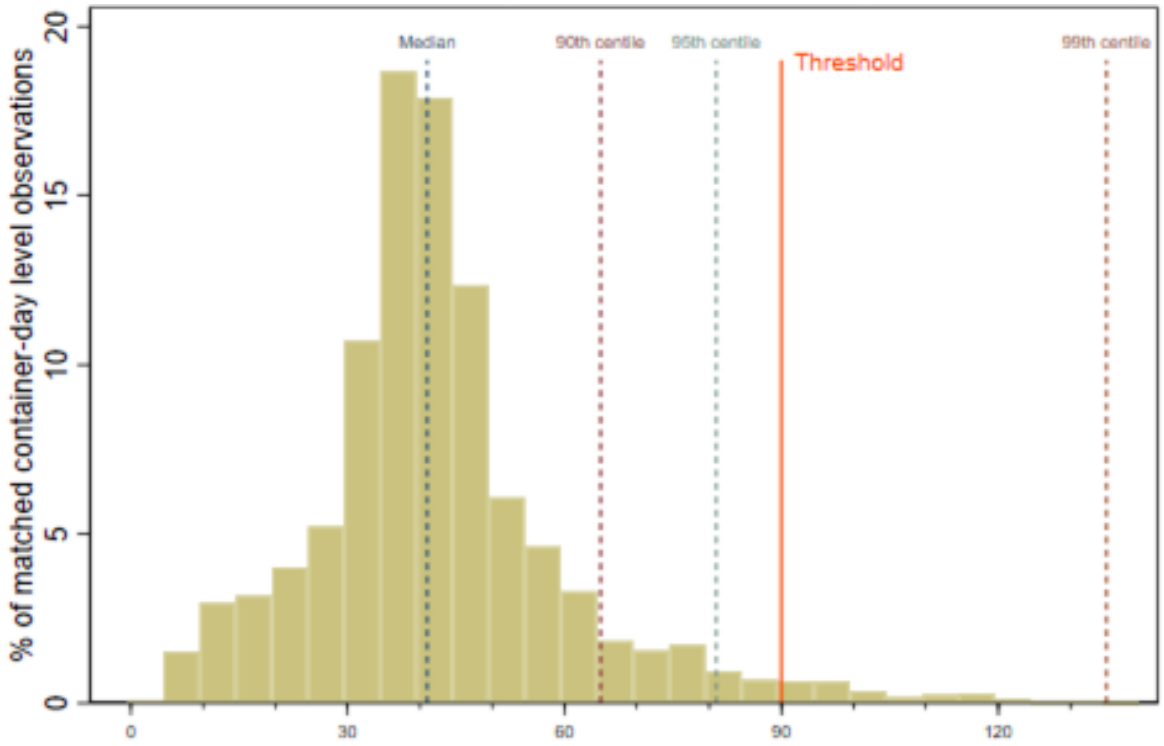
However, we need to account for potential transit time at both customs offices after the French declaration registration date and before the Madagascan declaration registration date. The travel time could also be increased depending on potential stopovers between France and Madagascar, or the use of alternative routes such as travelling though the Cape of Good Hope. After some data analyses using alternative thresholds, we find that 90 days would provide a good trade-off between the selection of the maximum of seemingly identical shipments and the removal of unrelated shipments with a similar container ID. We have 17 containers in the final matched container-day level database (Sample 7 in Appendix Table A1) which appear twice as distinct shipments traveling during different periods.

Since the period covered by Madagascan customs data in our sample was 2014-2016, we consider French customs data from October 3rd, 2013. Appendix Figure A2 shows the distribution of the date registration gap across matched container-day level observations. Most observations have a registration gap around the median of 41 days which is higher than the travel time provided by searates.com due to transit times. The 90-day threshold enables the inclusion of more than 95% of observations while removing only a fraction of observations more likely to refer to distinct shipments. The resulting sample corresponds to sample 5 in Appendix Table A1.

Multiple containers related to a unique declaration after matching: We remove from the sample observations for which all containers related to a unique French declaration could not be matched with a unique container (corresponding to a unique declaration) in Madagascan customs data and vice versa. The same principle applies for the few instances of multiple declarations related to multiple container IDs in French customs data because of the inclusion of multiple declarations related to a unique container.

Removal of small transactions and huge absolute value gaps: We remove from the sample observations whose export or import FOB value is below 1,000€ to limit large swings in value gaps, when considered in their log difference form, originating from those small values. We also remove containers which benefitted from at least one tax exemption in Madagascar. The

Figure A2 – Distribution of the registration gap between declarations (in days)

