Use of internal information, external information acquisition and customs underreporting

Cyril Chalendard

CENTRE D'ETUDES
ET DE RECHERCHES
SUR LE DEVELOPPEMENT
INTERNATIONAL

SÉRIE ÉTUDES ET DOCUMENTS

Use of Internal Information, External Information Acquisition and Customs Underreporting

Cyril Chalendard

Études et Documents n° 22

September 2015

To cite this document:

Chalendard C. (2015) "Use of Internal Information, External Information Acquisition and Customs Underreporting", *Études et Documents,* n° 22, CERDI.
http://cerdi.org/production/show/id/1722/type_production_id/1

CERDI
65 BD. F. MITTERRAND
63000 CLERMONT FERRAND – FRANCE
TEL. + 33 4 73 17 74 00
FAX + 33 4 73 17 74 28
www.cerdi.org

The author

Cyril Chalendard
PhD Student
Clermont Université, Université d'Auvergne, CNRS, UMR 6587, CERDI, F-63009 Clermont Fd
Email : cyril.chalendard@udamail.fr

This work was supported by the LABEX IDGM+ (ANR-10-LABX-14-01) within the program "Investissements d'Avenir" operated by the French National Research Agency (ANR).

Études et Documents are available online at: http://www.cerdi.org/ed

Director of Publication: Vianney Dequiedt
Editor: Catherine Araujo Bonjean
Publisher: Mariannick Cornec
ISSN: 2114 - 7957

Abstract

This paper identifies opportunities for improving the performance of revenue-collection authorities. To detect and combat fraud, we argue that revenue-collection authorities should, notably in the absence of reliable third-party information, exploit non-usual sources of information. Specifically, our micro-level study of customs evasion provides evidence that using internal or external available sources of information facilitates customs-enforcement. Estimates highlight that exploiting historical data and/or relying on an information provider — a pre-shipment inspection company —significantly reduces observed trade statistics discrepancies. The potential endogeneity of pre-shipment inspections is addressed by using instrumental variables. Robust to a variety of additional checks, findings notably suggest that governments should work on the introduction of an automatic exchange of information system across customs authorities to substantially offset the informational advantage of the importer.

Key words

Use of internal information, External Information acquisition, Customs enforcement, Tax evasion, Pre-shipment inspections.

JEL codes

H26, H83, K42, F13, O17

Acknowledgment

Supports from the International Centre for Tax and Development and the Programme d'Investissements d'Avenir of the French government (ANR-10-LABX-14-01) are gratefully acknowledged. I also wish to thank Thomas Cantens, Céline Carrère, Vianney Dequiedt, Anne-Marie Geourjon and Bertand Laporte (alphabetical order) for their useful comments/discussions and participants at the 10th Annual World Customs Organization Picard Conference. The responsibility for all interpretations and conclusions expressed in this article lies solely with the author.

1 Introduction

To offset the informational advantage of the taxpayer, revenue-collection authorities are continually seeking quality information. Regarding customs, paper trails do not really help in practice the officer as supporting documents can be easily forged (through for instance the falsification of the commercial invoice . . .)[1]. Yet, the authority can rely on other (mainly two) information sources in order to estimate quite accurately the transaction value. First, customs may infer the value from previous import declarations. In addition to relying on this internal information source, customs can also rely on an external information source. Specifically, through the implementation of pre-shipment inspection (henceforth PSI) programs, the government may contract with a private information provider — namely a PSI company—.

This paper presents the first empirical evidence that, by reducing the information asymmetry between the importer and customs, exploiting internal information and acquiring external information facilitate customs-enforcement. Results stress that using historical data and/or relying on an information provider significantly reduce evasion. We interpret the lack of complementarity/substitutability between these two information sources as evidence that the importer fully reacts to the information disclosed by the the PSI company (legal practice). Besides, as in Javorcik and Narciso (2008), we find that enforcement is product-varying. Finally, estimates confirm that higher taxes trigger more customs evasion.

Our contribution stresses that a promising measure to combat fraud would consist in setting-up an automatic exchange of information system across customs authorities. In the short term, we outline that revenue-collection authorities should first improve the performance of their risk analysis system through a systematic and full exploitation of the informational component of historical data.

In this study, we use a customs database with unusually detailed information on import transactions (see section 4) from a representative Sub-Saharan Africa customs authority. The country's name having to be kept confidential, we will name it country A. As in many African countries, the government's budget of country A continues to rely heavily on revenue from border collection. Customs evasion remains also, as in many developing countries, a highly topical issue (see e.g Zake (2011), Montagnat-Rentier and Parent (2012)). As the assessment of the customs value is still problematic, the government continues to rely on the PSI services of a private company.

Following Bhagwati (1964) 's approach, we proxy mis-reporting by discrepancies in mirror trade statistics (*i.e.*, the difference between the export value and the import value). Each observation is at the HS-6 product-trade partner level. Our key variables of interest are (i) the number of previous import declarations and (ii) the pre-shipment inspection frequency. While the first aimed at proxying the internal information available, the latter proxies the

[1] On the relationship between information trails and enforcement see *i.a.* Pomeranz (2015), Almunia and Lopez-Rodriguez (2013), Bachas and Jensen (2014).

external information acquisition. Consistently with the argument that customs officers infer the transaction value of a specific product from the declared value of similar products previously imported, estimates indicate that a one percentage increase in the number of previous import declarations leads to a reduction of approximately 0.4% of underreporting. Our analysis further indicates that a one percentage point increase in pre-shipment inspections results in a 0.7% decline in observed trade statistics discrepancies. The interaction variable between the quantity of internal information available (*i.e.*, the number of previous import declarations) and the amount of external information acquired (*i.e.,* the pre-shipment inspection frequency) is not significant. This lack of impact supports the hypothesis that the importer fully adjusts its customs declaration to the value estimated by the PSI company (legal practice). To address the potential endogeneity issue of pre-shipment inspections —coefficients may be biased toward zero—, we use instrumental variable (IV) techniques. Specifically, we use the fraction of import operations handled by the main customs clearance office — termed office #1— to instrument the private inspection frequency. This instrument is positively correlated to the suspect endogenous variable as the selection of the customs clearance office and the probability that a private inspection will be carry out have in common some specific criteria (*i.a.*, type of goods, containerised cargo). It is noteworthy that enforcement is homogeneous between customs offices (except landborder offices, a robustness check addresses this issue). Therefore, it is unlikely that the fact that imported goods are released by the customs office #1 have a specific impact on undervaluation. Our instrument is highly significant in the first stage regression. Even if IV estimations confirm the bias (the coefficient is double), other results remain largely unchanged. Results also hold to the inclusion of trade partner dummies, HS-2 industry dummies or HS-2 industry-trade partner dummies. Finally, findings are robust to a series of additional robustness checks such as using an alternative measure of underreporting, using an alternative measure of the quantity of internal information available, numerous changes in the sample size ...

This work contributes to the emerging literature on customs compliance which applies mainly theoretical findings from tax compliance. The benchmark economic approach of modelling tax-non compliance has been pioneered by Allingham and Sandmo (1972) (henceforth A-S). The taxpayer evades if the current benefits of doing so exceeds the expected costs (*i.e.,* tax adjustement and fine). Compliance arises because of the fear of detection and punishment. More broadly, the literature on tax compliance[2] outlines that increasing enforcement will reduce the amount of evaded income. Some theoretical studies have enriched the traditional framework by integrating the fact that enforcement may vary. The detection probability is thus endogenized. The A-S. contribution considers the simple case where the probability of detection varies with the amount reported. In the context of corporate tax evasion, Almunia and Lopez-Rodriguez (2013) consider enforcement as a non-linear function of reported revenue. Enforcement also depends on the type of income. The most recent influential contribution —Kleven, Knudsen, Kreiner, Pedersen, and Saez (2011)— shows that third-party reporting is an effective way to reduce evasion. Nevertheless, in response to their inability to cheat on third-party reported incomes, Carrillo, Pomeranz, and Sing-

[2]For a review, see among others, Andreoni, Erard, and Feinstein (1998), Slemrod and Yitzhaki (2002), Sandmo (2005) and Saez, Slemrod, and Giertz (2012).

hal (2014) outline that firms may find some alternative tax-evasion methods. In particular, under a weak enforcement environment, they show that using third-party information may ultimately have no (positive) impact on corporate tax revenues because of adjustments on less verifiable costs.

Overall, the empirical literature on customs compliance supports theoretical findings: estimates stress that enforcement reduces evasion. Typically, customs evasion is proxied by discrepancies in mirror trade statistics[3]. Evasion appears to be greater in countries where the rule of law is limited (see Jean and Mitaritonna (2010)) and the level of corruption is high (see Fisman and Wei (2009)). Fisman and Wei (2004)'s results are consistent with the conjecture that enforcement is invariant to the tax rate. The role of entrepot trade in facilitating evasion has been empirically studied (Fisman, Moustakerski, and Wei (2008)). As for Sub-Saharan Africa, the role of Benin, Togo and the Gambia for unofficial transit trade has been stressed (Golub (2012)). Customs evasion may also be facilitated by international networks (Rotunno and Vézina (2012)). Supporting the argument that an efficient border management highly depends on border point characteristics (e.g. existence of suitable secure infrastructure, advanced computerization), Mishra, Subramanian, and Topalova (2008) find that the evasion elasticity varies by the mode of entry of goods (airport, seaport or landport). Finally, it has been largely stressed that enforcement depends also on intrinsic characteristics of products (see for e.g. Javorcik and Narciso (2008)).

The remainder of the article is structured as follows. Section 2 presents available information sources. Section 3 describes our measure of underreporting and our empirical approach. Section 4 describes the data and provides some descriptive statistics. Section 5 presents and discusses the results and includes several robustness checks. Concluding remarks are offered in section 6.

2 Presentation of information sources

While firms have private information on imports value, the customs authority is uninformed about the true value. Therefore, besides (falsified or not) supporting documentation required (commercial invoice, bill of lading . . .), customs seek additional reliable information in order to determine the truthfulness of the declaration. Basically, customs can rely on two additional information sources— one internal and one external.

