Technology (Ab)use and Corruption in Customs

Cyril Chalendard[‡]

Ana M. Fernandes[†]

Gael Raballand[§]

Bob Rijkers[†]

July 2020

Abstract: This paper presents a new methodology to detect collusion in customs and applies it to Madagascar's main port. Manipulation of assignment of import declarations to inspectors is identified by detecting deviations from random assignment prescribed by official rules. Deviant declarations are more at risk of tax evasion, yet less likely to be deemed fraudulent by inspectors, who also clear them faster. An intervention in which inspector assignment was delegated to a third party validates our approach, but also triggered a novel manifestation of manipulation that perpetuated systemic corruption. Tax revenues for deviant declarations would have been 27% higher had their assignment not been tampered with.

Key words: corruption, collusion, tax enforcement, tariff evasion, trade policy

JEL codes: F13, D73, H26

‡ ITC & Development Research Group, World Bank, † Development Research Group, World Bank, § Governance Global Practice, World Bank.

We thank Madagascar Customs and GasyNet for sharing their data, feedback, and willingness to collaborate. We are especially grateful to Éric Rabenja, Dannick Fiononana, Antsa Rakotoarisoa, Cédric Catheline, Stéphane Manouvrier, Prisca Michea, Tolotra Hasinantenaina Ramarosandratana, and Tahina Rabesalama, whose support was critical to the success of this project. Aaditya Mattoo provided extremely useful comments and guidance, especially in the early stages of the project. We also thank Thomas Cantens, Vianney Dequiedt, Simeon Djankov, Henrik Kleven, Beata Javorcik, Joana Naritomi, Tristan Reed, Sandra Sequeira, Joana Silva, Joel Slemrod, René Tapsoba, Shang-Jin Wei, and seminar participants at the World Bank, the Tinbergen Institute, CESIfo, the ABCDE conference, the Midwest Economic Theory and International Trade Conference, and 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) and the Strategic Research Partnership on Economic Development. 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. All errors are our responsibility.

The authors can be contacted at cyril.chalendard@gmail.com, afernandes@worlbank.org, graballand@worldbank.org and brijkers@worldbank.org (corresponding author).

1 Introduction

State capacity to raise tax revenue is a an important enabler of development (Besley and Persson, 2009). Poorer countries mobilize less tax revenue as a share of GDP (Gordon and Li, 2009) and suffer higher levels of corruption. While tax evasion and weak bureaucratic performance are salient drivers of the differences in revenue mobilization across the development spectrum (Finan et al., 2017, Khan et al., 2015 Khan et al., 2019), less is known about who evades, and to what extent evasion is facilitated by (which) bureaucrats. Evidence on the effectiveness of reforms to remedy systemic corruption is also scant.

This paper breaks new ground on these issues. It presents a new methodology to detect and quantify the prevalence and costs of collusion in customs, and assesses the effectiveness of an intervention intended to eliminate such collusion. Across the globe, customs information technology (IT) systems prescribe random assignment of incoming declarations to inspectors, conditional on their productivity (in the task of clearing declarations) as a way to deter corruption. Our approach identifies potential manipulation of inspector assignment by evaluating whether certain inspectors are paired excessively frequently with certain brokers, deviating from what conditional random assignment would predict. To assess whether these deviations reflect collusion, we subsequently examine whether excess interaction between brokers and inspectors is associated with an increased risk of tax evasion and whether deviant declarations are treated preferentially by inspectors. The methodology is validated by studying the impact of an intervention that delegates the (randomization of) inspector assignment to a third party organization external to customs.

We apply our approach to Madagascar's main port, Toamasina, which provides an ideal setting for studying collusion in customs. First, like many other developing countries, Madagascar is heavily reliant on revenues collected at the border (Baunsgaard and Keen, 2010), which account for 48% of total tax revenues. Toamasina collects more than three quarters (78%) of non-oil tax revenues and employs a limited number of inspectors. Each inspector oversees the collection of 1.3% of total yearly taxes in Madagascar. Second, corruption appears rife in customs. A survey of inspectors reveals that only 6% believe non-ethical conduct is sanctioned, and only 23% believe their colleagues act with integrity. Third, inspectors repeatedly interact with a limited number of brokers, with whom they also share social ties. The combination of high stakes, a small number of players, limited sanctions for improper conduct and extensive repeated interactions is conducive to collusion. Last but not least, Madagascar's senior customs management were willing to undertake reforms to curb collusion and provided us unprecedented data access. They shared data for the period 2015-2018 covering rich details on each import declaration including declared value and weight, weight measured upon arrival at the port, taxes paid, the identity of the broker which registered it and the inspector assigned to it, whether fraud was recorded, all revisions to inspector assignment, value, weight, and tax liability made during the clearance process, as well as risk management information (inspection channel, risk scores, and valuation advice).

Our methodology comprises three steps. First, we detect potential manipulation of inspector assignment by identifying pairings of brokers with inspectors that occur much more frequently than would be expected on the basis of conditional random assignment. In Toamasina, 9.7% of all declarations are handled by inspectors whose assignment contravened the random inspector assignment prescribed by official rules. Second, these deviant declarations are shown to have characteristics commonly associated with an elevated risk of tariff evasion and to embody sizeable potential tax revenue losses. Third, we demonstrate that inspectors treat preferentially the declarations registered by brokers with whom they interact excessively frequently ceteris paribus. They clear them faster, are less likely to deem them fraudulent, and impose lower weight, value and tax adjustments, thus exacerbating disparities in tax revenue losses between deviant and non-deviant declarations. These findings are obtained in stringent econometric specifications that control for a rich set of declaration characteristics capturing the risk of tax evasion, as well as broker and inspector fixed effects. According to back of the envelope calculations average tax revenue per declaration non-randomly assigned would have been 27% higher in the absence of excess interaction. Total tax revenues collected in Toamasina would have been 4% higher.[1]

We argue that these patterns are consistent with a collusion scheme in which brokers bribe staff in the customs information technology (IT) department and/or the customs port manager to be paired with their preferred inspector, who agrees to clear the declarations that are the object of collusion faster, not to impose fines and penalties, not to insist on upward adjustment (or to request just a marginal one) of the customs declared value. The resulting tax savings are presumably shared with inspectors. Although we do not directly observe bribe payments, our findings are consistent with extensive circumstantial evidence collected during repeated field visits, IT audits, focus group discussions, and a survey of customs inspectors. Based on our findings, Madagascar's customs management felt compelled to sanction inspectors, to suspend the head of the IT department and to reform inspector assignment by divesting it to a third party outside customs. This re-randomization

[1] As explained in Section 8, these estimates do not reflect the overall cost of corruption in customs, but only the tax revenue losses associated with the specific collusion scheme we document.

intervention was so successful in eliminating deviations from random inspector assignment that it became standard practice.

Explanations other than collusion are difficult to reconcile with the totality of the observed patterns. Nonetheless, we demonstrate that the results are not driven by "familiarity" between inspectors and brokers nor by inspectors being sought out because they specialize in the clearance of different goods or have different levels of expertise in evaluating different shipments. By allowing for potential collusion between inspectors and importers in some specifications we show that brokers, rather than importers, seem to be the protagonists of the collusion scheme we unveil. This may be due to the fact that they handle more shipments (and thus have more to gain from collusive agreements), and interact much more frequently with inspectors than importers do. In addition, lobbying customs officials is the core of a broker's business in countries with weak governance.