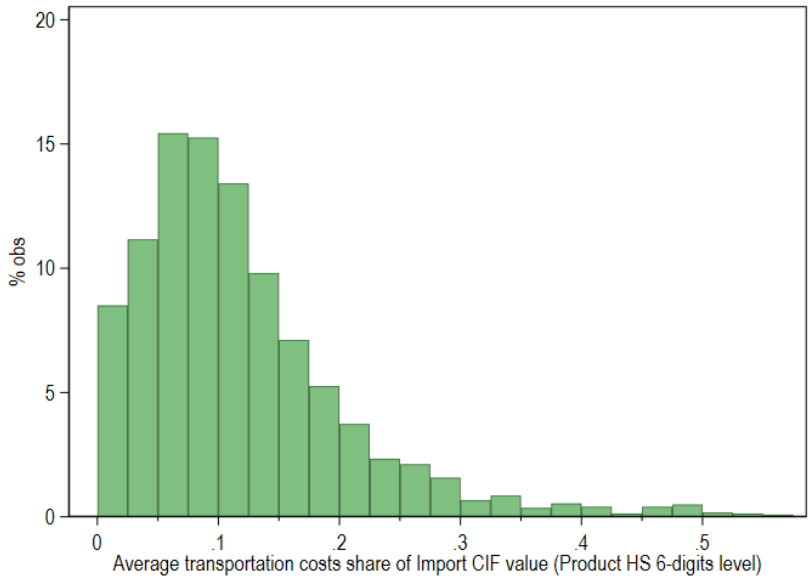


Note: the plot removes observations from the last centile of the distribution.

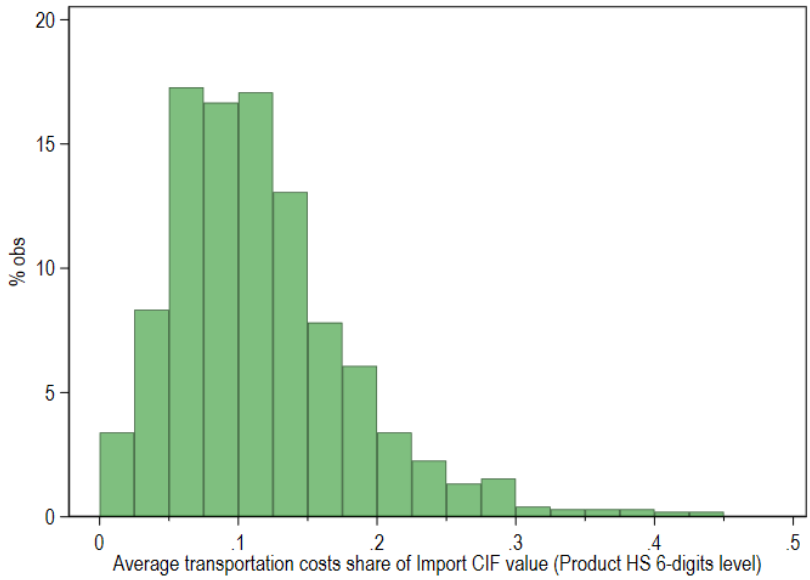
resulting sample corresponds to sample 6 in Appendix Table A1.

Removal of observations for which a hypothetical tariff cannot be computed: To assess the relationship between trade gaps and tariff rates, we rely on a hypothetical tariff rate measure given by the average of tariff rates which should have been used considering the HS 6-digit products reported in France and weighted by their relative export value share. The HS 6-digit product tariff rates used to compute the hypothetical tariff are the median tariffs applied in Madagascar as explained in Section 2.4. Consequently, we remove from the sample containers for which at least one HS 6-digit product reported in the French data could not be matched with the median tariff. Such products may be either those never reported in any Madagascan declarations or those in HS chapter 99 which are specific to the exporting country. The resulting sample corresponds to sample 7 in Appendix Table A1 and is our baseline sample for this study.