[3]The existing empirical literature on customs compliance has mainly attempted to estimate how customs compliance responds to taxes. Higher taxes are associated with more evasion. In the seminal work of Fisman and Wei (2004), estimates outline that the semi-elasticity of evasion increases with the tax rate for China. Specifically, they find that a one percentage point increase in tariffs rises evasion by about 3 %. Customs evasion has also been empirically studied for India (Mishra, Subramanian, and Topalova (2008)), North America (see Stoyanov (2012)), Eastern Europe (Javorcik and Narciso (2008), Sub-Saharan Africa (Van Dunem and Arndt (2009), Bouet and Roy (2012), Levin and Widell (2014)) . . . Finally, note that the literature also highlights that, in some specific cases, underreporting of exports may occur. For instance, Ferrantino, Liu, and Wang (2012) find evidence for underreporting of exports at Chinese border to avoid paying the value added tax.

2.1 Use of historical data: an internal information source

Examining the impact of alternative audit rules (relative *versus* fixed), Bayer and Cowell (2009) theoretically show that a response to the hidden information problem consists in using free-information provided by tax returns of competing firms[4]. Basically, this finding indicates that, when a firm makes an extremely low report compared with reports of competing firms, the tax authority should have serious doubts about the truthfulness of the tax return. Further investigations should then be conducted. Adapted to our specific context, this contribution suggests that exploiting information contained in "competing" import declarations — *i.e.*, previously registered import declarations— may be helpful in assessing the transaction value of a new import. If historical data have an informational component, then exploiting this internal source of information would facilitate customs-enforcement. The informational advantage of the importer is then partially offset[5]. Notice that, since declarations are secrets, the importer has no information on reports made by other importers. Therefore, apart from colluding at the declaration stage[6], importers are not able to internalize these (potential) informational externalities.

In Sub-Saharan Africa, to the best of our knowledge no customs authority currently exploits its internal information — historical data— in a systematic and formal way. In particular, no risk analysis system compares new import declarations to the previous. Nevertheless, as argued below, on the field, customs officers may make efficient use of information provided by previous import declarations.

Customs officers have to determine the import declaration compliance. To carry out this difficult task, they may, in addition to analysing supporting documents, draw on their knowledge/prior experience to issue an opinion on the compliance. In such a situation, the customs officer actually uses the information that he has learned from similar import declarations to better assess tax and duties payable. It would imply that the more a product is frequently imported, the more the product becomes difficult to underreport. In other words, enforcement may increase with the quantity of internal information available. The empirical analysis will try to address this specific question.

2.2 Pre-shipment inspection services: an external information source

Background on PSI programs. Over fifty countries have experienced pre-shipment inspection programs since mid-twentieth century. PSI mainly consists of a set of verification services to assist customs authorities in combating fraud. These services, performed by pri-

[4]Bayer and Cowell (2009) find that the tax authority has a clear advantage to set-up an audit rule comparing firm's and competing firms tax returns —*i.e.*, the audit probability depends on competing firms reports, the relative rule— instead of using a fixed rule —*i.e.*, the probability of inspection is fixed.

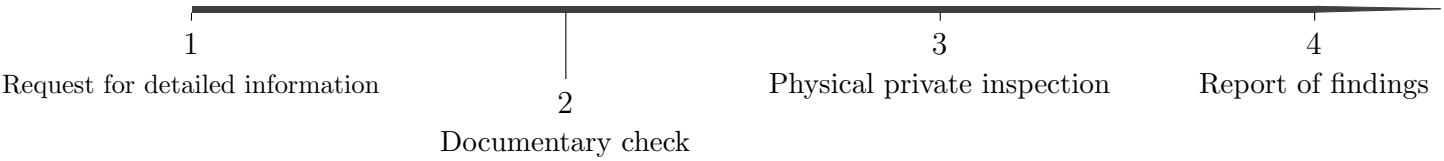
[5]Obviously, exploiting internal information is beneficial if and only if main characteristics of the product (*i.a.,* value-to-quantity-ratio, unit value, bulkiness) are strongly correlated among import transactions.

[6]On this, see Bayer and Cowell (2010).

vate companies operating in the country of export[7], aim at providing an opinion on main characteristics (value, quantity, classification ...) of the shipment after completing an inspection. For carrying out this activity, the surveillance company usually charges between 0.5% and 1.05% of the free-on-board (FOB) value of inspected merchandises with a minimum amount. Fees are paid either by the contracting government or by the importer. Finally, note that PSI contracts and PSI activities are regulated by the World Trade Organization (WTO) agreement on pre-shipment inspection and by the code of practice of the International Federation of Inspection Agencies (IFIA).

Given that several stages are needed to achieve an import, we will provide an overview of a standard import operation. In our country case study (country A), the administrative burden can be separated into two distinct (but complementary) procedures: (1) the pre-shipment inspection procedure and (2) the import customs clearance procedure. The timing of a classic import operation is detailed in the following two paragraphs.

PSI process. The PSI process takes place at the country of embarkation. A non-binding completed form —the request for detailed information (RDI)— is sent to the PSI firm local office. Analysing supporting documentation and the overall coherence of the RDI, the private company performs a preliminary price verification. Then, based on a risk analysis, certain shipments are subject to a physical inspection[8]. A report of findings (ROF) stating the estimated customs value, quantities and tariff classification[9] is finally issued[10]. Notice that the WTO agreement on customs valuation (henceforth ACV)[11] stipulates that the ROF can only be used by customs as an advisory document. The PSI process is summarized in the following diagram.



Import customs clearance procedures. When the shipment is at the seaport/airport/landport of destination, the economic operator makes the customs declaration (himself or by a representative namely the customs broker)[12]. Required documents (the bill of lading, the commer-

[7]The market structure is oligopolistic. Dequiedt, Geourjon, and Rota-Graziosi (2012) outline that SGS-BIVAC, Intertek and Cotecna account for more than 90% of the market.

[8]In some countries, goods are always physically inspected (e.g. Ghana or Togo).

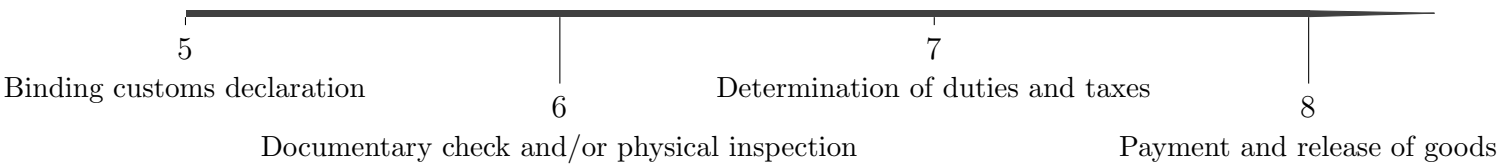
[9]The classification of goods is usually determined according to the the harmonized system of classification issued by the World Customs Organization.

[10]For sake of brevity, complaints and appeals procedures are not detailed.

[11]The WTO agreement on customs valuation (ACV) is available at the following adress: https://www.wto.org/english/docs_e/legal_e/20-val_01_e.htm.

[12]For further details on the declaration processing, see Keen (2003) page 74-78.

cial invoice, the ROF . . .) are filed with customs. Obviously, the PSI-supplied information acts as an incentive for the economic operator to make an import declaration in accordance with the report of findings. Differences may arise in case of disputes between the importer and the surveillance company or collusion between the importer and the customs officer[13]. From the risk analysis results[14], a documentary check and/or a physical inspection may be carried out. Using the ACV methodology and the customs code, key objectives of honest customs officers are then (i) to correctly assess the customs value (unit price and quantities) and (ii) to verify the classification of goods. Based on the inspector's report, customs announce the amount of tax and duties and eventually penalties to be paid. Once payment is complete, goods are released. The following diagram summarizes main customs clearance steps.



The efficiency of PSI services: an open question. The existing research literature on the evaluation of PSI programs does not provide a clear response (see De Wulf and Sokol (2005), McLinden et al. (2011)). Anson, Cadot, and Olarreaga (2006) studied four PSI programs. Results are mixed. Their empirical results suggest that PSI may have no impact (Indonesia), reduce (Philippines) or increase (Argentina) fraud. Jean and Mitaritonna (2010) find also an heterogeneous effect. In particular, their results indicate that PSI programs seem to be less efficient in least developed countries. Low (1995) argues that success depends on details of the PSI contract. Yang (2008a) stresses that PSI programs may create new forms of fraud rather than curbing it. Studying the Philippines PSI program. he finds evidence that, in response of the introduction of PSI programs, importers adapted their behaviors and found alternative duty-evasion methods. By splitting up shipments to get a shipment' value below the minimum value threshold or by importing via a duty-exempt export processing zone, bad practices continued. Nonetheless, switching to alternative methods became much more complex as corrective measures were set-up. Finally, Yang (2008b) outlines that, on average across countries, PSI programs are associated with a decrease of underreporting but that these improvements do not appear to persist over time. Thus, PSI programs may only have a short-term performance-enhancing effect.

[13]On this, see Dequiedt, Geourjon, and Rota-Graziosi (2012).

[14]For more details on modern risk analysis techniques, see Geourjon and Laporte (2005) and Geourjon and Laporte (2012).

3 Empirical strategy

3.1 Measure of underreporting

Our measure of underreporting is the so-called Trade gap[15]. The Trade gap is defined as the difference between the value of exports from each trade partner to country A as reported by each trade partner and the value of imports by country A from each trade partner as reported by country A. The logarithmic form aims at reducing the problem of non-normality of the residuals. Formally, we have:

$$Trade\ gap_{pe} = ln(Export\ value)_{pe} - ln(Import\ value)_{pe} \tag{1}$$

where $Export\ value_{pe}$ is the value of exports of the HS 6-digit product p reported by the partner country/exporter e to country A and $Import\ value_{pe}$ is the value of imports by country A from the partner country e of the HS 6-digit product p. As indicated by Bhagwati (1964), a positive gap will reveal an underreporting of imports.