Alternative explanations also fail to explain why the re-randomization intervention virtually eliminated the prevalence of deviations from random inspector assignment. IT manipulation resurfaced after a few months, however, albeit in a different guise. Customs IT staff figured out a new way manipulate inspector assignment and bypass the re-randomization. This bypasssing was identified by assessing whether the entire set of declarations registered by brokers was shared with the third party for inspector random re-assignment. We show that 7.2% of all import declarations were withheld from re-randomization.[2] The circumvention of the re-randomization not only attests to the difficulties inherent in the dislodging systemic corruption but also provides quasi-experimental variation in exposure to the re-randomization intervention.

The bypassing resulted in the resurgence of excess interaction between inspectors and brokers, driven exclusively by withheld declarations.[3] Interestingly, withheld declarations were disproportionately assigned to inspectors with whom brokers had interacted excessively frequently in the period before the re-randomization intervention, suggesting persistence in the collusion schemes we unveil. These withheld declarations were on average more risky, subject to higher taxes, more undervalued, and embodied larger tax revenue losses, especially when their eventual (non-random) assignment resulted in excess interaction between inspectors and brokers. Inspectors only provide preferential treatment to withheld

[2] In practice, such bypassing appears to have been the result of the temporary disabling of a re-randomization trigger, such that all declarations registered during specific time intervals when this trigger was deactivated were withheld from being sent to the third party to be re-randomized (including those that were the subject of a collusive agreement).

[3] Excess interaction was not observed for declarations handled by inspectors whose assignment was re-randomized.

declarations if registered by brokers with whom they interact excessively frequently. These findings validate our methodology and interpretation that the documented patterns reflect collusion.

To start with, we contribute to the literature on the measurement of corruption (Bardhan, 1997, Olken and Pande, 2012) and its development consequences (Shleifer and Vishny, 1993, Shleifer and Vishny, 2002), as well as the nascent literature on the performance of bureaucrats as a determinant of state effectiveness and tax collection (Olken and Pande, 2012; Dincecco and Ravanilla, 2017; Pepinsky et al., 2017; Xu, 2018; Xu et al. 2018). The overhwelming majority of inspectors in Toamasina took part in the collusion scheme, which was enabled by staff in the IT department. The systemic nature of the corruption helps explain why dislodging it proved so difficult, and why institutional change is typically slow. Our results also contribute to the smaller literature on the effectiveness of anti-corruption interventions (e.g., Ferraz and Finan, 2008; Niehaus and Sukhtankar, 2013) by demonstrating that IT solutions can help curb corruption (see also Lajaaj et al., 2019), but are not a panacea (see also Casaburi et al., 2019), because they can also serve as a conduit to it.

Second, we contribute to the literature on the determinants of tax enforcement (Kleven et al., 2011; Pomeranz and Vila Belda, 2019; Slemrod, 2019), and specifically the literature on tariff evasion (Bhagwati, 1964; Fisman and Wei, 2004; Yang, 2008; Dutt and Traca, 2010; Sequeira and Djankov, 2014; Rijkers et al., 2017; Sequeira, 2016; Wier, 2020) by pinpointing which brokers and inspectors cheat, and which import declarations are most likely to be undervalued. Our focus on brokers as protagonists of corruption schemes is novel, given that their practices are rarely studied in spite of their near ubiquity in international supply chains. Our uniquely rich data enable construction of declaration-specific estimates of tariff evasion. These permit more precise quantification of the tax revenue losses associated with corruption than is possible based on aggregate product-level data, which remain the norm in the literature.

Our results are also relevant for the understanding of trade costs and market distortions in developing countries (Atkin and Khandelwal, 2019). As pointed out by Sequeira (2016), in the presence of corruption, tariffs and other import taxes may not be as burdensome as they appear on paper. We complement her findings by showing that corruption not only impacts tax enforcement and *de facto* tariff incidence, but also another margin of trade costs, notably clearance times. Firms participating in collusion schemes gain a double advantage over their competitors, being able to import goods both faster and at lower cost.

Finally, our paper contributes to the growing field of forensic economics (see e.g., Jacob and Levitt, 2003), reviewed in Zitzewitz (2012), by evaluating the effectiveness of random assignment as a means of deterring corruption. While our study focuses on customs, random assignment is used to prevent corruption in a plethora of settings including the assignment of cases to judges and prosecutors.[4] We believe our approach can fruitfully be adapted to these other contexts.[5]

The remainder of this paper is organized as follows. Section 2 describes the context and the customs clearance process while Section 3 presents our data. Section 4 describes our methodology to detect deviations from official rules in inspector assignment to declarations. Section 5 examines whether deviant declarations are at a higher risk of tax evasion. Section 6 assesses whether there is differential treatment of deviant declarations by inspectors. Section 7 validates our approach by analyzing the impact of the re-randomization intervention. Section 8 provides estimates for the costs of collusion in terms of tax losses. Section 9 concludes.

2 Context: Customs Clearance Process in Madagascar

This section describes the customs clearance process and argues that the conditions in Toamasina are conducive to collusion: there are few players who interact repeatedly, the stakes are high, and there is almost no punishment for improper conduct.

Taxes and duties collected by customs accounted for 48 percent of overall tax revenue in Madagascar in 2019, despite substantial tariff evasion (Chalendard et al., 2019). Most of this revenue was collected in Toamasina, which accounted for 78% percent of non-oil tax revenue and 52% percent of non-oil imports and employed on average 17 inspectors per year during our sample period. Each inspector oversees the collection of 17 million USD worth of tax revenue per year on average, representing 1.3% of total taxes collected.

Jobs in the customs administration - especially inspector jobs in Toamasina - are among the most sought-after jobs in Madagascar. They are secure, well-paid, and offer several benefits. Inspectors earn a salary of roughly 11,000 USD per year (21 times annual GDP per capita of 527 USD) and receive as bonus 5 to 20 percent of the fines they issue. They can also earn performance bonuses of up to 1,000 USD per quarter if they are among the top inspectors in terms of clearance speed, fraud detection, and tax revenue mobilization.

[4] Around the world 162 countries have adopted random assignment of cases to judges to deter judicial corruption (Doing Businesss, 2020)

[5] In fact, random assignment has been exploited to study discrimination (Price and Wolfers, 2010; Di Tella and Schargrodsky , 2013), judicial outcomes and their effects (Aizer and Dyle, 2015; Cohen and Yang, 2019; Bhuller et al., 2020). To our knowledge we are the first to exploit inadherence with random assignment to detect corruption.

However, these performance rewards may not sufficiently incentivize inspectors to act with integrity. Corruption appears pervasive, possibly due to the virtual absence of sanctions for improper conduct and to threats from economic operators, and because compensation is low relative to opportunities for graft (Chalendard et al., 2020). According to a nationwide survey of inspectors that we conducted in 2017, only 23% believe that their colleagues act with integrity, only 6% claim non-ethical behavior is sanctioned and only 12% believe promotions are merit-based. Close to a third of inspectors claim being subjected to threats from economic operators on a regular basis.[6] Undervaluation of imports - which results in tax evasion - was widely agreed to be the main type of customs fraud in Madagascar.[7]

The inspectors in Toamasina interact with a limited number of brokers (*commissionnaires agrées en douane*). In a typical semester, there are on average 51 brokers who each handle 181 import declarations, on behalf of 34 different importers.[8] Brokers must have a license, which is issued by the customs administration and they administer the customs clearance process on behalf of the importer by fulfilling customs formalities and submitting documentation. They are accountable for the payment of taxes, duties and potential penalties and are penalized (with a fine) in case of non-compliance. In principle, repeated non-compliance can result in the revocation of the broker's license. In practice, suspension of brokers due to misconduct is rare. Customs officials and brokers frequently socialize in the small town of Toamasina and many brokers either have served as customs officials themselves, or deliberately recruit former customs officials because of their expertise and networks. Thus there is extensive repeated interaction between inspectors and brokers, both inside and outside of the customs premises.