Figure A3 – Transport costs as share of imports CIF across HS 6-digit products



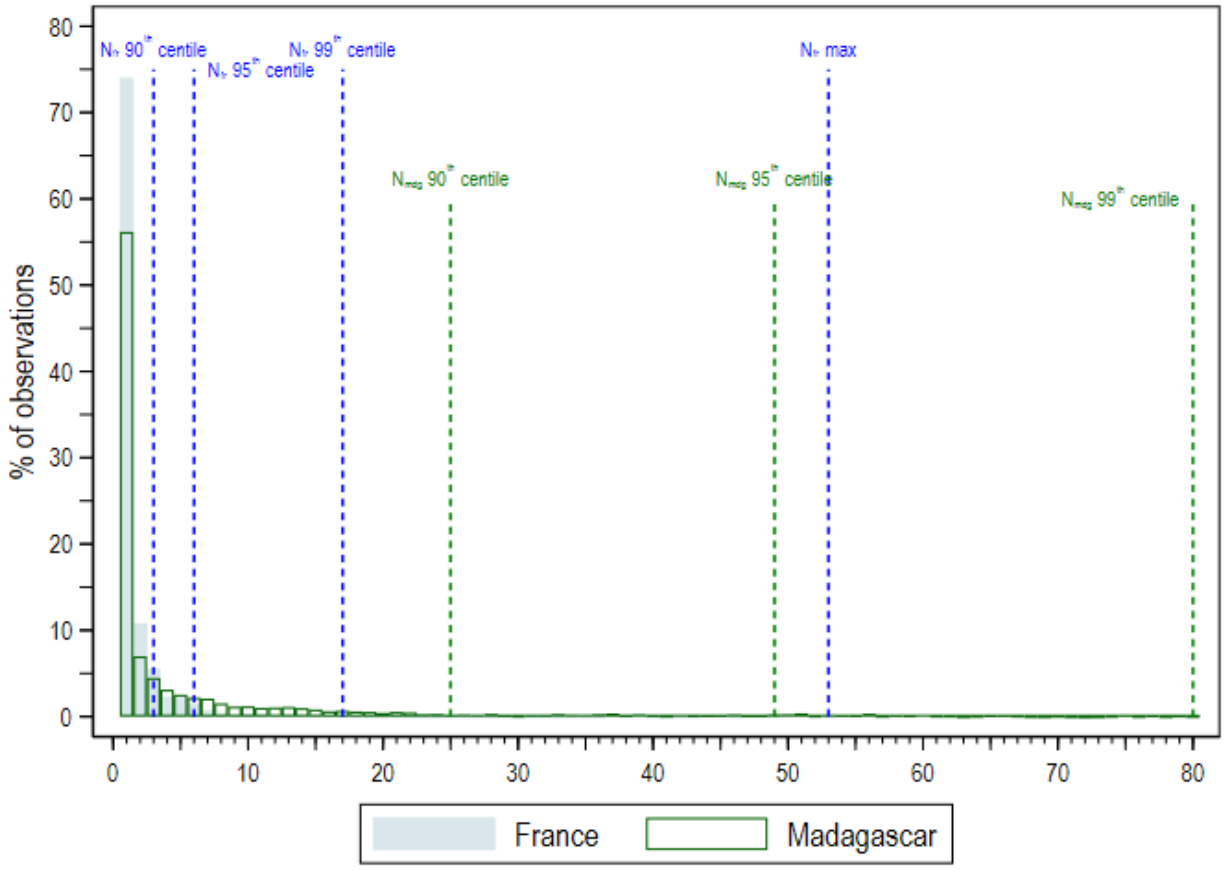
(a) All containers (N=2244)



(b) At least 10 containers (N=981)

Note: the plots exclude observations (products) in the last centile of the distribution of transport costs to import CIF. ratios

Figure A4 – Number of products per container



Notes: N_{fr} is the number of products declared in France and $N_{mdg}$is the number of products declared in Madagascar. The plot excludes the last centile in the distribution of N_{mdg}.

A2. Additional Tables and Figures

1

Table A1 – Representativeness of the matched container sample

Sample	Total	% Total	% Total among non-missing French containers
N. French declarations			
(0) Total exports	74456	100.00%	
(1) Non-missing container ID	16838	22.60%	100.00%
(2) ISO compliant container ID	16380	22.00%	97.30%
(3) Matched container-day obs.	15566	20.90%	92.40%
(4) - with 1 decl. per container	8521	11.40%	50.60%
(5) - arrived within 90 days	8203	11.00%	48.70%
(6) - after removing small transactions	6630	8.90%	39.40%
(7) - with information on tariffs (final sample)	5868	7.90%	34.80%
N. container IDs			
(0) Total exports	13205	100.00%	
(1) With container ID	13205	100.00%	100.00%
(2) With ISO compliant container ID	12929	97.90%	97.90%
(3) Matched container-day obs.	11826	89.60%	89.60%
(4) - with 1 decl. per container	9685	73.30%	73.30%
(5) - arrived within 90 days	9378	71.00%	71.00%
(6) - after removing small transactions	8004	60.60%	60.60%
(7) - with information on tariffs (final sample)	7461	56.50%	56.50%
Export FOB value (million €)			
(0) Total exports	1218.605	100.00%	
(1) With container IDs	337.615	27.70%	100.00%
(2) With ISO compliant IDs	330.862	27.20%	98.00%
(3) Matched container-day obs.	309.6478	25.40%	91.70%
(4) - with 1 decl. per container	238.752	19.60%	70.70%
(5) - arrived within 90 days	234.038	19.20%	69.30%
(6) - without small transactions	193.942	15.90%	57.40%
(7) - with information on tariffs (final sample)	184.348	15.10%	54.60%
Export weight (million kg)			
(0) Total exports	231.431	100.00%	
(1) With container ID	174.838	75.50%	100.00%
(2) With ISO compliant container ID	172.774	74.70%	98.80%
(3) Matched container-day obs.	159.2019	68.80%	91.10%
(4) - with 1 decl. per container	141.800	61.30%	81.10%
(5) - arrived within 90 days	137.944	59.60%	78.90%
(6) - after removing small transactions	125.434	54.20%	71.70%
(7) - with information on tariffs (final sample)	118.63	51.30%	67.90%
Unit Price (€/kg)			
(0) Total exports			5.27
(1) With container ID			1.93
(2) With ISO compliant container ID			1.91
(3) Matched container-day obs.			1.95
(4) - with 1 decl. per container			1.68
(5) - arrived within 90 days			1.70
(6) - after removing small transactions			1.55
(7) - with information on tariffs (final sample)			1.55