3.2 Baseline specification

Econometric specification. We draw on the Fisman and Wei (2004)'s specification. We augment their model by mainly adding (i) our key variables of interest and (ii) some fixed effects (HS-2 industry - trade partner dummies). The following model is then obtained,

$$\text{Trade gap}_{pe} = \alpha + \gamma \text{Ln(\# imports)}_{pe} + \kappa \text{PSI frequency}_{pe} + \sigma \text{Taxes}_{pe} + \zeta \text{Differentiated product}_p + D_{ie} + \epsilon_{pe} \tag{2}$$

where γ and κ are our key parameters of interest; Trade gap_{pe} is the measure of underreporting (as defined in eq.1) and ϵ_{pe} is the error term. In order to remove potentially omitted variable bias, we include pairwise HS-2 industry - trade partner dummies D_{ie}. These dummies control for all factors— whether observable or unobservable– that are constant over HS-6 products belonging to the same HS-2 industry i and coming from the same trade partner/exporter e. Following Fisman and Wei (2004), we cluster standard errors at the HS-4 digit level to account for potential heteroskedasticity.

Use of the internal information available: the Ln(# imports) variable. Following subsection 2.1, we want to estimate the impact of the number of previous import declarations (in logarithm) —Ln(# imports)— on evasion. If there exists a negative relationship between our product-trade partner varying variable and evasion; then it will suggest that the more import declarations there has been, the easier it is for a customs officer to determine the transaction value of a present import. Notice that, in our final dataset, one import transaction tallies with one import declaration. As any declaration may be inspected in country A, the informational component of previous import declarations is potentially systematically

[15]See Fisman and Wei (2004), Javorcik and Narciso (2008), Mishra, Subramanian, and Topalova (2008), Jean and Mitaritonna (2010), Bouet and Roy (2012), Ferrantino, Liu, and Wang (2012), among others.

used[16]. Due to potential large differences in countries' product quality (see Hallak and Schott (2011) for example), it may be useless to compare the declared value of a product which has to be assessed with the declared value of a similar product coming from an another trade partner. For instance, it is likely that obtaining information on the value of Chinese dresses does not really help to assess the value of a French dress. Therefore, our preferred internal information variable only counts, for each HS-6 product, the number of import transactions from the exporting country (and not from the world). In the Appendix B, we will consider the number of import transactions from the world. The information quality associated with this latter variable being probably lower, we expect a smaller negative coefficient.

External information acquisition: the PSI frequency variable. As mentioned in the introductory section, country A uses the services of an information provider: a pre-shipment inspection company. Bear in mind that, prior to shipment, company's employees have to (i) assess the transaction value, (ii) verify the quantity and (iii) share their opinion through establishing a report of findings. Additional details of the local contract may be provided upon request. Exploiting the fact that pre-shipment inspections are not systematic (see subsection 2.2), we will estimate the causal relationship between the frequency of PSI and our evasion measure[17]. If the PSI company provides useful information to customs, we expect that undervaluation will be mitigated for observations with high PSI rates. Endogeneity issue will be discussed later.

Control variables. In our regressions, we control for the import taxation rate (Taxes) and the degree of differentiation of the product (Differentiated product).
Import taxation rate. Following Pritchett and Sethi (1994), we use collected taxes — *i.e.,* the sum of duties, excises and value added tax (VAT) effectively paid — as our measure of revenue. Divided by the import value, we get the *de facto* tax rate (Taxes). *De facto* means that the rate is calculated from taxes effectively paid and thus may differ from those mentioned in the customs code. Differences are due to granting preferential treatments. Some minor taxes — the regional integration tax and the statistical tax[18]— are not included since reliable data at the HS-6 product level are regrettably not available. Note that σ is the semi-elasticity of evasion with respect to taxes. Fisman and Wei (2004) argue that tax differentials may create an incentive to misreport the goods classification of a shipment. The misreporting hypothesis will be tested by including the average tariff on similar products.
Degree of differentiation. Mishra, Subramanian, and Topalova (2008) outline that the level of enforcement is product-varying. A proxy typically used is the degree of differentiation of the product. Following Javorcik and Narciso (2008), we use the classification of Rauch (1999). Whereas homogenous products[19], differentiated products (e.g. jackets) are those not having a reference price or not quoted on an organized exchange. We conjecture that, due

[16]Since customs officers are allowed to reroute a declaration from a facilitation channel to an inspection channel in case of suspicion, any declaration may be physically inspected. However, such requests must be duly reasoned in order to prevent malpractices.

[17]Our final database indicates for each couple HS6 product-trade partner the fraction of imports submitted to a pre-shipment inspection, see Appendix C.

[18]The statistical tax is a tax imposed to raise customs' administration revenues, see Doe (2006).

[19]A typical example of homogenous products is sugar.

to the lack of reference price, assessing the transaction value of differentiated products is much more difficult which in turn makes fooling easier. A larger evasion for differentiated products is thus expected.

Endogeneity issue and potential bias. Evasion may be estimated with measurement errors. Indeed, some parts of discrepancies in trade statistics are due to factors other than tax evasion such as statistical reasons —difference in (c.i.f - f.o.b) valuation, see Nitsch (2012)—, exchange rate conversion issues (see Carrère and Grigoriou (2014)), unintentional classifications (see Jean and Mitaritonna (2010)) or weak export control laws (see Stoyanov (2012)). However, to the extent that measurement errors in the evasion measure are not related to the error term, the estimator remains unbiased (but less efficient, see Jean and Mitaritonna (2010)). In the case of export subsidies granted by the trade partner, over-reporting of exports may arise. Nevertheless, under the reasonable assumption that export misreporting is not related to our independent variables, estimations are not biased.

It is noteworthy that due to the common external tariff — country A is member of a regional integration zone— tariff rate applies to country A. The government is thus not able to modify tariffs according to tax evasion. Concerning VAT, the general tax code sets out only two distinct VAT rates — the general rate and the zero rate for basic necessities –. Therefore, it is very unlikely that the government determines the VAT rate on the basis of customs evasion.

As regard as the pre-shipment inspection variable, the fact that shipments below a minimum value are exempted from private inspections may create an incentive to underreport. Yang (2008b) shows that, at the beginning of the Philippine PSI program, this exemption was exploited by tax evaders. Basically, the usual strategy consists in declaring a value below the threshold to avoid a pre-shipment inspection. Graphical representations of cumulative distribution function of shipments with a value around the PSI threshold (see figure 1) suggest that these fraudulent activities are not common. One remaining major concern is that an effective PSI risk analysis implies that riskiest import declarations are inspected more frequently. The PSI frequency variable may therefore be affected by the Trade gap. In such a situation, the error term ϵ_{pe} is correlated with the pre-shipment inspection variable meaning that the independent and identically distributed (i.i.d.) assumption is violated. Therefore, we use an instrumental variable approach to deal with potential reverse causation. Specifically, we use the fraction of import declarations handled by a specific customs office —office #1— to instrument the frequency of private inspections. This instrument varies by product and trade partner. Office #1 is the main customs clearance office of country A. It is worth mentioning that, according to the shipment's characteristics (type of goods, container size ...), import declarations are transmitted to the most appropriate customs local office. Guidance criteria are thus exclusively based on technical criteria. Put it differently, it means that office #1's activity does not rely on any risk analysis. Let us now discuss instrument exogeneity and instrument relevance. Given that performance between customs offices are close —a detailed justification is available upon request—, we argue that enforcement does not vary between customs offices. And since, for each shipment, the selection of the customs

clearance office and the probability to carry out a private inspection are partly based on similar criteria (*i.a.*, type of goods, containerised cargo), we argue that there is a positive association between these two variables. Therefore, we conjecture that the proposed instrument is valid. Future statistical tests will support the hypothesis.

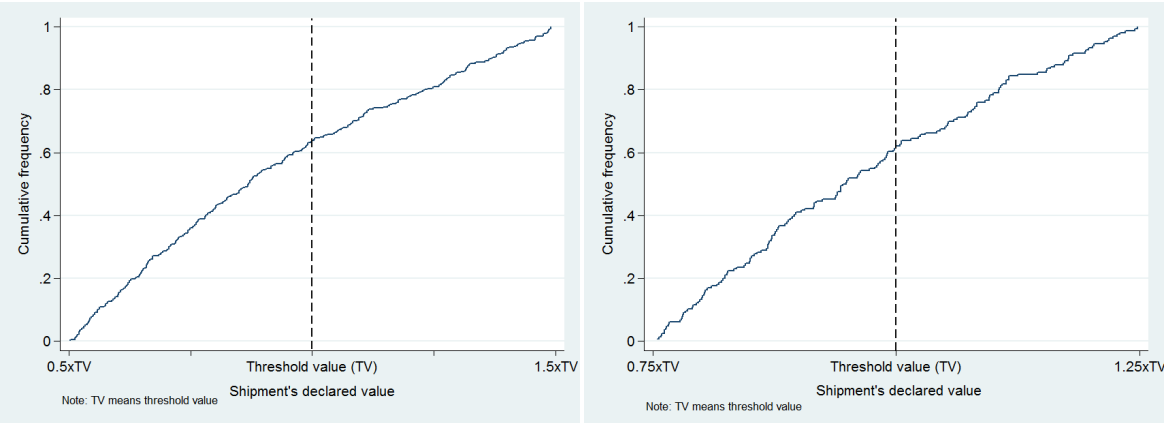


Figure 1: Statistical distribution of shipments with a value close to the PSI minimum value, bandwidths: 0.5 × the minimum value (left) or 0.25 × the minimum value (right)
Observation: Displayed graphical representations are quite linear and distributions have no really peak suggesting that declaring shipment values below the threshold to avoid PSI is not a common practice.

4 Data and descriptive overview

Dataset. The dataset is drawn from an export database and an import database. Country-disaggregated export data are provided on an annual basis and come from the well-known UN COMTRADE database. Local customs have provided transaction-level import data. Interestingly, each customs declaration provides in-depth information including the HS-11 digit product code, the customs office of entry, the presence of any attached administrative document (*i.a.*, the report of findings), the date of registration, the exporting country, the declared value, the declared quantity and the declared weight. Since export data cannot be disaggregated beyond the 6-digit level, import data have been aggregated on a 6-digit basis. Import values are expressed in the local monetary unit. A conversion into U.S. dollars has been performed using the exchange rate at the date of registration of the declaration. Exchange-rate miscalculations are thus strongly reduced. Since we consider the European Union as a single trade partner, the "Rotterdam effect" — *i.e.*, differences in import and export reports due to the transit trade, see Herrigan, Kochen, and Williams (2005)— is strongly minored. Trade partners that do not report their exports to UN COMTRADE (e.g. Egypt, Morocco, Nigeria, United Arab Emirates) are excluded from the sample. Due to important variations in the nomenclature used in recent years, our study covers only one year (2013). Main explanatory variables have been obtained from the Information Technology department, local customs. Appendix C provides a full list of variables, definitions, and sources.