There is significant information asymmetry between importers and brokers given that the latter are much better informed about customs procedures and are the first point of contact for customs in case disputes arise. Some brokers have transparent pricing schemes which typically depend on the size and contents of the cargo, but others charge a fixed amount (inclusive of potential tax liabilities) per container cleared, irrespective of its content, which implies that their profits directly depend on the amount of tax they remit on behalf of the importer. To understand how collusion may happen it is instructive to consider the customs clearance process, a stylized version of which is depicted in Figure 1.

[6]There is a strong *esprit de corps*, with the majority of inspectors feeling proud to work for customs. During personal interviews, many of them expressed strong feelings of loyalty towards their colleagues.

[7]Administrative data on fraud records classify 67.2% of all fraud in Madagascar customs as underreporting of value, 27.4% as underreporting of quantities, and the remainder as product misclassification (4.9%) or misreporting country of origin (0.5%).

[8]These averages do not consider small brokers (i.e., those handling less than 50 declarations per semester) since they will not be part of our estimating sample (described in Section 3).

1. **Registration.** The first step in the process is the electronic registration of an import declaration by the broker on behalf of the importer via the Automated System for Customs Data (ASYCUDA)++ customs clearance IT system.[9]

2. **Risk analysis.** The second step consists in risk analysis conducted by both GasyNet, a third party service provider that assists Madagascar customs with risk analysis and logistics, and the customs risk management unit.[10] For each declaration, (i) a risk score is issued based on GasyNet's proprietary risk model, (ii) a clearance channel is recommended along with a qualitative justification. If the yellow channel is selected the inspector only needs to check the documentation. If the red channel is selected the inspector is expected to physically inspect the cargo. However, the inspector is at liberty to change the clearance channel based on her own judgment. In addition, (iii) for a very small subset of high-risk declarations for which the accuracy of the declared import value is questionable, GasyNet issues a valuation advice: a detailed report on what the value of the specific declaration is likely to be.

3. **Inspector assignment.** The third and, for our purposes, crucial step is the assignment of the declaration to a particular inspector by the ASYCUDA IT system. Official rules prescribe that a newly registered declaration should be assigned to whichever inspector has the lowest workload (i.e., has the fewest pending declarations on his/her desk) and is active (i.e., is connected to the IT system and can therefore receive new declarations). Official rules do allow for productivity differences across inspectors: a highly productive inspector will get, on average, more declarations than a poorly productive inspector. Yet, the assignment of declarations to inspectors is supposed to be random conditional on her/his productivity. We will exploit this crucial feature of the official rules for identifying deviant (collusive) declarations in Section 4.

   However, the customs port manager, the *Chef des Opérations Commerciales* (COPCO), has the authority to override the IT system's initial assignment and re-assign a declaration to a different active inspector. Such re-assignments are warranted in case of unanticipated absenteeism (due to illness or because inspectors simply fail to show up on time, or at all) and, should, *a priori*, happen only randomly.[11]

4. **Assessment.** The fourth step is the assessment of the declaration by the assigned

[9] ASYCUDA is an integrated customs management system developed by United Nations Conference on Trade and Development (UNCTAD) that has been adopted by more than 90 countries.

[10] In reality the second step (risk analysis) and the third step (inspector assignment) happen simultaneously.

[11] Such reassignments occur for 6% of the import declarations.

inspector based on the documentation submitted by the broker on behalf of the importer, the risk analysis diagnostics provided by the risk management unit and GasyNet, and the results of a potential physical inspection. She has to decide which (if any) adjustments to the import value, quantity and/or product classification are to be made and report whether fraud was perpetrated. She then assesses what duties, taxes and potential penalties are to be paid based on the (potentially revised) final value and product classification of the import declaration.

5. **Clearance.** In the final step in which goods are cleared, the importer (or the broker on behalf of the importer) pays the taxes, duties and potential penalties and goods are released from customs.

Our analysis of collusion will focus both on manipulation of the assignment of declarations to inspectors (by IT department staff and/or the customs port manager) done in step 3, and on differential treatment of manipulated declarations by inspectors during assessment in step 4.

3 Data

Our study combines the following databases.

- **Customs transactions data** From Madagascar's customs administration we obtained highly disaggregated administrative data tracking imports at the transaction level for the period January 2015-November 2018. For each import declaration, the data covers the HS 8-digit products included (designated as items), their source country, the dates/times of registration, inspector assignment, assessment, and clearance, the broker, the importing firm, and, crucially, the customs inspector assigned to handle the declaration. The data also contain a number of unique variables that are important for our analysis. For each item, they contain information on both the initially declared and the finally registered import value, weight, and taxes paid (tariff and value added tax as well as exemptions). These variables enable us to evaluate inspector modifications of value, weight, and tax liabilities. In addition, for each declaration we can track any modifications made to the IT system's initial inspector assignment by the customs port manager.[12] This will allow us to disentangle the role

[12]This data was obtained from the customs administration's internal control systems and was merged to the administrative customs data.

of IT department staff from that of the customs port manager in generating deviations from official rules in inspector assignment.

- **Fraud records** We obtained fraud records from from the Legal Department (*Service des Affaires Juridiques et du Contentieux*). For each declaration, we know both whether and if so what type(s) of fraud was detected and the amount of taxes recovered (if any). Information on whether and how much inspectors modified tax yield is important for assessing the role of inspectors and, to the best of our knowledge, has not been exploited in prior literature on tariff evasion.

- **Risk management data** From the customs risk management unit we received for each import declaration information on the initial and finally-used clearance channel (documentary control/yellow channel, physical nspection/red channel or no inspection/blue channel). From GasyNet we received the risk score assigned to each import declaration (related to the risk of non-compliance with customs regulations ranging from 1 to 9) and valuation advice in case it was issued.

- **(Objective) container weight data** We obtained from the company in charge of managing Toamasina's container terminal - Madagascar International Container Terminal Services Limited (MICTSL) - data on the weight of containers that arrive in Toamasina as measured by weighing at a scale upon arrival for the period 2015-2017. This port authority weight data is merged to the customs data at the declaration level, for declarations whose goods fill completely one or more containers. For declarations that share containers with other declarations this information is missing. These port authority weight data provide a very useful benchmark for verifying whether the weight registered by the broker is correct.

- **UN COMTRADE data** We also rely on an international trade data source UN COMTRADE to obtain export flows - values and quantities (weight) - at the country-HS 6-digit-year level for all of Madagascar's trading partners in 2015-2018. We use this mirror data for flows being imported by Madagascar to construct exogenous benchmark/reference prices to which we will compare the unit prices of the items included the import declarations in the Madagascar customs data (as will be described below).