Table A2 – Prevalence of discrepancies - by size of the value gap (VG)

Size of gap	Contributions %Total				Prevalence by type of misclassification			
	N	Value Gap	Tariff Revenue Gap	Imports FOB	None (Correct)	Tariff-decreasing	Tariff-neutral	Tariff-increasing
VG $\in (-5\%, 5\%)$	53.3%	0.2%	33.6%	67.0%	57.1%	17.0%	14.7%	11.2%
VG <= -5%	21.2%	-57.4%	-21.5%	17.1%	39.7%	22.5%	16.7%	21.0%
VG > 5%	25.5%	157.3%	87.9%	15.9%	41.0%	21.4%	18.0%	19.5%
Total	100.0%	100.0%	100.0%	100.0%	49.3%	19.3%	16.0%	15.4%

Table A3 – Prevalence of discrepancies - by misclassification

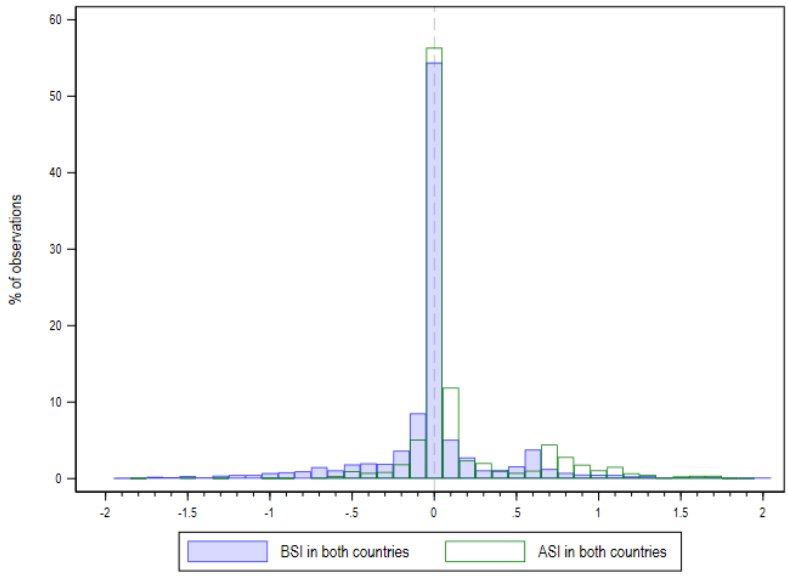
Mislassification	% Total				Value gap (VG)		
	N	Value Gap	Tariff Revenue Gap	Imports FOB	VG $\in (-5\%, 5\%)$	VG <- 5%	VG > 5%
None (correct)	49.3%	17.9%	13.6%	39.4%	61.7%	17.1%	21.2%
Tariff-decreasing	19.3%	25.4%	81.8%	20.5%	47.0%	24.8%	28.2%
Tariff-neutral	16.0%	27.3%	22.1%	23.2%	49.0%	22.3%	8.8%
Tariff-increasing	15.4%	29.4%	-17.5%	17.0%	38.7%	29.0%	32.3%
Total	100.0%	100.0%	100.0%	100.0%	53.3%	21.2%	25.5%

Table A4 – Global evidence: more distinct HS 6-digit products reported by importing countries than exporting countries

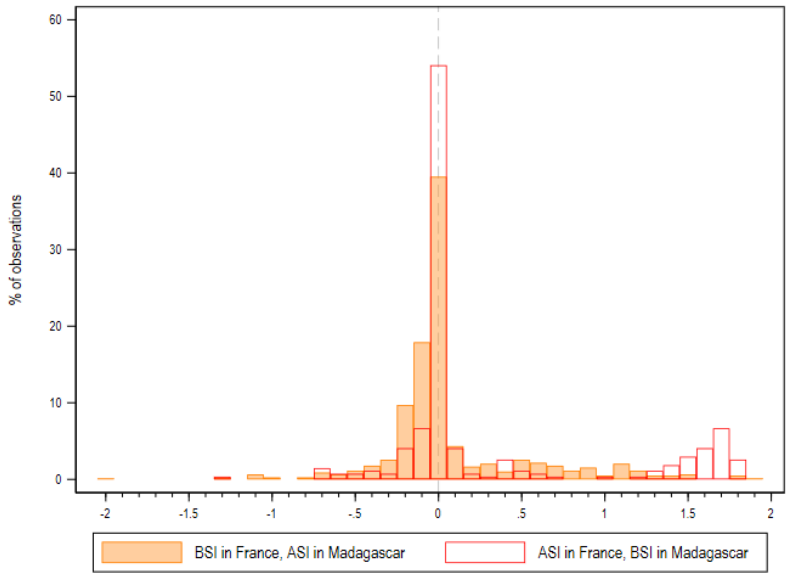
% of exporting country-importing country-year observations with product gap<0 (more distinct HS 6-digit products reported by the importer than by the exporter)						
Year	France as exporting country	Madagascar as importing country	All exporting countries	HIC exporting country to non-HIC importing country	France's rank among exporting countries (2014-2016)	Madagascar's rank among importing countries (2014-2016)
2014	66.00%	88.30%	69.80%	62.40%		
2015	72.40%	83.60%	71.60%	71.00%	90/156	50/156
2016	72.90%	80.00%	71.50%	73.80%		