In order to assess the quality of our database, we can rely on the match rate between imports and exports. This ratio is defined as the number of observations for which a non-zero export and a non-zero import are registered divided by the total number of observations (unit of analysis: HS6 product-trade partner level). The average match rate is quite large (66%) indicating that data seem to be reliable[20]. The rate is slightly higher than in Van Dunem and Arndt (2009) (55%) and Mishra, Subramanian, and Topalova (2008) (65%). Table 1 presents for the top ten partners the corresponding match rate. It is noteworthy that, for the European Union, the coverage is particulary strong (91%). Non-matched observations are removed from our final dataset. Then, the sample is formed by 52 trade partners and 3,085 products. The number of exported HS6 products per country is available upon request.

Table 1: Match rates across main trade partners

Country name	Share of HS6 products that are in both databases (%)
European Union	91
China	86
United States	87
India	57
South Africa	58
Turkey	52
Japan	81
Taiwan	56
Switzerland	61
Republic of Korea	62

Descriptive statistics. Table 2 displays the summary statistics of the data used, in which each observation is at the HS-6 product-trade partner level. For dummy variables, the mean is the part of observations for which the variable is equal to one. Our sample has 8,749 observations. With no evasion, we expect a negative gap due to the c.i.f-f.o.b margin. However, observe that mean and median of Trade gap are positive. This feature supports the hypothesis that tax evasion is quite widespread. Note that these discrepancies are larger than in Tanzania (mean: 0.004, see Levin and Widell (2014)) but lower than in Mozambique (mean: 0.290, see Van Dunem and Arndt (2009)) and Kenya (mean: 0.724, see Levin and Widell (2014)). Finally, observe that over 25% of the discrepancies are negative (p25$<$ 0, row 4) suggesting that some factors reduce ability to cheat.

Distributions of our measure of underreporting is fairly normally distributed (see figure 2). Therefore, we can perform a Pearson's correlation to assess the relationship between our dependent variable (Trade gap) and our regressors[21], see table 3. In *prima facie*, the

[20]The average match rate is weighted by the number of traded products. For further details, see Mishra, Subramanian, and Topalova (2008).

[21]Note that the Pearson's correlation coefficient is a measure of linear relationship. Therefore, a not significant value does not imply that there is no relationship between the variables.

Table 2: Summary statistics

	Mean	Median	Min	Max	SD	p5	p25	p75	p95	Observation
Ln(Export value)	10.130	10.290	0.000	18.940	2.843	5.075	8.294	12.130	14.500	8,749
Ln(Import value)	9.943	10.060	-1.874	18.810	2.784	5.256	8.081	11.890	14.360	8,749
Trade gap	0.188	0.070	-11.240	11.810	2.241	-3.507	-0.876	1.261	4.051	8,749
PSI frequency	0.639	0.720	0.000	1.000	0.353	0.000	0.385	1.000	1.000	8,749
Ln(# of imports)	1.978	1.792	0.000	9.290	1.603	0.000	0.693	3.045	4.956	8,749
Ln(# of imports from the world)	4.403	4.431	0.000	9.568	1.782	1.386	3.135	5.746	7.189	8,749
Taxes	0.306	0.293	0.000	6.561	0.231	0.000	0.111	0.494	0.577	8,749
Differentiated product	0.831	1.000	0.000	1.000	0.375	0.000	1.000	1.000	1.000	7,073

Summary statistics for the dependent variable Trade $\text{gap}_{pe} = \ln(\text{Export value})_{pe} - \ln(\text{Import value})_{pe}$ and for the independent variables. Please refer to the Appendix C for a detailed definition of these variables. SD is standard deviation, p stands for percentile in this table. Observations are at the HS 6 product-trade partner level. Such high maximal taxation rate is explained by the presence of products subject to compound tariffs (*i.e.*, a combination of ad valorem and specific tariffs). In a robustness check, we will exclude from the sample products subject to an administrative value.

quantity of internal information appears to mitigate underreporting. The correlation coefficient between the gap and pre-shipment inspections is not significant. Nevertheless, notice that since correlation only quantifies the extent to which two variables go together, correlation is not causality. The positive and significant linear correlation between Taxes and the Trade gap supports the hypothesis of tax evasion. Large rates are associated with large gaps. Evasion seems to be more prevalent for differentiated products. Furthermore, table 3 points that none of the pairwise correlation coefficients between two explanatory variables are in absolute value greater than 0.8 (*i.e.*, the high correlation threshold usually recognized[22]) indicating that there is no multicollinearity between any of two explanatory variables. Finally, note that our instrument is strongly and positively correlated with PSI frequency — the Pearson's correlation coefficient is positive (0.701) and significant at 1%— but uncorrelated with the dependent variable — the coefficient (0.015) is not significant at 1% —. Thus, the fraction of import declarations handled by the main customs clearance office (office #1) appears to be a plausible instrument (relevant and exogenous).

[22]It is usually acknowledged that when the Pearson correlation (in absolute value) is greater than 0.8 the issue of multicollinearity arises.

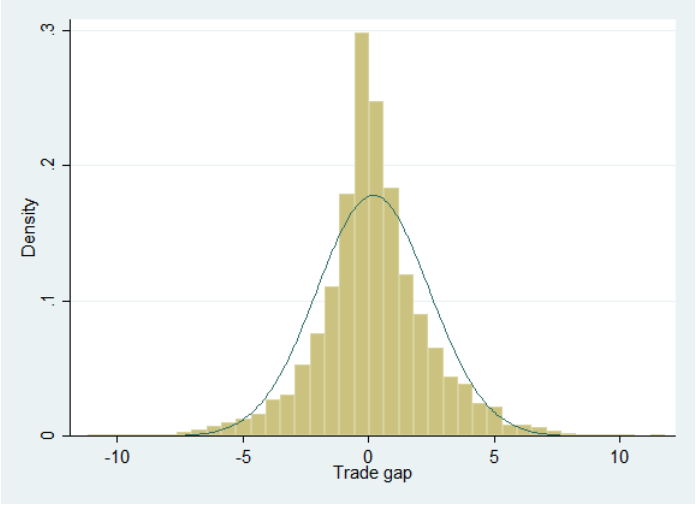


Figure 2: Density distribution of the Trade gap

Table 3: Pearson's correlations

Variable	Trade gap	Taxes	PSI frequency	Ln(# of imports)	Differentiated product
Taxes	0.137***	1			
PSI frequency	0.007	0.346***	1		
Ln(# of imports)	-0.224***	0.000	-0.037***	1	
Differentiated product	0.040***	0.040***	-0.144***	0.067***	1

Notes:
(i) Observations are at the observation is at the HS-6 product-trade partner level.
(ii) Trade $\text{gap}_{pe} = \ln(\text{Export value})_{pe} - \ln(\text{Import value})_{pe}$.
(iii) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level. For further details, please see the text.

5 Results and robustness

5.1 Basic results

We estimate equation (2). The regression results are displayed in table 4. We use a number of dummy variables to ensure robust econometric identification. Column (2) includes trade partner dummies. The inclusion of trade partner dummies capture any individual partner characteristic such as the geographical distance or differences in the quality of export monitoring. Column (3) adds HS-2 industry dummies. Industry dummies D_i take into account the unobserved heterogeneity across industries — *i.e.,* determinants of gaps not industry-varying—. Column (4) replaces trade partner and HS-2 industry dummies with couple dummies — *i.e.*, HS-2 industry-trade partner dummies —. Regression 4 is more refined relative to other regressions as it controls for industry-trade partner characteristics. Overall, we find that (i) acquiring external information through carrying out pre-shipment inspections and (ii) using the internal information provided by previous import declarations as a help for customs valuation have a positive impact on enforcement.

Impact of the use of information sources on underreporting. The negative coefficient on *PSI frequency* suggests that pre-shipment inspections reduce evasion. The coefficient is quite stable across specifications. Quantitatively, column (4) indicates that a one percentage point increase in pre-shipment inspections leads to a reduction of approximately 0.357% (i.e., $e^{0.356/100}-1$) of underreporting. Notice that this estimated ordinary least squares (OLS) coefficient is biased due to endogeneity. Two stage least squares (TSLS) estimates will highlight that the downward bias is substantial. Although these results *per se* do not justify contracting with a PSI firm[23], they strongly support the idea that reducing importers' ability to cheat involves contracting with information providers. As stressed by Zucman (2013) in tax matters, completing and implementing an international agreement on automatic exchange of customs information would strongly increase enforcement[24].

As for internal information, column (4) indicates that every one percentage increase in import declarations results in a 0.41% decline in underreporting. This estimate is consistent with the hypothesis that customs officers use information they have learned from similar import declarations to assess quite accurately the import value. More generally, this result highlights that setting-up a risk analysis system fully exploiting in a systematic and formal way the free-information provided by previous import declarations may substantially improve customs performance. Finally, observe that the magnitude of the coefficient of the (natural logarithm of the) number of import declarations from the world —Ln(# of imports from the world)— is one-half (-0.2, see results on table 7, Appendix A). This result simply indicates that the informational component of historical data increases with the relevance of data used.

[23] A cost-benefit analysis would be necessary to draw any conclusions concerning the PSI effectiveness.

[24] To promote communication between customs, the World Customs Organization has recently launched the Customs Enforcement Network (CEN) digital platform. For more details on this application, the reader may refer to Han and McGauran (2014).

Impact of other explanatory variables. Control variables are significant with the expected sign. Clearly, the estimated semi-elasticity of underreporting with respect to taxes is positive and significant. Our estimates outline that a one percentage point increase in taxes (tariff + excise + VAT) leads to about a 1.57 % (*i.e.* $exp^{(1.553/100)} - 1$) decrease in declared imports, see column (4)[25]. In other words, a one percentage point increase in taxes will cause a 1.57% increase in Trade gap. The estimate is slightly larger with that found for Mozambique (1.38%, see Van Dunem and Arndt (2009)) but lower than the estimate of Tanzania (2.36%, see Levin and Widell (2014))[26]. The coefficient associated to the product differentiation dummy is significant and positive indicating that the quality of enforcement is product-varying. This result confirms that fooling is easier when the price is not well known. Enforcement is thus weaker for differentiated products[27].