- **Re-randomization of inspector assignment and IT manipulation** On November 18, 2017 the assignment of inspectors to declarations was delegated to GasyNet.

By comparing daily their list of declarations (that their system randomly assigned to some inspector) to the list of declarations that cleared customs from the customs administration, GasyNet was able to identify declarations that were withheld from the re-randomization - as will be discussed in Section 7. They provided us with the list of withheld declarations.

Madagascar's raw customs data covers all formal import transactions made under several regimes: final imports for consumption (imports for home use), re-imports, temporary admissions, inward processing, warehouse, and other. Our analysis focuses on import declarations that are subject to taxation and to a physical or a documentary control by customs inspectors in Toamasina.[13] This implies focusing only on imports for home use and re-imports and excluding declarations from importers that are members of the "*Procédure Accélérée de Dédouanement*" (PAD), a trusted trader program that allows member firms to benefit from expedited clearance procedures with minimal controls at the border. In order to increase the relevance of our collusion proxy, we remove from the sample (i) declarations registered by brokers that do not interact frequently with customs (i.e., brokers that register less than 50 declarations per semester); (ii) declarations assigned to inspectors that worked less than two months in Toamasina over the course of a semester. Our final estimating sample accounts for an average of 76.9% of declarations, 78.9% of collected taxes, and 76.5% of total import value for import declarations subject to taxation and to a physical or a documentary control cleared in Toamasina, across the period ranging from January 1 2015 to November 17 2018.[14]

To analyze which declarations are most likely to be subject to collusive agreements we will use measures of excess interaction between inspectors and brokers as proxies for IT manipulation described in Sections 4 and 7. The definition of all variables is provided in Appendix Tables 10, 11, and 12. Here we briefly describe the declaration-level customs outcomes on which we will estimate the impacts of collusion. These are clearance time (measured as the log number of hours from the time the declaration was (last) assigned to an inspector to her assessment of the declaration), a dummy for whether or not fraud was recorded, the change in log value (finally registered - initially declared), tax adjustment, and hypothetical tax revenue losses described below. As additional declaration-level customs outcomes used in robustness exercises we consider: the change in log weight (finally registered - initially declared) and the gap between the port authority weight and the initially declared

[13]Imports subject to specific clearance procedures (oil and vehicles) are excluded.

[14]Our sample ends one year after the start of the re-randomization of inspector assignment to the third party organization and a few days before the unveiling of the IT manipulation of the re-randomization.

weight (for simplicity called weight gap).

Hypothetical tax revenue losses for a declaration are computed based on the difference between hypothetical tax yield and actual tax yield. Measuring hypothetical tax yield is notoriously challenging given that it is unobserved. Our baseline measure of a declaration's hypothetical tax yield considers as a reference price for each of its items the median unit price (ratio of value to weight) reported across Malagasy importers for the same origin country and year. For each item included in the declaration the relevant reference price is multiplied by the item's weight and the item' actual tax rate. Summing the resulting hypothetical item-level tax yield across all items included in the declaration yields the declaration-level hypothetical tax yield. This is a very conservative measure, for it assumes that the median unit price is not itself under-reported. Our alternative measure of a declaration's hypothetical tax yield considers as a reference price for each of its HS 6-digit products the unit price reported by the exporting country in that year in UN COMTRADE multiplied by the products' weights and by the products' actual tax rates and sums these across all products in the declaration.[15] This measure has the advantage of using prices that are more likely to be exogenous to tax evasion in Madagascar.[16]

Two additional measures of hypothetical tax revenue losses are constructed for two subsets of declarations. For declarations for which port authority weight data is available, hypothetical tax yield is constructed also correcting for underreporting of quantities assuming that the measured port authority weight is correct.[17] For declarations for which valuation advice was issued, hypothetical tax yield is constructed as the declaration's reference value multiplied by the average tax rate.

As determinants of collusion (and subsequently as controls for evasion risk) we rely on the following *ex-ante* risk characteristics of the import declaration: the tax rate (tariffs and other taxes), the risk score, a dummy for the red channel, a dummy for being a mixed shipment (i.e., one that includes different items), the share of differentiated products as per Rauch (1999)'s classification, and a dummy for receiving GasyNet's valuation advice. In robustness exercises we consider other declaration characteristics: the log of the initially

[15] An HS 6-digit product's weight is obtained by summing across the weights of all the corresponding items. An HS 6-digit product's tax rate is obtained as the ratio between the sum of actual taxes and the sum of finally declared import value across all corresponding items

[16] It is possible that the firms behind a given export flow collude with importers in Madagascar and issue fake invoices for importers to present to customs to minimize their tax liabilities. In addition, export unit prices may be downward biased since they are typically recorded as Free On Board (FOB) whereas import prices are recorded Cost Insurance Freight (CIF) and therefore include transportation and insurance costs.

[17] We cannot correct quantities declared at the item level since the port authority weight is available only for the declaration as a whole. By implication we are assuming that the weight of all items in a given declaration is underreported to the same extent.

declared value, the log of the initially declared weight, the initial unit price relative to median import unit price and the initial hypothetical tax revenue loss (using as reference price the median import price). Summary statistics on all customs outcomes and declaration characteristics are shown in Appendix Tables 13 and 14.

4 Identifying Deviant Declarations

Our identification of declarations suspect of collusion relies on detecting deviations from official rules in the assignment of incoming declarations to customs inspectors. Recall from Section 2 that according to official rules, incoming declarations should be randomly assigned to inspectors conditional on their productivity.

These official rules imply that the share of all declarations that a given inspector handles, which we will refer to as her inspection share (analogous to the concept of "market" share in industrial organization) can vary across inspectors, depending on their productivity, but should not vary systematically across brokers, unless inspector assignment was manipulated. All inspectors should have, for a given broker, an inspection share close to their average inspection share.

To assess whether this is indeed the case we consider the import declarations registered by a specific broker during a semester (this defines a "market") and we define the inspection share of an inspector as the percentage of its declarations handled by that inspector:

$$S_{ibt} = \frac{n_{ibt}}{\sum_{i=1}^{k} n_{ibt}} \tag{1}$$

where n_{ibt} is the number of declarations registered by broker b in semester t, assigned to inspector i and k is the total number of inspectors active in a given semester t.

We compare the observed inspection share in Equation (1) to the expected inspection share that would arise if the assignment of import declarations to inspectors respected the official rules, which we predict using a multinomial distribution.[18] We define inspector productivity p_{it} to be the ratio between the total number of import declarations handled by inspector i in semester t and the total number of import declarations registered in Toamasina

[18] An alternative strategy to identify deviations from official rules would have been to calculate the workload for each inspector and to evaluate whether an incoming declaration is indeed assigned to the active inspector with the lowest workload when it is registered. This would in principle enable us to identify which specific declarations were non-randomly assigned. Unfortunately, the IT system does not keep a log of which inspectors were connected at what time, which makes implementing this strategy infeasible.

in semester t (where $\sum_{i=1}^{k} p_{it} = 1$). The probability of observing a particular distribution $(X_{1bt}, X_{2bt}, ..., X_{kbt})$ of declarations of a given broker b across inspectors $1, 2, ...k$ in semester t (where $\sum_{i=1}^{k} X_{ibt} = n_{bt}$, the number of declarations in semester t registered by broker b) given their productivities $(p_{1t}, p_{2t}, ..., p_{kt})$ is:

$$(X_{1bt}, X_{2bt}, ..., X_{kbt} | p_{1t}, p_{2t}, ...p_k) = \frac{n_{bt}!}{\prod_{i=1}^{k} X_{ibt}!} \prod_{i=1}^{k} p_{it}^{X_{ibt}} \tag{2}$$

By implication, what we are assuming is that for each inspector i, the probability of receiving X_{ibt} import declarations from the number of declarations registered by broker b in semester t follows a binomial distribution with probability p being equal to p_{it}, which is the inspector share of the total number of import declarations in the semester. This enables us to assess for each broker and inspector whether the observed inspection share is statistically significantly different from the predicted inspection share under random assignment.