Notes: computed over 156 countries reporting trade data to COMTRADE as an exporting country and as an importing country over 2014-2016. The product gap refers to the number of distinct HS 6-digit products declared as an exporting country minus the number of distinct HS 6-digit products declared as an importing country. HIC indicates high-income countries and non-HIC all other countries. The rank of France [Madagascar] is in the likelihood to report less distinct products than the importing.

Figure A5 – Distribution of value gap depending on incoterms



(a) Consistent incoterms



(b) Conflicting incoterms

Table A5 – Additional summary statistics

	Mean	S.D.	N	Min	Median	Max
Madagascan declaration features						
Risk score	4.135	3.495	7 454	1	2	10
Physical inspection (red channel)	0.373	0.484	7 461	0	0	1
Fraud record	0.036	0.185	7 461	0	0	1
Scan exam	0.253	0.434	7 461	0	0	1
French declaration features						
Absence of EORI number	0.099	0.299	7 461	0	0	1
French multiproduct shipment	0.26	0.439	7 461	0	0	1
Incoterms declared						
After shipment incoterm in both countries	0.326	0.469	7 385	0	0	1
BSI in France but ASI in Madagascar	0.037	0.188	7 385	0	0	1
ASI in France but BSI in Madagascar	0.107	0.309	7 385	0	0	1
Importer categories						
Importer under PAD program (PAD)	0.248	0.432	7 461	0	0	1
Firm with government participation (SOE)	0.05	0.218	7 461	0	0	1
Free zone enterprise (FZE)	0.168	0.374	7 461	0	0	1
Unregistered importer	0.046	0.209	7 461	0	0	1
Importer features						
Importer size - ln(sum of import values)	13.991	2.069	7 119	6.923	14.736	16.33
Importer-exporter	0.573	0.495	7 119	0	1	1
New importer	0.044	0.205	7 119	0	0	1
Young importer	0.142	0.35	7 119	0	0	1
Old importer	0.814	0.389	7 119	0	1	1

Table A6 – Determinants of value gaps: alternative specifications

	Value gap (diff-ln)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hypothetical tariff rate	-0.087	-0.065	-0.077	-0.020	-0.019	-0.087	-0.087
	(0.286)	(0.283)	(0.287)	(0.293)	(0.287)	(0.286)	(0.286)
After shipment incoterm in both countries	0.086***	0.084***	0.087***	0.086***	0.084**	0.086***	0.086***
	(0.033)	(0.032)	(0.033)	(0.033)	(0.035)	(0.033)	(0.033)
BSI in France but ASI in MDG	0.029	0.028	0.028	0.026	0.050	0.029	0.029
	(0.038)	(0.038)	(0.038)	(0.038)	(0.046)	(0.038)	(0.038)
ASI in France but BSI in MDG	0.152***	0.152**	0.153***	0.152**	0.136**	0.152***	0.152***
	(0.059)	(0.059)	(0.059)	(0.059)	(0.056)	(0.059)	(0.059)
Hypothetical tariff rate * After shipment incoterm in both countries							-0.129
							(0.328)
Hypothetical tariff rate * BSI in France but ASI in MDG							-0.027
							(0.465)
Hypothetical tariff rate * ASI in France but BSI in MDG							0.377
							(0.703)
Larger export value	0.094**	0.094**	0.095**	0.091**	0.086**	0.094**	0.094**
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
Hypothetical tariff rate * Larger export value	1.419***	1.428***	1.418***	1.377***	1.386***	1.419***	1.419***
	(0.373)	(0.374)	(0.372)	(0.382)	(0.367)	(0.373)	(0.373)
Multiproduct shipment (Fr)	0.047	0.053	0.047	0.050	0.061	0.047	0.047
	(0.043)	(0.043)	(0.043)	(0.042)	(0.046)	(0.043)	(0.043)
Hypothetical tariff rate * Multiproduct shipment (Fr)	0.109	0.082	0.110	0.076	-0.053	0.109	0.109
	(0.303)	(0.301)	(0.303)	(0.302)	(0.317)	(0.303)	(0.303)
SOE	-0.034	-0.029	-0.035	-0.028	-0.035	-0.034	-0.034
	(0.049)	(0.049)	(0.049)	(0.049)	(0.051)	(0.049)	(0.049)
FZE	-0.099	-0.091	-0.098	-0.109	-0.081	-0.099	-0.099
	(0.070)	(0.071)	(0.071)	(0.071)	(0.080)	(0.070)	(0.070)
Unregistered importer	-0.085	-0.086	-0.083	-0.092	-0.100	-0.085	-0.085
	(0.085)	(0.085)	(0.085)	(0.084)	(0.089)	(0.085)	(0.085)
PAD	-0.013	-0.011	-0.012	-0.012	-0.018	-0.013	-0.013
	(0.043)	(0.042)	(0.044)	(0.043)	(0.038)	(0.043)	(0.043)
Hypothetical tariff rate * SOE	-1.107**	-1.119**	-1.115**	-1.136**	-1.248**	-1.107**	-1.107**
	(0.518)	(0.521)	(0.513)	(0.515)	(0.555)	(0.518)	(0.518)
Hypothetical tariff rate * Unregistered importer	1.278**	1.289**	1.276**	1.254**	1.337**	1.278**	1.278**
	(0.580)	(0.581)	(0.580)	(0.577)	(0.596)	(0.580)	(0.580)
Hypothetical tariff rate * PAD	0.018	-0.007	0.015	-0.008	0.010	0.018	0.018
	(0.312)	(0.312)	(0.316)	(0.320)	(0.297)	(0.312)	(0.312)
Non-commercial exporter	-0.365***	-0.363***	-0.364***	-0.337***	-0.294***	-0.365***	-0.365***
	(0.115)	(0.115)	(0.115)	(0.120)	(0.107)	(0.115)	(0.115)
Hypothetical tariff rate * Non-commercial exporter	0.228	0.233	0.229	0.068	0.203	0.228	0.228
	(0.745)	(0.745)	(0.745)	(0.762)	(0.723)	(0.745)	(0.745)
Non tariff measure		-0.035					
		(0.026)					
Minimum value			0.016				
			(0.032)				
Used good				-0.137**			
				(0.065)			
Importer size - ln(sum(M))						-0.017	
						(0.011)	
Importer-Exporter						-0.100***	
						(0.037)	
New importer						-0.105*	
						(0.054)	
Young importer						-0.048	
						(0.038)	
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
French and Madagascan HS 2-digit product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Madagascar customs office FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
French customs office FE	No	No	No	No	Yes	No	No
Inspector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Broker FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	7385	7384	7385	7385	7385	7385	7385
R^2	0.489	0.490	0.489	0.491	0.508	0.489	0.489
Adjusted R^2	0.460	0.460	0.460	0.462	0.476	0.460	0.460