Causality issues: IV approach. To address the potential endogeneity of PSI frequency, we use the instrumental variable approach. First-stage regression results confirm the instrument relevance (see table 6, Appendix A). The coefficient of the instrument is positive and statistically significant at the 1% level. It is noteworthy that Office #1 frequency explains about 50% percent of the variance of the PSI variable. To test the instrument exogeneity, we conduct a statistical test. Precisely, we regress the Trade gap on the Office #1 frequency. The coefficient is not significant indicating that it is very unlikely that the exclusion restriction is violated[28].

The results of the IV estimation — second stage regressions— are displayed in column (4)-(8), table 4. First-stage F statistics —Kleibergen-Paap Wald statistics— indicate that the instrument is not weak[29]. The qualitative nature of our results is unaffected. Instrumenting for the potential endogeneity of PSI frequency, the magnitude of the PSI coefficient is, as expected and in absolute terms, larger than in the OLS estimations. Specifically, instru-

[25]From eq.2, we get: $\frac{\partial \text{Trade gap}}{\partial \text{Taxes}} = \sigma$. Using eq.1, it yields: $\frac{\frac{\partial \text{Import value}}{\partial \text{Taxes}}}{\text{Import value}} = \frac{\frac{\partial \text{Export value}}{\partial \text{Taxes}}}{\text{Export value}} - \sigma$. Observe that, as in Fisman and Wei (2004), this reduced form highlights that increasing taxes may induce a negative value effect —*i.e.,* a decline of exports to country A, see the first term— and a negative compliance effect —*i.e.,* the fraction of true imports reported to local customs is reduced, see the second term $(-\sigma)$—.

[26]Of course, when we use the official/*de jure* tax rate as our tax measure, the estimated coefficient is (slightly) lower. Results are not reported, but they are available upon request.
Misclassification of imported goods — *i.e.*, mislabelling a higher-taxed product as a lower-taxed similar product — may be a useful strategy to evade taxes. To test for misreporting, we follow Fisman and Wei (2004) and we add the the average tax rate of the HS-4 digit category as an additional regressor. As in Levin and Widell (2014) for Tanzania, we find no hard evidence of evasion by misclassification within 4-digit classifications (results are not reported but they are available upon request). It suggests that understating import value at customs clearance is the main channel of evasion. Notice that, due to our specific targets, products that illegally enter the territory are mechanically excluded from the final sample. The smuggling evasion channel is therefore not considered in this paper.

[27]Contrary to Mishra, Subramanian, and Topalova (2008), the effect of this "ease-of-enforcement" measure seems to be linear, results are available upon request.

[28]For the sake of brevity, results are not reported. They are obviously available upon request.

[29]Kleibergen-Paap Wald statistics are much greater than the Stock-Yogo weak ID test critical values of 16.38 suggesting that the assumption of weak identification is rejected. Finally, note that, in this study of a single endogenous regressor, the Kleibergen-Paap Wald statistics is the heteroskedasticity-robust first-stage F statistics.

Table 4: Extended regression analysis: OLS and IV

Variables	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PSI frequency	-0.209**	-0.287***	-0.324***	-0.356***	-0.304*	-0.502***	-0.611***	-0.694***
	(0.106)	(0.101)	(0.114)	(0.134)	(0.163)	(0.170)	(0.190)	(0.206)
Ln(# imports)	-0.316***	-0.419***	-0.417***	-0.414***	-0.316***	-0.420***	-0.417***	-0.414***
	(0.017)	(0.018)	(0.018)	(0.020)	(0.017)	(0.018)	(0.018)	(0.019)
Taxes	1.404***	1.549***	1.486***	1.553***	1.454***	1.653***	1.616***	1.710***
	(0.197)	(0.209)	(0.265)	(0.297)	(0.220)	(0.235)	(0.293)	(0.314)
Differentiated product	0.265***	0.370***	0.149	0.210*	0.251***	0.338***	0.130	0.189
	(0.070)	(0.074)	(0.114)	(0.124)	(0.073)	(0.077)	(0.112)	(0.115)
Constant	0.290***	0.368	0.239	0.578***	0.348***	0.569*	0.0717	1.514***
	(0.103)	(0.315)	(0.315)	(0.148)	(0.124)	(0.338)	(0.420)	(0.197)
Trade partner dummies		Yes	Yes			Yes	Yes	
HS-2 industry dummies			Yes				Yes	
Couple dummies				Yes				Yes
Estimator	OLS	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS
Observations	7,073	7,073	7,073	7,073	7,073	7,073	7,073	7,073
K-P Wald F stat					688.489	793.338	1,207.291	1,063.406
R-squared	0.071	0.119	0.135	0.266	0.071	0.119	0.134	0.264
Adjusted R-squared	0.071	0.113	0.117	0.158	0.071	0.112	0.116	0.156

Notes:
(i) Observations are at the HS-6 product-trade partner level.
(ii) The dependent variable is Trade $\text{gap}_{pe} = \ln(\text{Export value})_{pe} - \ln(\text{Import value})_{pe}$.
(iii) Standard errors are in parentheses, clustered by HS-4 products.
(iv) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level.
(v) Couple dummies: HS-2 industry-trade partner dummies.
(vi) OLS is Ordinary Least Squares, 2SLS is Two Stage Least Squares.
(vii) K-P Wald F stat means Kleibergen-Paap Wald F statistics.

mental variables estimations outline that a one percentage point increase in pre-shipment inspections reduce underreporting by 0.696% (i.e., $e^{0.694/100} - 1$), see column (8), table 4.

Use of the internal information and external information acquisition: independent, substitute, or complementary? To test whether the combination of the two information sources is associated with a lower gap —*i.e.,* sources are complementary—, we introduce the multiplicative variable PSI frequency × Ln(# of imports). The coefficient associated is not significant. We interpret this result as evidence that the economic operator fully adjusts its import declaration to the report of findings (see timing in section 2.2). This behavior would explain why PSI-supplied information seems to be not useful for the customs officer. Such importer's response is in line with Carrillo, Pomeranz, and Singhal (2014)'s findings. Empirical results are omitted in the interest of parsimony but are available from the authors upon request.

Use of an alternative dependent variable. Interpretations so far have been based on the assumption that the number of previous import operations does not impact the $\frac{c.i.f}{f.o.b}$factor: $\frac{\partial \frac{c.i.f}{f.o.b}\text{factor}_{pe}}{\partial \text{Ln}(\#\text{ imports})_{pe}} = 0$. Equation 3 outlines that the derivative of the Trade gap is the sum of derivatives of (i) the Quantity gap, (ii) the Unit value gap and (iii) the $\frac{c.i.f}{f.o.b}$ factor; with Quantity gap $= ln(\frac{\text{Exported quantity}}{\text{Imported quantity}})$ and Unit value gap $= ln(\frac{\text{Exported unit value}}{\text{Imported unit value}})$.

$$\frac{\partial \text{ Trade gap}}{\partial \text{Ln}(\#\text{ imports})} = \frac{\partial \text{Quantity gap}}{\partial \text{Ln}(\#\text{ imports})} + \underbrace{\frac{\partial \text{ Unit value gap}}{\partial \text{Ln}(\#\text{ imports})} + \frac{\partial \frac{c.i.f}{f.o.b}\text{factor}}{\partial \text{Ln}(\#\text{ imports})}}_{\frac{\partial Ln(\frac{\text{Exported unit value}_{f.o.b}}{\text{Imported unit value}_{c.i.f}})}{\partial \text{Ln}(\#\text{ imports})}} \tag{3}$$

Proof is detailed in the Appendix D.

If a relationship between Ln(# imports) and the $\frac{c.i.f}{f.o.b}$ factor exists, then the estimated marginal effect differs from the specific impact of previous imports declarations on underreporting. To overcome this potential problem, we will use a not $\frac{c.i.f}{f.o.b}$ factor-varying dependent variable. Specifically, we will restrict our attention to discrepancies in volume[30]. If there was no exploitation of historical data, (the natural logarithm of) the number of previous import declarations would not impact this underreporting measure. Given that the unit of measurement for the supplementary quantity is product-varying (item, liter, meter, ...), we will use the net weight —always expressed in kilograms— as our quantity variable. Following the Trade gap, the Weight gap is thus defined as follows: Weight $\text{gap}_{pe} = ln(\frac{\text{Exported weight}}{\text{Imported weight}})_{pe}$.

Based on equation 2, the following equation is then estimated:

[30]Bear in mind that, by definition, cheating on the total value implies cheating on the volume (quantity) and/or on the unit value (as total value = quantities × unit value).

$$\text{Weight gap}_{pe} = \alpha_w + \gamma_w \text{Ln(\# imports)}_{pe} + \kappa_w \text{PSI frequency}_{pe} + \sigma_w \text{Taxes}_{pe} + \zeta_w \text{Differentiated product}_p + D_{pi} + u_{pe} \quad (4)$$

where u_{pe} is the error term and subscript w stands for the Weight gap specification.

Assessing accurately the quantity may be a difficult task especially when goods are containerized and/or in case of large shipment. Accumulating information, knowledge about the usual weight, the usual bulkiness through the exploitation of historical data may help the customs officer in determining quantities imported. Therefore, we expect that $\gamma_w < 0$. Given that PSI companies have to produce a detailed opinion of quantities shipped, a negative relationship between the PSI frequency and underreporting of quantities is also expected.

Table 5 provides additional evidence in favor of the exploitation of historical data to combat fraud. The coefficient on the Ln(# of imports) variable very slightly changes. Results also reiterate the role played by pre-shipment inspections in reducing underreporting. Furthermore, estimates highlight that quantity-based evasion increases with taxes. As observed in previous estimations (see table 4), when we include HS-2 industry dummies the coefficient associated with the product differentiation variable is not significant due to a lack of intra-industry variability.