Figure 2 shows overlaid histograms of the observed distribution of the share of declarations of a given broker cleared by a specific inspector (the darker bars) and the predicted distribution expected if official rules were adhered to (the lighter bars). The equality of these two distributions is rejected at the 1% significance level. Clearly, the observed density distribution of inspector shares by broker has a significantly higher dispersion and more mass in the upper tail than the predicted distribution. This implies that, relative to the distribution of expected inspection shares, the observed assignment of declarations is characterized by excess interaction between some inspectors and some brokers.

Our measure of potential manipulation of inspector assignment is the deviation between actual assignment and random assignment of declarations to inspectors. Specifically, we define the *excess interaction share* ES_{ibt} as the difference between the share of broker b's declarations handled by inspector i in semester t (S_{ibt}) and the predicted share ($\overline{S_{ibt}}$) she would be expected to handle if declaration assignment to inspectors followed official rules:

$$ES_{ibt} = S_{ibt} - \overline{S_{ibt}} \tag{3}$$

The excess interaction share varies across pairs of inspectors and brokers within a semester but all declarations of a given broker handled by a particular inspector are characterized by the same excess interaction share. Inevitably some of those declarations may not have been manipulated but will be characterized by excess interaction, which implies that we may be overestimating the prevalence of manipulation of inspector assignment but underestimating differences between manipulated and non-manipulated declarations.

For some of the analysis done subsequently in the paper it is useful to define an indicator for *significant excess interaction* which takes the value of one if (i) the excess interaction share exceeds 0.05 and (ii) the null hypothesis that the observed inspection share does not exceed the predicted inspection share (i.e., that declarations were randomly assigned to inspectors) can be rejected at the 0.001 percent significance level and of zero otherwise.[19]

Table 1 documents the prevalence of non-random assignment in Toamasina over the period ranging from January 1 2015 to the start of the delegated re-randomization to the third party, November 18 2017. For 9.7% of declarations the inspector handling them interacts significantly more frequently with a broker than would be predicted based on conditional random assignment according to this definition. Prima facie, this is evidence of deviations from official rules in the assignment of import declarations of a given broker across inspectors and thus of potential collusion.

Who is responsible for this non-random assignment: the IT team that manipulates the IT system's initial assignment or the customs port manager who manually and voluntarily erases the initial assignment and reassigns declarations? To answer this question, Figure 3.a plots the density distributions of the initial inspector assignment made by the IT system (the short-dashed line) as well as the final assignment (the long-dashed line) which reflects both the initial assignment and potential re-assignments of declarations to inspectors made by the customs port manager. The distributions of initial and final inspector assignment are very similar, and both deviate markedly from the predicted distribution that would be observed if official rules were adhered to (the dotted line). In fact, for 9.3% of all declarations the inspector initially assigned to handle them appears to interact excessively frequently with the broker, relative to what official rules would predict. Thus, manipulation of the IT system appears to be the predominant driver of non-random assignment.

However, this does not mean that the customs port manager is not part of the collusion scheme. Focusing *only* on declarations re-assigned by the customs port manager, Figure 3.b reveals that these re-assignments exacerbate, rather than reduce, non-random assignment. Instead of offsetting collusion, the port manager appears to be reinforcing it. If he were

[19] The 0.05 threshold corresponds to (approximately) a standard deviation in the observed distribution of excess interaction shares. While conservative, this threshold ensures that the indicator only focuses on inspectors with substantial excess interaction (for which we are confident that the observed excess pairing is not accidental).

to choose inspectors randomly when reassigning declarations, we would expect the final distribution to have been less skewed. Quantitatively however, his actions account only for a small share (4%) of overall non-random assignment.

The fact that certain brokers' declarations are not randomly assigned to inspectors was confirmed in inspector interviews in Toamasina. One inspector mentioned *"I have been here 7 months, but there are certain brokers whose declarations I have never handled"*. Another complained *"I never get the good declarations"*. Our interpretation that such non-random assignment results from IT manipulation is consistent with the remarks by an external auditor of Madagascar's customs IT system of an *"over-reliance on IT administrator account, which is typically used at most a few times a year to make major systemic changes, but was used multiple times a day in Madagascar. The IT administrator account allows you to override basic settings."* and of *"...surprising and suspiciously long queues outside the office of the head of the IT department, which normally is not a client-facing function"*.

Based on the findings of an early incarnation of this paper a number of customs inspectors were sanctioned and removed from their posts. The assignment of declarations was divested to the third party GasyNet, who agreed to re-randomize the assignment of declarations to inspectors, which constitutes a unique opportunity to assess whether we are indeed identifying IT manipulation that we will exploit in Section 7.

5 Do Deviant Declarations Exhibit a Higher Risk of Tax Evasion?

Are the declarations treated by inspectors who interact excessively frequently with a broker special? If excess interactions were the product of accidental deviations from official rules in inspector assignment, then the characteristics of these declarations should not systematically differ from those of other declarations. In contrast, if excess interactions are the product of deliberate IT manipulation to assign a specific declaration to a preferred inspector with whom the broker has a collusive agreement, then a higher risk of customs fraud, which would indicate higher susceptibility to tax evasion, would be expected for such declarations.

On average, declarations characterized by higher excess interaction shares have higher risk scores and are subject to higher tax rates, as is shown in Figures 4.a and 4.b, which present polynomial plots of these risk characteristics against excess interaction. By contrast, initial unit prices relative to median import unit prices tend to fall with the excess interaction share, as shown in Figure 4.c, suggesting that declarations of brokers that interact excessively with some inspectors are more likely to be undervalued. Excess interaction is indeed correlated with hypothetical tax revenue losses calculated on the basis of the initial registration of the declaration by the broker (that is, before the inspector assesses the declaration and insists on potential modifications), as shown in Figure 4.d.

Table 2 presents estimates of unconditional bivariate ordinary least squares (OLS) regressions of declaration characteristics commonly associated with tax evasion on the excess interaction share. A 10 percentage point (0.10) increase in the excess interaction share is associated with a 3.1 percentage point higher tax rate, an increase in the risk score of half a point, a 7.8 percentage point increase in the probability of the declaration being mixed, a 5 percentage point increase in the share of the declaration's value accounted for by differentiated products, a 9.4 percentage point increase in the probability of valuation advice being issued, and a 5.9 percentage point decrease in the initial price relative to median import price. These significantly lower initial prices may explain why the excess interaction share is not significantly correlated with the initially declared value, despite being associated with a higher initially declared weight.

Put simply, declarations characterized by excess interaction have characteristics commonly associated with an elevated risk of tax evasion. A 10 percentage point increase in the excess interaction share is associated with a 6.3 percentage point increase in hypothetical tax revenue losses. Appendix Table 15 presents regressions examining the determinants of the excess interaction share. The estimates suggest that the tax rate and initial undervaluation are among the most salient predictors of deviations from conditional random assignment of inspectors to declarations.