Notes: The table presents OLS estimates of the determinants of value gaps. Standard errors are clustered by importerand presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. FE stands for fixed effects.

Table A7 – Determinants of alternative measures of gaps

	Value gap			**Weight gap**	**Unit price gap**
		alternative exchange rate	no transport cost adjustment		
	(1)	(2)	(3)	(4)	(5)
Hypothetical tariff rate	-0.087	-0.095	0.187	-0.333**	0.246
	(0.286)	(0.286)	(0.272)	(0.156)	(0.266)
After shipment incoterm in both countries	0.086***	0.086***	0.058**	0.025*	0.061*
	(0.033)	(0.033)	(0.029)	(0.014)	(0.032)
BSI in France but ASI in MDG	0.029	0.027	0.007	0.006	0.022
	(0.038)	(0.038)	(0.035)	(0.030)	(0.044)
ASI in France but BSI in MDG	0.152***	0.152***	0.155***	0.027	0.125**
	(0.059)	(0.059)	(0.057)	(0.019)	(0.057)
Larger export value	0.094**	0.093**	0.249***	-0.017	0.111***
	(0.040)	(0.040)	(0.041)	(0.021)	(0.039)
Hypothetical tariff rate * Larger export value	1.419***	1.422***	0.900***	0.187	1.232***
	(0.373)	(0.374)	(0.344)	(0.158)	(0.321)
Multiproduct shipment (Fr)	0.047	0.044	0.040	0.061*	-0.014
	(0.043)	(0.043)	(0.041)	(0.034)	(0.032)
Hypothetical tariff rate * Multiproduct shipment (Fr)	0.109	0.129	0.159	-0.178	0.287
	(0.303)	(0.302)	(0.284)	(0.207)	(0.236)
SOE	-0.034	-0.027	-0.040	0.000	-0.034
	(0.049)	(0.050)	(0.056)	(0.034)	(0.057)
FZE	-0.099	-0.093	-0.030	-0.010	-0.089
	(0.070)	(0.070)	(0.057)	(0.036)	(0.057)
Unregistered importer	-0.085	-0.086	-0.069	-0.014	-0.071
	(0.085)	(0.085)	(0.074)	(0.048)	(0.078)
PAD	-0.013	-0.013	-0.040	-0.038	0.025
	(0.043)	(0.043)	(0.039)	(0.033)	(0.045)
Hypothetical tariff rate * SOE	-1.107**	-1.144**	-1.062*	0.099	-1.206**
	(0.518)	(0.513)	(0.550)	(0.153)	(0.518)
Hypothetical tariff rate * Unregistered importer	1.278**	1.279**	0.952*	0.462	0.816
	(0.580)	(0.579)	(0.565)	(0.372)	(0.606)
Hypothetical tariff rate * PAD	0.018	0.016	0.189	0.072	-0.054
	(0.312)	(0.313)	(0.287)	(0.186)	(0.334)
Non-commercial exporter	-0.365***	-0.366***	-0.318***	-0.136**	-0.229*
	(0.115)	(0.115)	(0.116)	(0.056)	(0.124)
Hypothetical tariff rate * Non-commercial exporter	0.228	0.226	-0.048	0.886**	-0.657
	(0.745)	(0.745)	(0.767)	(0.381)	(0.772)
Month-year FE	Yes	Yes	Yes	Yes	Yes
French and Madagascan HS 2-digit product FE	Yes	Yes	Yes	Yes	Yes
Madagascar customs office FE	Yes	Yes	Yes	Yes	Yes
Inspector FE	Yes	Yes	Yes	Yes	Yes
Broker FE	Yes	Yes	Yes	Yes	Yes
N	7385	7385	7385	7385	7385
R^2	0.489	0.489	0.527	0.210	0.443
Adjusted R^2	0.460	0.459	0.499	0.164	0.411