Table 5: Weight gap as an alternative dependent variable

Variables	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PSI frequency	-0.491***	-0.509***	-0.615***	-0.661***	-0.702***	-0.750***	-1.081***	-1.218***
	(0.107)	(0.112)	(0.122)	(0.146)	(0.178)	(0.186)	(0.202)	(0.230)
Ln(# of imports)	-0.337***	-0.420***	-0.418***	-0.414***	-0.338***	-0.421***	-0.417***	-0.414***
	(0.019)	(0.022)	(0.022)	(0.024)	(0.019)	(0.022)	(0.022)	(0.022)
Taxes	0.536***	0.680***	0.429**	0.418**	0.645***	0.795***	0.638***	0.673***
	(0.154)	(0.161)	(0.183)	(0.202)	(0.175)	(0.184)	(0.216)	(0.231)
Differentiated product	-0.306***	-0.189**	-0.074	0.023	-0.337***	-0.225***	-0.104	-0.013
	(0.083)	(0.084)	(0.127)	(0.139)	(0.086)	(0.087)	(0.125)	(0.130)
Constant	0.767***	1.062***	0.608**	2.505***	0.899***	0.606***	-1.752	-1.215
	(0.116)	(0.278)	(0.286)	(0.171)	(0.146)	(0.198)	(2.008)	(1.978)
Trade partner dummies		Yes	Yes			Yes	Yes	
HS-2 industry dummies			Yes				Yes	
Couple dummies				Yes				Yes
Estimator	OLS	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS
Observations	6,836	6,836	6,836	6,836	6,836	6,836	6,836	6,836
R-squared	0.053	0.084	0.114	0.243	0.053	0.083	0.112	0.240
Adjusted R-squared	0.053	0.077	0.095	0.131	0.052	0.076	0.093	0.128

Notes:
(i) Observations are at the HS-6 product-trade partner level.
(ii) The dependent variable is Weight $\text{gap}_{pe} = \ln(\text{Exported weight})_{pe} - \ln(\text{Imported weight})_{pe}$.
(iii) Standard errors are in parentheses, clustered by HS-4 products.
(iv) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level.
(v) Couple dummies: HS-2 industry- trade partner dummies.
(vi) OLS is Ordinary Least Squares, 2SLS is Two Stage Least Squares.

5.2 Further robustness checks

We validate previous findings through a number of additional robustness checks. Results are reported in the Appendix B. In table 8, we start by re-estimating the full specification by employing an alternative variable quantifying the internal information available. The objective is to use a less disaggregated variable to avoid any potential measurement error. Specifically, we use a high correlated variable with the number of import declarations from the world namely the number of trade partner exporting the HS6-product[31]. Figure 3 illustrates the strong uphill positive relationship. Estimates outline that entrance of a new supplier country lowers underreporting by 3,2% (see column 2). Thus, irrespective of how we quantify the internal information available, we find a negative and statistically significant impact on evasion. Then, as an additional robustness check, we use the Rauch classification of goods based on the liberal definition instead of the conservative definition. Next, in order to increase the sample size, we exclude from regressors the differentiation variable. As for the alternative dependent variable (Weight gap), previous findings are also confirmed (see table 10).

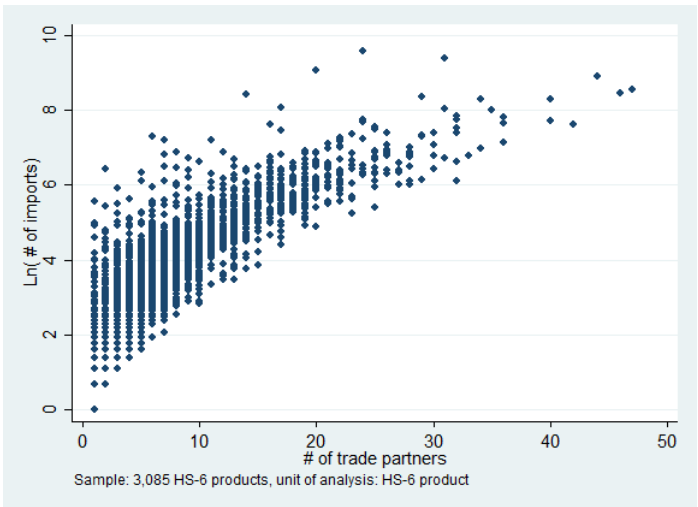


Figure 3: A positive relationship between the # of exporting countries and the (ln) # of imports from the world

Finally, results are also robust to a variety of sample restrictions. In table 9 columns (1)-(12), the sample is restricted by successively removing (i) products subject to specific tariffs, (ii) main problematic HS-2 -industries (worn clothing; textile; mineral fuels, mineral oils; miscellaneous manufactured articles; art, collectors' piece and antiques[32]) and (iii) potential outliers (first and last 0.05 quantile). We also present regressions using the alternative dependent variable (table 11). Estimated coefficients remain essentially unchanged.

[31] Of course, the Pearson's correlation coefficient is high and positive (0.867) and significant at 1%. The correlation coefficient between the number of import declarations from the trade partner (in logarithm) and the number of exporting countries is also positive (0.236) and significant at 1%; unit of analysis: HS-6 product - trade partner.

[32] We exclude following chapters: 61, 62, 63, 94, 95, 96, 97.

A final concern with our empirical estimates may be related to the intra-regional trade. As neighbouring countries do not report their exports to UN COMTRADE, intra-regional transactions are excluded from the sample. If biggest tax evaders have a strategy which consists of importing through customs offices where enforcement is the weakest —*i.e.*, through landports in country A—, then we underestimate the semi-elasticy of evasion. However, it is noteworthy that implementing such a strategy may actually induce little (or no) benefits since intra-regional transport costs are large due to a lack of road infrastructure and poor trade logistics performances. To test whether this potential tax evasion strategy may occur, we remove all HS-6 products having been imported at least once through a land border[33]. This sample restriction yields similar results in terms of magnitude of coefficients and significance levels (see columns (13)-(14), tables 9 and 11).

6 Concluding remarks

The contribution of this micro-level study to the extensive literature on evasion mainly consists in providing evidence that using any relevant source of information — internal or external— facilitates customs-enforcement. Specifically, we quantify the impact of using the internal information and acquiring external information on evasion. Estimates suggest that a one percentage increase in import declarations implies a 0.41% decline in underreporting. Further, results point that PSI programs significantly reduce observed trade statistics discrepancies. We leave for future research the cost-benefit analysis of the pre-shipment inspection program. These findings are supported by a variety of robustness checks.

From an operational perspective, this article suggests that revenue-collection authorities should exploit non-usual sources of information to improve their performances. In particular, authorities should modernize their risk analysis system in order to fully exploit the informational component of historical data. This paper also sheds light that relying on an (ideally free) information provider may be a valuable help to reach the revenue collection target. A powerful free-information provider may consist in setting-up an automatic exchange of information system across customs authorities.

[33]Since informal cross-border trade is not recorded by local customs, we are unfortunately not able to precisely identify HS-6 products coming informally into country A (see e.g. Golub (2015)). Nevertheless, notice that a large part of HS-6 products partly coming informally into country A are presumably excluded in at least one of our various sample restrictions.

References

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

Almunia, M., and D. Lopez-Rodriguez (2013): "Firms' responses to tax enforcement strategies: Evidence from Spain," *Munich Personal RePEc Archive*.

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

Anson, J., O. Cadot, and M. Olarreaga (2006): "Tariff evasion and customs corruption: Does pre-Shipment inspection help?," *Contribution to Economic Analysis and Policy*, 5(1), 1600–1623.

Bachas, P., and A. Jensen (2014): "An empirical test of information trails: Do firm finance and size matter for tax collection?," Discussion paper.

Bayer, R., and F. Cowell (2009): "Tax compliance and firms' strategic interdependence," *Journal of Public Economics*, 93(11), 1131–1143.

——— (2010): "Tax compliance by firms and audit policy," *London School of Economics STICERD Research Paper. London: London School of Economics.*

Bhagwati, J. (1964): "On the underinvoicing of imports," *Bulletin of the Oxford University Institute of Economics & Statistics*, 27(4), 389–397.

Bouet, A., and D. Roy (2012): "Trade protection and tax evasion: Evidence from Kenya, Mauritius, and Nigeria," *The Journal of International Trade & Economic Development*, 21(2), 287–320.

Carrère, C., and C. Grigoriou (2014): "Can mirror data help to capture informal international trade?," *Trade and commodities research study series. Geneva: United Nations Conference on Trade and Development.*

Carrillo, P., D. Pomeranz, and M. Singhal (2014): "Dodging the taxman: Firm misreporting and limits to tax enforcement," *Harvard Business School Working Paper. Boston: Harvard University.*

De Wulf, L., and J. B. Sokol (2005): *Customs modernization handbook*. World Bank Publications.

Dequiedt, V., A.-M. Geourjon, and G. Rota-Graziosi (2012): "Mutual supervision in preshipment inspection programs," *Journal of Development Economics*, 99(2), 282–291.

Doe, L. K. (2006): "Reforming external tariffs in Central and Western African Countries," *IMF working paper. Washington: International Monetary Fund.*

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., P. Moustakerski, and S.-J. Wei (2008): "Outsourcing tariff evasion: A new explanation for entrepot trade," *The Review of Economics and Statistics*, 90(3), 587–592.

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

——— (2009): "The smuggling of art, and the art of smuggling: Uncovering the illicit trade in cultural property and antiques," *American Economic Journal: Applied Economics*, 1(3), 82–96.

Geourjon, A.-M., and B. Laporte (2005): "Risk management for targeting customs controls in developing countries: A risky venture for revenue performance?," *Public Administration and development*, 25(2), 105–113.

——— (2012): "Risk management in customs: Initial lessons from experiences in some African countries," *Revue d'économie du développement*, 20(3), 67–82.

Golub, S. (2015): "Informal cross-border trade and smuggling in Africa," *Handbook on Trade and Development*, p. 179.

Golub, S. S. (2012): "Entrepôt trade and smuggling in West Africa: Benin, Togo and Nigeria," *The World Economy*, 35(9), 1139–1161.

Hallak, J. C., and P. K. Schott (2011): "Estimating cross-country differences in product quality," *The Quarterly Journal of Economics*, 126(1), 417–474.