6 Are Deviant Declarations Treated Differently?

This section assesses whether inspectors treat the deviant declarations differently - in a preferential manner - from other declarations. If excess interactions were accidental, then inspectors should provide no differential treatment to deviant declarations, beyond the increased scrutiny that may be legitimately expected as these declarations were shown to be at a higher risk for tax evasion in Section 5. Similarly, if IT department staff was simply bribed to assign certain declarations to the least competent inspector, we would not necessarily expect the chosen inspector to treat manipulated declarations any differently from the way she handles other declarations. Inspectors complicit in a collusive agreement, by contrast, would plausibly provide, in exchange of a bribe, preferential treatment to manipulated declarations.

To assess whether inspectors treat deviant declarations - those of brokers with whom they interact excessively - differently than other declarations, the following specification is estimated by OLS:

$$Y_d = \beta_E ES_{ibt} + \beta_X X_d + \mu_i + \nu_b + \kappa_c + \pi_p + \tau_m + \epsilon \qquad (4)$$

where Y_d is one of the declaration-level customs outcomes described in Section 3 (clearance time, fraud records, value and tax adjustments, hypothetical tax revenue losses). The main regressor of interest is the excess interaction share ES_{ibt} defined in Section 4. The vector of declaration characteristics X_d includes the tax rate, the risk score, a dummy for the red channel, a dummy for being a mixed shipment, the share of differentiated products, a dummy for GasyNet's valuation advice as well as inspector fixed effects μ_i, broker fixed effects ν_b, HS 2-digit product fixed effects π_p, source country fixed effects κ_c, and month-year fixed effects τ_m. The independent and identically distributed (i.i.d) error is ϵ.

The inclusion of inspector fixed effects accounts for heterogeneity across inspectors in their average productivity, ability, work ethic, and other time-invariant characteristics that may impact their performance. Similarly, broker fixed effects account for heterogeneity in their import patterns, efficacy, record-keeping, and other characteristics that may impact customs clearance. This specification is thus very stringent, in that it is identifying whether the excess interaction share has predictive power for customs outcomes even after controlling for average differences across inspectors and brokers. Standard errors are clustered by inspector.[20]

6.1 Main Findings

The results from estimating Equation (4) are shown in Table 3. Inspectors assess declarations registered by brokers with whom they interact excessively frequently significantly faster than other declarations. Column (1) implies that a 10 percentage point increase in the excess interaction share is associated with a 20 percentage point (or approximately a 4-hour) reduction in clearance times. Declarations characterized by excess interaction are also less

[20] Due to the inclusion of a large set of fixed effects, our estimates are obtained using the reghdfe Stata command drawing on Guimaraes and Portugal (2010). The current version of the command eliminates from the number of observations singletons and adjusts standard errors for their exclusion. A singleton is an observation unique in the sample in having a given fixed effect equal to one: e.g., a singleton would be a declaration with imports from source country A if no other declaration reports importing from country A.

likely to be deemed fraudulent: column (2) shows that a 10 percentage point increase in the excess interaction share is associated with a 2.8 percentage point reduction in the likelihood of fraud being recorded. This is a large effect given that the unconditional probability of fraud being recorded is 8 percentage points (see Appendix Table 13).

In the same vein, Columns (3) and (4) show that value and tax adjustments are significantly lower for declarations characterized by excess interaction. A 10 percentage point increase in the excess interaction share is linked to a 0.8 percentage point lower increase in value and a 0.9 percentage point lower increase in tax yield. These are again sizeable effects given that the unconditional averages of value and tax adjustment are 2 percentage points. The significantly lower likelihood of the tax burden being revised upwards is perturbing since declarations characterized by excess interaction are more likely to be undervalued to start with, as shown in Section 5. Inspectors thus seem to exacerbate, rather than reduce, the disparities between declarations characterized by excess interaction and other declarations. As a result, excess interaction is associated with sizeable tax revenue losses. Column (5) implies that a 10 percentage point increase in the excess interaction share is associated with a tax revenue loss of 3.9 percentage points.

In summary, inspectors treat the declarations of brokers with whom they interact excessively frequently preferentially: they clear these declarations more quickly and subject them to significantly laxer tax enforcement. If inspectors were honest, no preferential treatment should be observed.

6.2 Alternative Explanations

This section evaluates the most salient alternative explanations for the findings of differential preferential treatment of deviant declarations by inspectors by running a set of additional tests. To start with, particularly for outcomes related to trade facilitation, one possibility is that our excess interaction share merely reflects "familiarity" between inspector and broker, whereby the fact that certain brokers interact very frequently with an inspector reduces fixed inspection costs. Alternatively, inspectors may update their prior beliefs about brokers' likely compliance based on their past interactions with them and consequently be less likely to scrutinize brokers with whom they interact frequently for which they have a sizable pool of past interactions to base their inferences on. To assess the validity of these explanations for our results, we add to Equation (4) a measure of "familiarity": the total number of prior transactions of that same broker cleared by the same inspector over the preceding semester.[21] The results in Table 4 (Panel A) show that the familiarity measure itself has some predictive power: it is linked to slightly faster clearance times and higher tax revenue losses, but does not predict the incidence of fraud or value adjustment. More importantly for our purposes, controlling for familiarity only marginally reduces the impact of the excess interaction share which remains strongly statistically significant in all specifications. Put differently, the results do not appear to be driven by familiarity or learning, which, in any case, cannot explain why deviant declarations would be more risky to start with.

A second possible explanation for differential treatment is that it reflects congestion and fluctuations in inspectors' workload. Specifically, when inspectors get very busy they may be tempted to exert less scrutiny and speed up clearance merely to be able to manage increased traffic. If this increase in their workload is generated by absenteeism of other inspectors, we might see a simultaneous increase in the excess interaction share and a decrease in scrutiny and clearance times. To control for such congestion, we add to Equation (4) the number of declarations assigned to a given inspector over the course of the calendar month as a proxy for their workload. While Table 4 (Panel B) shows that this measure of workload is clearly positively correlated with clearance time, the impact of the excess interaction share on the other customs outcomes is hardly impacted by its inclusion. Similarly, Appendix Table 16 (Panel A) shows that the results are robust to controlling for inspector-month and broker-month fixed effects, which can also proxy for workload and congestion.

Third and related, one may worry that the patterns we document are an artefact of dubious declarations being more likely to be registered outside of regular business hours, i.e., late in the evening, at night, or during the weekend. This could help explain excess interaction since there are typically much fewer inspectors active and they may monitor incoming declarations less aggressively because they are fatigued and/or want to go home. However, Table 4 (Panel C) shows that the results are robust to excluding declarations registered outside of regular business hours, which account for less than 3% of all declarations.