Notes: The table presents OLS estimates of the determinants of value gaps in columns 1 through 3, weight gaps in coulmn 4 and unit price gap in column 5. Standard errors are clustered by importer and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. FE stands for fixed effects.

Table A8 – Determinants of value gaps: alternative samples

	Value gap							
Sample	Baseline	1 French declaration per container	Without casual traders	Without conflicting incoterms	Small weight gap	Correct classification	Strategic misclassification	Non strategic misclassification
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hypothetical tariff rate	-0.087	-0.221	-0.007	-0.262	-0.153	-0.479	1.127	0.740
	(0.286)	(0.298)	(0.304)	(0.280)	(0.315)	(0.363)	(0.941)	(0.634)
After shipment incoterm in both countries	0.086***	0.059**	0.088***	0.055*	0.058*	0.020	0.211***	0.135***
	(0.033)	(0.028)	(0.032)	(0.031)	(0.033)	(0.028)	(0.057)	(0.048)
BSI in France but ASI in MDG	0.029	0.002	0.015		-0.098**	-0.118***	0.201**	0.048
	(0.038)	(0.035)	(0.037)		(0.038)	(0.034)	(0.079)	(0.046)
ASI in France but BSI in MDG	0.152***	0.140**	0.154**		0.050	0.009	0.048	0.363***
	(0.059)	(0.059)	(0.060)		(0.041)	(0.034)	(0.080)	(0.130)
SOE	-0.034	-0.017	-0.040	-0.008	-0.113*	-0.024	0.389	-0.108
	(0.049)	(0.042)	(0.054)	(0.050)	(0.066)	(0.057)	(0.276)	(0.113)
FZE	-0.099	-0.091	-0.060	-0.018	-0.231*	-0.141	0.000	0.007
	(0.070)	(0.059)	(0.072)	(0.060)	(0.130)	(0.115)	(0.000)	(0.103)
Unregistered importer	-0.085	-0.231***		-0.118	-0.086	-0.968***	-0.475**	0.053
	(0.085)	(0.070)		(0.088)	(0.088)	(0.217)	(0.212)	(0.129)
PAD	-0.013	-0.032	-0.017	-0.022	-0.032	-0.079*	0.157	-0.074
	(0.043)	(0.038)	(0.042)	(0.038)	(0.041)	(0.044)	(0.164)	(0.071)
Hypothetical tariff rate * SOE	-1.107**	-1.143***	-1.164**	-0.991**	-0.018	-1.305**	-3.374**	-1.382
	(0.518)	(0.433)	(0.586)	(0.419)	(0.684)	(0.612)	(1.644)	(1.020)
Hypothetical tariff rate * Unregistered importer	1.278**	1.977***		1.392**	0.875	4.693***	2.922**	0.928
	(0.580)	(0.557)		(0.620)	(0.748)	(1.293)	(1.404)	(1.124)
Hypothetical tariff rate * PAD	0.018	0.170	-0.000	0.182	0.429	0.893***	-0.662	-1.459***
	(0.312)	(0.297)	(0.303)	(0.270)	(0.282)	(0.270)	(0.930)	(0.492)
Non-commercial exporter	-0.365***	-0.208*		-0.362***	-0.365***	-0.153	-0.362	-0.135
	(0.115)	(0.120)		(0.127)	(0.131)	(0.279)	(0.286)	(0.131)
Hypothetical tariff rate * Non-commercial exporter	0.228	-0.652		0.365	0.207	0.460	-1.224	-0.909
	(0.745)	(0.882)		(0.761)	(0.883)	(1.879)	(1.752)	(1.249)
Larger export value	0.094**	0.066*	0.080**	0.044	0.069	-0.056	0.370*	0.219**
	(0.040)	(0.039)	(0.039)	(0.040)	(0.070)	(0.045)	(0.214)	(0.103)
Hypothetical tariff rate * Larger export value	1.419***	1.514***	1.340***	1.692***	1.831***	2.470***	-0.226	1.061*
	(0.373)	(0.432)	(0.422)	(0.429)	(0.621)	(0.660)	(1.165)	(0.638)
Multiproduct shipment (Fr)	0.047	0.025	0.051	0.026	0.033	0.055	-0.186	0.053
	(0.043)	(0.036)	(0.048)	(0.046)	(0.060)	(0.072)	(0.188)	(0.062)
Hypothetical tariff rate * Multiproduct shipment (Fr)	0.109	-0.143	-0.174	0.157	-0.247	-1.290**	2.020**	-0.132
	(0.303)	(0.281)	(0.315)	(0.341)	(0.442)	(0.550)	(1.001)	(0.438)
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
French and Madagascan HS 2-digit product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MDG customs office FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Inspector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Broker FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	7385	6832	6631	6325	4260	3651	1416	2318
r2	0.489	0.487	0.530	0.475	0.557	0.540	0.599	0.639
r2_a	0.460	0.455	0.501	0.440	0.517	0.505	0.479	0.568