Han, C.-R., and R. McGauran (2014): "Tracing trails: implications of tax information exchange programs for customs administrations1," *World Customs Journal*, 8(2), 3–13.

Herrigan, M., A. Kochen, and T. Williams (2005): "Analysis of asymmetries in intra-community trade statistics with particular regard to the impact of the Rotterdam and Antwerp effects," Discussion paper, Statistics and Analysis of Trade Unit, HM Revenue & Customs,United Kingdom.

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 custom duties," *CEPII Working Paper. Paris: Centre d'études prospectives et d'informations internationales.*

Keen, M. (2003): *Changing customs: Challenges and strategies for the reform of customs administration.* International Monetary Fund.

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.

Levin, J., and L. M. Widell (2014): "Tax evasion in Kenya and Tanzania: Evidence from missing imports," *Economic Modelling*, 39, 151–162.

Low, P. (1995): *Preshipment inspection services*. Washington: World Bank Publications.

McLinden, G., et al. (2011): *Border management modernization*. Washington: World Bank Publications.

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

Montagnat-Rentier, G., and G. Parent (2012): *Customs administration reform and modernization in Francophone Sub-Saharan Africa, 1995–2010*. Washington: International Monetary Fund.

Nitsch, V. (2012): *Trade mispricing and illicit flows*. World Bank Publications.

Pomeranz, D. (2015): "No taxation without information: Deterrence and self-enforcement in the value added tax," *American Economic Review, forthcoming*.

Pritchett, L., and G. Sethi (1994): "Tariff rates, tariff revenue, and tariff reform: some new facts," *The World Bank Economic Review*, 8(1), 1–16.

Rauch, J. E. (1999): "Networks versus markets in international trade," *Journal of international Economics*, 48(1), 7–35.

Rotunno, L., and P.-L. Vézina (2012): "Chinese networks and tariff evasion," *The World Economy*, 35(12), 1772–1794.

Saez, E., J. Slemrod, and S. H. Giertz (2012): "The elasticity of taxable income with respect to marginal tax rates: A critical review," *Journal of Economic Literature*, 50(1), 3–50.

Sandmo, A. (2005): "The theory of tax evasion: A retrospective view," *National Tax Journal*, 58(4), 643.

Slemrod, J., and S. Yitzhaki (2002): *Handbook of Public Economics*chap. Tax avoidance, evasion, and administration, pp. 1423–1470. Elsevier.

Stoyanov, A. (2012): "Tariff evasion and rules of origin violations under the Canada-US Free Trade Agreement," *Canadian Journal of Economics/Revue canadienne d'économique*, 45(3), 879–902.

Van Dunem, J. E., and C. Arndt (2009): "Estimating border tax evasion in Mozambique," *The Journal of Development Studies*, 45(6), 1010–1025.

Yang, D. (2008a): "Can enforcement backfire? Crime displacement in the context of customs reform in the Philippines," *The Review of Economics and Statistics*, 90(1), 1–14.

——— (2008b): "Integrity for hire: an analysis of a widespread customs reform," *Journal of Law and economics*, 51(1), 25–57.

Zake, J. O. (2011): *Customs Administration Reform and Modernization in Anglophone Africa-Early 1990s to Mid-2010*. Washington: International Monetary Fund.

Zucman, G. (2013): "The Missing Wealth of Nations: Are Europe and the US net Debtors or net Creditors?," *The Quarterly journal of economics*, 128(3), 1321–1364.

A Appendix A: Supplementary tables

Table 6: First-stage regressions

Variables	PSI frequency	PSI frequency	PSI frequency	PSI frequency	PSI frequency
	(1)	(2)	(3)	(4)	(5)
Office #1 frequency	0.634***	0.585***	0.626***	0.579***	0.588***
	(0.017)	(0.022)	(0.022)	(0.017)	(0.018)
Ln(# of imports)		-0.000	-0.002	0.003	0.001
		(0.003)	(0.004)	(0.002)	(0.003)
Taxes		0.171***	0.179***	0.222***	0.229***
		(0.035)	(0.035)	(0.042)	(0.045)
Differentiated product		-0.036***	-0.028***	-0.010	-0.019
		(0.011)	(0.010)	(0.016)	(0.017)
Constant	0.344***	0.345***	0.409***	0.405***	0.309***
	(0.016)	(0.014)	(0.065)	(0.053)	(0.020)
Trade partner dummies			Yes	Yes	
HS-2 industry dummies				Yes	
Couple dummies					Yes
Estimator	OLS	OLS	OLS	OLS	OLS
Observations	8,749	7,073	7,073	7,073	7,073
R-squared	0.491	0.492	0.518	0.604	0.677
Adjusted R-squared	0.491	0.492	0.514	0.596	0.630

Notes:
(i) Observations are at the HS-6 product-trade partner level.
(ii) The dependent variable is PSI frequency. Please refer to the Appendix D for a detailed definition.
(iii) Standard errors are in parentheses, clustered by HS-4 products.
(iv) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level.
(v) OLS is Ordinary Least Squares.

Table 7: Robustness check: Ln(# imports from the world)

Variables	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PSI frequency	-0.235**	-0.343***	-0.383***	-0.392***	-0.298*	-0.604***	-0.704***	-0.755***
	(0.105)	(0.106)	(0.118)	(0.136)	(0.166)	(0.178)	(0.198)	(0.211)
Ln(# of imports from the world)	-0.211***	-0.245***	-0.232***	-0.227***	-0.212***	-0.249***	-0.234***	-0.229***
	(0.019)	(0.018)	(0.019)	(0.020)	(0.019)	(0.018)	(0.019)	(0.018)
Taxes	1.487***	1.557***	1.494***	1.537***	1.521***	1.684***	1.641***	1.707***
	(0.201)	(0.218)	(0.280)	(0.304)	(0.223)	(0.249)	(0.313)	(0.323)
Differentiated product	0.466***	0.420***	0.165	0.207	0.458***	0.385***	0.147	0.185
	(0.078)	(0.079)	(0.118)	(0.127)	(0.081)	(0.081)	(0.115)	(0.118)
Constant	0.416***	0.876***	0.819***	0.785***	0.457***	1.136***	0.846**	1.940***
	(0.112)	(0.220)	(0.248)	(0.156)	(0.130)	(0.253)	(0.343)	(0.201)
Trade partner dummies		Yes	Yes			Yes	Yes	
HS-2 industry dummies			Yes				Yes	
Couple dummies				Yes				Yes
Estimator	OLS	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS
Observations	7,073	7,073	7,073	7,073	7,073	7,073	7,073	7,073
K-P Wald F stat					738.231	861.820	1,200.115	1,057.022
R-squared	0.045	0.077	0.094	0.231	0.045	0.076	0.092	0.230
Adjusted R-squared	0.045	0.070	0.075	0.118	0.045	0.069	0.074	0.116

Notes:
(i) Observations are at the HS-6 product-trade partner level.
(ii) The dependent variable is Trade $\text{gap}_{pe} = \ln(\text{Export value})_{pe} - \ln(\text{Import value})_{pe}$.
(iii) Standard errors are in parentheses, clustered by HS-4 products.
(iv) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level.
(v) Couple dummies: HS-2 industry-trade partner dummies.
(vi) OLS is Ordinary Least Squares, 2SLS is Two Stage Least Squares.
(vii) K-P Wald F stat means Kleibergen-Paap Wald F statistics.

B Appendix B: Additional robustness checks

Table 8: Robustness checks, continuation

Variables	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap
	(1)	(2)	(3)	(4)	(5)	(6)
PSI frequency	-0.410***	-0.748***	-0.356***	-0.690***	-0.450***	-0.725***
	(0.136)	(0.210)	(0.134)	(0.206)	(0.115)	(0.176)
Quantity of internal information available	-0.031***	-0.032***	-0.416***	-0.416***	-0.414***	-0.415***
	(0.004)	(0.004)	(0.020)	(0.019)	(0.018)	(0.017)
Taxes	1.475***	1.633***	1.532***	1.687***	1.702***	1.837***
	(0.292)	(0.310)	(0.296)	(0.313)	(0.268)	(0.279)
Differentiated product	0.160	0.142	0.241**	0.229**		
	(0.126)	(0.117)	(0.116)	(0.108)		
Constant	-0.784***	1.545***	2.207***	1.488***	-7.172***	1.630***
	(0.136)	(0.195)	(0.147)	(0.187)	(1.239)	(0.161)
Couple dummies	Yes	Yes	Yes	Yes	Yes	Yes
Alternative internal information variable	Yes	Yes				
Differentiated product: liberal definition			Yes	Yes		
No differentiation variable					Yes	Yes
Estimator	OLS	2SLS	OLS	2SLS	OLS	2SLS
Observations	7,073	7,073	7,073	7,073	8,749	8,749
R-squared	0.222	0.221	0.266	0.265	0.262	0.261
Adjusted R-squared	0.108	0.106	0.158	0.156	0.167	0.166

Notes:
(i) Observations are at the HS-6 product-trade partner level.
(ii) The dependent variable is Trade $\text{gap}_{pe} = \ln(\text{Export value})_{pe} - \ln(\text{Import value})_{pe}$.
(iii) Quantity of internal information available: # of exporting countries (columns (1) and (2)) or Ln(# of imports) (columns (3)-(6)).
(iv) Standard errors are in parentheses, clustered by HS-4 products.
(v) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level
(vi) Couple dummies: HS-2 industry - trade partner dummies.
(vii) OLS is Ordinary Least Squares, 2SLS is Two Stage Least Squares.