Fourth, one may be concerned that the results are driven by (excess) interaction between inspectors and importers themselves rather than brokers, who are supposed to represent the interests of importers. We address this possibility in two ways. In Table 4 (Panel D) we control for importer fixed effects in Equation (4). This hardly impacts the qualitative

[21] Our excess interaction share measure is based on identifying deviations in the share of a given broker's declarations handled by a given inspector. By contrast, the familiarity measure is based on the absolute number of interactions between the broker and the inspector. Whereas inspectors will interact more with large brokers, and hence be more "familiar" with them, they will not necessarily interact excessively with large brokers, since our excess interaction share is a relative measure.

pattern of results.[22] In Table 4 (Panel E) we add to Equation (4) the excess interaction share between importers and inspectors. The measure is defined analogously to the excess interaction share between brokers and inspectors, for importers that lodged at least 50 declarations during the semester, which leads to a dramatic reduction in sample size. The excess interaction share between inspectors and importers neither significantly predicts fraud, nor value nor tax adjustment, and does not seem correlated with tax revenue losses. By contrast, the excess interaction share between inspectors and brokers remains robustly significant. These results justify our focus on brokers rather than importers. The fact that brokers seem to be the primary protagonists of the specific collusion scheme we document may be because they have more to gain from it; there are far fewer brokers than importers, and brokers interact more frequently with inspectors than importers do. Moreover, lobbying customs is the core business of brokers in many developing countries.

Fifth, given the limited number of inspectors working in Toamasina one may worry that our results are driven by a few individuals, rather than reflecting widespread collusion. Table 4 (Panel F) replicates our baseline results but excluding for each semester the top three inspectors with the greatest share of declarations with significant excess interaction.[23] Though this reduces the coefficient estimates on the excess interaction share, they remain strongly statistically significant. The results are thus not driven by a select few inspectors.

Sixth, another potential concern is that results might be driven by inspectors specializing in clearing different goods. This concern is mitigated by the fact that, formally, there is no specialization across different inspectors: they all clear the same set of goods. However, one may nonetheless wonder whether the IT department staff who are manipulating assignment are systematically assigning declarations containing certain products to unwitting inspectors that do not have the requisite expertise to adequately evaluate them; they may be seeking out inspectors that are the worst at detecting fraud for particular sets of products. To address this concern, Table 5 presents regressions where the unit of observation is an item (recall that a declaration can contain multiple items). The dependent variables are the log of the initially declared unit price, adjustments in that unit price, the finally registered unit price, the adjustment in weight (finally registered - initially declared) and an item-specific measure of the hypothetical tax revenue loss. The main regressor of interest is still the

[22] In Appendix Table 16 (Panel C) we add importer-broker fixed effects and show that this does not impact the qualitative pattern of results either.

[23] We average for each inspector each semester the indicator for significant excess interaction across all their declarations and identify the three inspectors with the highest averages. Note that when we include this indicator for significant excess interaction in Equation (4) (instead of the excess interaction share), in Appendix Table 16 (Panel B) we find results that are qualitatively similar to our main results in Table 3.

excess interaction share and the set of controls now includes HS 8-digit product-inspector fixed effects, broker fixed effects, source country fixed effects, month-year fixed effects, and a vector of both declaration characteristics (the risk score, a dummy for the red channel, a dummy for being a mixed shipment, a dummy for GasyNet's valuation advice) and item characteristics (tax rate and dummy for being a differentiated item). The HS 8-digit product-inspector fixed effects capture the comparative advantage of the inspector in detecting fraud in different types of products. The item-level initially declared unit price is significantly negatively correlated with excess interaction (column (1)). A 10 percentage point increase in the excess interaction share leads to a 12.5 percentage point reduction in the initial unit price. Adjustments in the unit price are also negatively correlated with the excess interaction share (column (2)). As a result, the final unit price is even more negatively correlated with excess interaction (column (3)), with a 10 percentage point increase in the excess interaction share being linked to a 12.9 percentage point reduction in the item-level final unit price. Inspectors are also significantly less likely to revise the weight of items contained in a declaration registered by a broker with whom they interact excessively frequently (column (4)). As a result, item-level tax revenue losses are strongly correlated with excess interaction. A 10 percentage point increase in the excess interaction share results in a 3.8 percentage point higher tax revenue loss.

Seventh, evidence of heterogeneity in the differential treatment of deviant declarations also helps to eliminate explanations other than collusion for our main results. We estimate Equation (4) allowing the excess interaction share to be interacted with, respectively, the tax rate and a proxy for the initial hypothetical tax revenue loss. Differential treatment by inspectors that interact excessively frequently with a given broker appears especially pronounced for declarations subject to higher taxes: these are especially less likely to be deemed fraudulent (column (2) in Appendix Table 17 (Panel A)) and exhibit significantly higher tax revenue losses (column (5) in Appendix Table 17 (Panel A)). Similarly, there is some (weak) evidence that declarations with higher tax revenue losses are less likely to be deemed fraudulent and valued upwards (columns (7) and (8) in Appendix Table 17 (Panel B)). This heterogeneity in differential treatment is consistent with collusion since inspectors have stronger incentives to privilege those declarations that are likely to command the highest bribes, but would not be obviously predicted by competing explanations.

Some final evidence consistent with collusion is provided by the analysis of inspector re-assignments made by the customs port manager. Such re-assignments are much more likely when declarations are initially assigned to an inspector with whom the broker is not interacting excessively frequently (see Appendix Table 18). This is inconsistent with re-assignments being random. Moreover re-assigned declarations typically yield higher fraud findings, value and tax adjustments. This is especially the case if they are taken away from inspectors with initial excess interaction, suggesting that these non-randomly assigned declarations were more risky to start with. By contrast, re-assigned declarations from inspectors without excess interaction towards inspectors with excess interaction do not yield increased fraud findings or tax adjustments, as is shown in Appendix Table 19.

6.3 Measuring Tax Revenue Losses

This section examines the impact of deviations from official rules in the assignment of declarations to inspectors on the various measures of hypothetical tax revenue losses described in Section 3.[24] In Section 8, we will use the estimates of this impact to quantify the costs of collusion by evaluating how much higher tax revenues would have been in the absence of excess interaction between inspectors and brokers.

To obtain estimates of β_E we estimate two variants of Equation (4). First, Table 6 (Panel A) examines differential treatment by inspectors using the same specification as in column (5) of Table 3. The resulting $\widehat{\beta_E}$ may constitute a downward-biased estimate of the overall impact of collusion on tax revenue losses because we are conditioning on variables that are potentially endogenous to collusion, such as inspector and broker fixed effects, and the risk score. Table 6 (Panel B) therefore presents the results from estimating a variant of Equation (4) that includes only controls that are exogenous to collusion: the tax rate, the dummy for mixed shipment, the share of differentiated products, source country fixed effects, HS 2-digit product fixed effects and month-year fixed effects.

To start with, Panel A, column (1) provides evidence that excess interaction is associated with underreporting of quantities, captured by the weight gap (relative to the port authority weight) for the declaration. A 10 percentage point increase in the excess interaction share is associated with underreporting of quantities by 1.6 percentage points. Measures of tax revenue losses that consider undervaluation but do not capture this margin of evasion yield downward-biased estimates of the costs of collusion. By implication, the impact of collusion on our baseline measure of tax revenue losses in column (2) is overly conservative. Indeed when we calculate measures of tax revenue losses that correct for underreporting of quantities as well as prices in column (3) we find a stronger impact of the excess interaction share on tax revenue losses.[25]

[24]Some of these measures are available only for selected sub-samples of declarations. The attendant sample selection bias is explored in Appendix Table 20.