Notes: The table presents OLS estimates of the determinants of value gaps. Standard errors are clustered by importer and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. FE stands for fixed effects.

Table A9 – Value gaps and corruption risk

	Value gap								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Hypothetical tariff rate	1.301***	1.111***	1.051***	1.278***	1.493***	0.944**	0.870**	0.913**	0.788*
	(0.265)	(0.267)	(0.265)	(0.260)	(0.286)	(0.426)	(0.421)	(0.450)	(0.452)
Sign. excess interaction	0.259***	0.110	-0.221	-0.331**	-0.285*	-0.290*	-0.249*	-0.259*	-0.235*
	(0.089)	(0.067)	(0.176)	(0.155)	(0.150)	(0.151)	(0.138)	(0.140)	(0.139)
Hypothetical tariff rate * Sign. excess interaction			2.159*	2.829**	2.407**	2.468**	2.075**	2.151**	1.950**
			(1.195)	(1.093)	(1.058)	(1.056)	(0.972)	(0.962)	(0.952)
After shipment incoterm in both countries				0.209***	0.140***	0.132***	0.128***	0.134***	0.114***
				(0.033)	(0.034)	(0.034)	(0.035)	(0.035)	(0.035)
BSI in France but ASI in MDG				0.171***	0.128***	0.129***	0.109**	0.105**	0.127**
				(0.052)	(0.047)	(0.047)	(0.050)	(0.050)	(0.051)
ASI in France but BSI in MDG				0.260**	0.221*	0.216*	0.205*	0.211*	0.193*
				(0.120)	(0.115)	(0.114)	(0.113)	(0.113)	(0.111)
Larger export value					0.298***	0.173**	0.163**	0.181**	0.175**
					(0.045)	(0.076)	(0.074)	(0.075)	(0.076)
Hypothetical tariff rate * Larger export value						0.953*	0.923*	0.812	0.860*
						(0.488)	(0.482)	(0.500)	(0.493)
Multiproduct shipment (Fr)					0.046	0.053	0.037	0.029	0.016
					(0.041)	(0.061)	(0.061)	(0.063)	(0.063)
Hypothetical tariff rate * Multiproduct shipment (Fr)						-0.063	0.111	0.206	0.339
						(0.380)	(0.389)	(0.400)	(0.384)
FZE							0.125	0.124	0.139
							(0.096)	(0.097)	(0.088)
SOE							-0.099**	-0.045	-0.050
							(0.040)	(0.073)	(0.073)
Unregistered importer							0.059	-0.012	0.033
							(0.047)	(0.113)	(0.113)
PAD							-0.014	-0.127	-0.143
							(0.041)	(0.081)	(0.091)
Hypothetical tariff rate * SOE								-0.497	-0.231
								(0.643)	(0.644)
Hypothetical tariff rate * Unregistered importer								0.585	0.277
								(0.929)	(0.943)
Hypothetical tariff rate * PAD								0.757*	1.034**
								(0.425)	(0.441)
Non-commercial exporter							-0.283***	-0.177	-0.200
							(0.078)	(0.157)	(0.157)
Hypothetical tariff rate * Non-commercial exporter								-0.879	-0.801
								(1.105)	(1.095)
Risk score									0.015***
									(0.004)
Red channel									-0.027
									(0.036)
Scanned									0.055
									(0.039)
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
French and Madagascan HS 2-digit product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Inspector FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Broker FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. obs.	2502	2502	2502	2475	2475	2475	2475	2475	2473
R^2	0.518	0.602	0.603	0.623	0.649	0.650	0.658	0.659	0.663
Adjusted R^2	0.482	0.557	0.558	0.579	0.607	0.609	0.617	0.617	0.621

Notes: The table presents OLS estimates of the determinants of value gaps. Standard errors are clustered by importer and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. FE stands for fixed effects. The sample is smaller than in the baseline results due to the availability of the corruption proxy that is subject to some restrictions described in Chalendard, Fernandes, Raballand, and Rijkers (2023).