Table 9: Robustness checks, restricted sample

Variables	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap	Trade gap
	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
PSI frequency	-0.506***	-0.982***	-0.529***	-0.945***	-0.200**	-0.401***	-0.314**	-0.596***
	(0.127)	(0.198)	(0.132)	(0.209)	(0.090)	(0.146)	(0.145)	(0.216)
Ln(# imports)	-0.393***	-0.393***	-0.399***	-0.400***	-0.255***	-0.255***	-0.441***	-0.439***
	(0.020)	(0.019)	(0.021)	(0.019)	(0.016)	(0.015)	(0.024)	(0.023)
Taxes	2.157***	2.455***	2.171***	2.434***	1.021***	1.116***	1.486***	1.613***
	(0.192)	(0.197)	(0.195)	(0.202)	(0.140)	(0.151)	(0.331)	(0.344)
Differentiated product	0.170	0.136	0.169	0.138	0.178**	0.165**	0.177	0.165
	(0.126)	(0.117)	(0.125)	(0.117)	(0.088)	(0.082)	(0.138)	(0.129)
Constant	-0.376*	1.122***	-9.522***	0.101	1.429***	1.770***	0.671***	1.611***
	(0.192)	(0.151)	(0.912)	(0.155)	(0.105)	(0.113)	(0.166)	(0.221)
Couple dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excluding specific tariffs	Yes	Yes						
Excluding sensitive HS-2 industry			Yes	Yes				
Excluding first and last 0.05 quantile					Yes	Yes		
Excluding HS6 products imported from a neighbouring country							Yes	Yes
Estimator	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Observations	6,683	6,683	6,226	6,226	6,367	6,367	5,723	5,723
R-squared	0.264	0.261	0.258	0.256	0.232	0.231	0.257	0.256
Adjusted R-squared	0.153	0.149	0.151	0.149	0.111	0.110	0.142	0.141

Notes:
(i) Observations are at the HS-6 product-trade partner level.
(ii) The dependent variable is Trade $\text{gap}_{pe} = \ln(\text{Export value})_{pe} - \ln(\text{Import value})_{pe}$.
(iii) Standard errors are in parentheses, clustered by HS-4 products.
(iv) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level.
(v) Couple dummies: HS-2 industry-trade partner dummies.
(vi) OLS is Ordinary Least Squares, 2SLS is Two Stage Least Squares.

Table 10: Robustness checks, Weight gap

Variables	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap
	(1)	(2)	(3)	(4)	(5)	(6)
PSI frequency	-0.714***	-1.269***	-0.663***	-1.219***	-0.743***	-1.164***
	(0.149)	(0.233)	(0.145)	(0.230)	(0.127)	(0.198)
Quantity of internal information available	-0.029***	-0.031***	-0.413***	-0.413***	-0.414***	-0.414***
	(0.005)	(0.005)	(0.024)	(0.022)	(0.022)	(0.020)
Taxes	0.338*	0.595***	0.422**	0.678***	0.376**	0.580***
	(0.202)	(0.230)	(0.204)	(0.234)	(0.180)	(0.199)
Differentiated product	-0.037	-0.066	-0.015	-0.034		
	(0.140)	(0.130)	(0.148)	(0.140)		
Constant	0.683***	-1.796	2.504***	-1.194	-2.791***	-1.240
	(0.149)	(2.170)	(0.170)	(1.979)	-0.845	(1.978)
Couple dummies	Yes	Yes	Yes	Yes	Yes	Yes
Alternative internal information variable	Yes	Yes				
Differentiated product: liberal definition			Yes	Yes		
No differentiation variable					Yes	Yes
Estimator	OLS	2SLS	OLS	2SLS	OLS	2SLS
Observations	6,836	6,836	6,836	6,836	8,431	8,431
R-squared	0.207	0.204	0.243	0.240	0.238	0.236
Adjusted R-squared	0.090	0.086	0.131	0.128	0.138	0.135

Notes:
(i) Observations are at the HS-6 product-trade partner level.
(ii) The dependent variable is Weight $\text{gap}_{pe} = \ln(\text{Exported weight})_{pe} - \ln(\text{Imported weight})_{pe}$.
(iii) Quantity of internal information available: # of exporting countries (columns (1) and (2)) or Ln(# of imports) (columns (3)-(6)).
(iv) Standard errors are in parentheses, clustered by HS-4 products.
(v) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level.
(vi) Couple dummies: HS-2 industry-trade partner dummies.
(vii) OLS is Ordinary Least Squares, 2SLS is Two Stage Least Squares.

Table 11: Robustness checks, Weight gap, restricted sample

Variables	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap	Weight gap
	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
PSI frequency	-0.765***	-1.400***	-0.751***	-1.369***	-0.399***	-0.739***	-0.664***	-1.133***
	(0.148)	(0.239)	(0.156)	(0.250)	(0.101)	(0.158)	(0.160)	(0.243)
Ln(# imports)	-0.382***	-0.383***	-0.384***	-0.384***	-0.222***	-0.222***	-0.450***	-0.447***
	(0.023)	(0.022)	(0.024)	(0.023)	(0.018)	(0.016)	(0.029)	(0.027)
Taxes	0.510**	0.905***	0.593***	0.981***	0.194	0.351**	0.440*	0.649**
	(0.219)	(0.235)	(0.228)	(0.245)	(0.132)	(0.139)	(0.228)	(0.255)
Differentiated product	0.021	-0.025	-0.001	-0.046	0.006	-0.019	-0.001	-0.022
	(0.143)	(0.134)	(0.140)	(0.131)	(0.098)	(0.093)	(0.157)	(0.146)
Constant	-0.364**	-1.211	2.475***	1.467***	-2.264***	0.686	0.885***	-1.175
	(0.158)	(1.972)	(0.185)	(0.179)	(0.624)	-1.524	(0.191)	(1.976)
Couple dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Excluding specific tariffs	Yes	Yes						
Excluding sensitive HS-2 industry			Yes	Yes				
Excluding first and last 0.05 quantile					Yes	Yes		
Excluding HS6 products imported from a neighbouring country							Yes	Yes
Estimator	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Observations	6,451	6,451	6,000	6,000	6,146	6,146	5,523	5,523
R-squared	0.239	0.235	0.234	0.231	0.218	0.215	0.242	0.239
Adjusted R-squared	0.124	0.119	0.123	0.119	0.093	0.090	0.123	0.121

Notes:
(i) Observations are at the HS-6 product-trade partner level.
(ii) The dependent variable is Weight $\text{gap}_{pe} = \ln(\text{Exported weight})_{pe} - \ln(\text{Imported weight})_{pe}$.
(iii) Standard errors are in parentheses, clustered by HS-4 products.
(iv) Coefficient is statistically different from zero at the ***1% , **5%, and *10% level.
(v) Couple dummies: HS-2 industry-trade partner dummies.
(vi) OLS is Ordinary Least Squares, 2SLS is Two Stage Least Squares.

C Appendix C: Variable definitions

Differentiated product: Equals one if the HS6-product is traded on a organised market or listed in a trade publication. Conservative definition (if not specified). Source: Rauch classification of goods.

Office #1 frequency: At the HS-6 product-trade partner level, the fraction of import declarations handled by the customs office #1. Computed from information on the customs office of entry provided by the IT department, local customs.

Export value: Value of exports (U.S. dollars) by the trade partner to country A at the HS 6-digit HS level. The nomenclature is HS 2012. Source: UN COMTRADE.

Import value: Value of imports (U.S. dollars) from the trade partner into country A at the HS 6-digit HS level. Using the exchange rate of the day of the import, values, expressed in local monetary unit, are converted into U.S. dollars.

Ln(# of imports): At the HS-6 product-trade partner level, the Naperian logarithm of the total number of import transactions. Computed from information provided by the IT department, local customs.

Ln(# of imports from the world): At the HS-6 product level, the Naperian logarithm of the total number of import transactions (from the world). Computed from information provided by the IT department, local customs.

PSI frequency: At the HS-6 product-trade partner level, the fraction of import declarations in the original database submitted to a pre-shipment inspection. When a pre-shipment inspection is carried out, a report of findings is sent to the local customs service. This document specifies, among others, the total value, the incoterm, the exchange rate, the classification of the goods .... Information about the presence of a ROF attached to the declaration is provided by the IT department, local customs.

Taxes: Sum of effectively paid taxes — tariffs, excises, value added tax— divided by the value of imports. Sources: IT department, local customs.

D Appendix D: $\frac{\partial \textbf{Trade gap}}{\partial \textbf{ln(\# of imports)}}$

Let:

- k: $\frac{f.o.b}{c.i.f}$ factor,
- $ln(\#)$: ln(# of imports),
- p_M: Imported unit value, incoterm: f.o.b,

- p_X: Exported unit value, incoterm: f.o.b,
- q_M: Imported quantity,
- q_X: Exported quantity.

Hence:

$$\text{Trade gap} = ln(\frac{q_X \times p_X}{q_M \times p_M \times k})$$

We get:

$$\frac{\partial \text{Trade gap}}{\partial ln(\#)} = \frac{\partial q_X}{\partial ln(\#)} \times \frac{1}{q_X} + \frac{\partial p_X}{\partial ln(\#)} \times \frac{1}{p_X} - [\frac{\partial q_M}{\partial ln(\#)} \times \frac{1}{q_M} + \frac{\partial p_M}{\partial ln(\#)} \times \frac{1}{p_M} + \frac{\partial k}{\partial ln(\#)} \times \frac{1}{k}]$$

Since we have:

$$\frac{\partial ln(\frac{q_X}{q_M})}{\partial ln(\#)} = \frac{\partial q_X}{\partial ln(\#)} \times \frac{1}{q_X} - \frac{\partial q_M}{\partial ln(\#)} \times \frac{1}{q_M}$$

$$\frac{\partial ln(\frac{p_X}{p_M} \times \frac{1}{k})}{\partial ln(\#)} = \frac{\partial p_X}{\partial ln(\#)} \times \frac{1}{p_X} - [\frac{\partial p_M}{\partial ln(\#)} \times \frac{1}{p_M} + \frac{\partial k}{\partial ln(\#)} \times \frac{1}{k}]$$

We obtain:

$$\frac{\partial \text{Trade gap}}{\partial ln(\#)} = \frac{\partial ln(\frac{q_X}{q_M})}{\partial ln(\#)} + \frac{\partial ln(\frac{p_X}{p_M} \times \frac{1}{k})}{\partial ln(\#)}$$

It yields:

$$\frac{\partial \text{ Trade gap}}{\partial \ln(\#)} = \frac{\partial \text{ Quantity gap}}{\partial \ln(\#)} + \frac{\partial ln(\frac{\text{Exported unit value}_{FOB}}{\text{Imported unit value}_{CIF}})}{\partial \ln(\#)}$$