[25]The difference in coefficients is in part driven by sample selection, as is shown in Appendix Table

A second reason why our baseline impact may be downward-biased is that the price correction is based on median import unit prices which may themselves be underreported. To circumvent this problem, columns (4) and (5) present estimates of the impact of the excess interaction share on hypothetical tax revenue losses based on prices reported by countries exporting to Madagascar, which are arguably less likely to be endogenous to underinvoicing in Madagascar (it is possible that exporters collude with importers such that even these measures of tax revenue losses would be conservative). Using exporter prices to benchmark hypothetical tax yield leads to a near doubling of the coefficient on the excess interaction share. Moreover, the explanatory power of our model increases significantly, as is evidenced by the higher R^2s. Column (6) presents estimates of hypothetical tax revenue losses calculated using transaction-specific valuation advice provided by the third-party GasyNet based on its own proprietary data. This yields estimates of β_E that are slightly lower than our baseline estimates. This could reflect sample selection bias since valuation advice is only issued for a very small sub-sample of declarations that are highly undervalued to start with.

Panel B replicates all these regressions but eliminating from the vector of controls the declaration characteristics that might be impacted by collusive agreements. While the excess interaction share is no longer significantly correlated with the weight gap (column (7)), it has consistently higher impacts on tax revenue losses than in Panel A. According to our preferred estimates of tax revenue loss (column (11)) which rely on a hypothetical tax yield based on exporter prices and corrected for potential underreporting of quantities, a 10 percentage point increase in the excess interaction share is associated with a 21 percentage point increase in tax revenue losses. The excess interaction share has a similar impact on hypothetical tax yield based on transaction-specific third-party reference values (column (12)) as in Panel A, but this impact may be driven by sample selection bias. We will use Panel B's estimates of β_E to quantify the costs of collusion in Section 8.

7 Did Delegated Re-Randomization of Inspector Assignment Curb Collusion?

After presenting a preliminary version of this paper to the Director General (DG) of customs, internal audits were launched and a number of inspectors were either sanctioned or strongly

20. Nonetheless, regressions whose dependent variable is a measure of tax revenue losses that corrects for underreporting of quantities and prices consistently yield higher coefficients on the excess interaction share.

encouraged to opt for voluntary retirement. The head of the IT department was suspended. The DG also decided to reform the assignment of declarations to inspectors,by delegating it to the third-party GasyNet, instructing them to re-randomize the initial assignment made by the customs IT system. This intervention provides us with a unique opportunity to assess whether the excess interactions we document are indeed the product of IT manipulation and hence to validate our methodology to detect collusion. It simultaneously offers a case study of the effectiveness of IT interventions to curb corruption and reduce fraud.

7.1 Prevalence of Excess Interaction During Re-Randomization Period

The re-randomization of inspector assignment started on November 18 2017 and led to the virtual disappearance of excess interaction, as is shown in Figure 5 which plots the evolution of the share of declarations characterized by significant excess interaction.[26] While the prevalence of excess interaction trended upward in the period preceding the re-randomization intervention, it suddenly and precipitously fell to nearly zero after the start of re-randomization indicated by the vertical bar in the graph. The re-randomization intervention thus effectively eliminated excess interaction between inspectors and brokers.

However, approximately four months after the start of the re-randomization intervention excess interaction resurfaced, plausibly driven by a new form of IT manipulation: the withholding of certain declarations from re-randomization. IT department staff complicit in collusion schemes figured out a way to temporarily shut down the automatic notification that GasyNet receives when a declaration is registered, thus preventing GasyNet from randomizing the inspector assignment of these declarations. Approximately 7% of declarations (1,275 declarations out of 17,736 declarations registered in the re-randomization period) were withheld from re-randomization. These declarations were readily identified by comparing the set of declarations re-randomized by GasyNet to the set of declarations that cleared customs daily. The set of declarations withheld from re-randomization likely includes declarations that were not deliberately "targeted" to bypass the randomization. Disabling the automatic notifications for some period implied that none of the declarations registered during that period were re-randomized, whether or not they were part of a collusive agreement.[27]

[26]Recall that the significant excess interaction indicator is equal to one when the excess interaction share is at least 0.05 and it is statistically significantly different from 0 at the 0.001 significance level

[27]The withholding of declarations subject to collusive agreements likely involves coordination between brokers and customs IT department staff: they are likely to agree on a particular time slot during which the re-randomization is temporarily shut down and the declaration is registered. However other brokers, who

The evolution of the withholding of declarations from re-randomization is depicted by the line with squares in Figure 5 and is remarkably similar to the evolution of significant excess interaction. In fact, 36% of the declarations that were withheld are characterized by significant excess interaction. Conversely, 63% of the declarations characterized by significant excess interaction in the re-randomization period were withheld from re-randomization. Interestingly, non-random assignment is persistent: for a given pairing of a broker with a particular inspector the share of withheld declarations is correlated with past deviations from random assignment, as shown in Appendix Table 22, suggesting the withholding of declarations from random assignment reflects a continuation of collusive agreements.

To ascertain that IT manipulation is driving the excess interaction we conduct a simple placebo test: we calculate the prevalence of excess interaction for the sub-sample of declarations that were re-randomized. Any excess interaction in this sub-sample should be purely accidental. Indeed, there is hardly any excess interaction in this sub-sample, as is shown by the line for "random excess interaction" in Figure 5. The only period with some some excess interaction is 5-7 months after the start of the re-randomization intervention, when a number of inspectors went on repeated strikes (resulting in a higher average workload, and possibly higher excess interaction shares, for the remaining inspectors). Put differently, without the bypassing of the re-randomization there would not have been a resurgence of excess interaction between inspectors and brokers.

7.2 Excess Interaction and Evasion Risk During Re-Randomization Period

Random excess interaction (i.e., excess interaction in the sample of declarations whose assignment was re-randomized) is not correlated with declaration characteristics commonly associated with tax evasion, as is shown in Table 7 (Panel A) which replicates some of the key specifications presented in Table 2 for the sample of re-randomized declarations. There is no correlation between random excess interaction and any of the key tax evasion risk characteristics; all the R^2s are 0 and none of the coefficients are significant.

By contrast, Panel B shows that declarations withheld from re-randomization are not only characterized by significantly higher excess interaction shares but are also significantly more at risk of tax evasion on average than declarations that were re-randomized. They are subject to tax rates that are 8.8 percentage points higher, have risk scores that are 1.2

are not part of collusive agreements may also register declarations during these time slots, which implies that not all declarations that are withheld from randomization are part of collusive agreements.

points higher, are significantly heavier, and exhibit 19.7 percentage points lower initial unit prices relative to median import unit prices. As a result, these declarations exhibit 19.9 percentage points higher tax revenue losses than similar declarations whose assignment to inspectors was re-randomized.

Even in the re-randomization period, excess interaction is significantly correlated with declaration characteristics associated with tax evasion risk, as is shown in Panel C which replicates Table 2 using the entire sample of declarations (re-randomized and withheld from re-randomization) in this period. However, these correlations are entirely driven by declarations withheld from re-randomization as is shown in Panel D in which we interact the excess interaction share with the dummy for being withheld from re-randomization. While being withheld from re-randomization continues to significantly predict tax evasion risk, the excess interaction share only has predictive power when interacted with being withheld from randomization (consistent with the results in Panel B). The declarations withheld from re-randomization and cleared by inspectors with a higher excess interaction share have significantly lower initial unit prices and significantly higher initial tax revenue losses (columns (20) and (21)). This suggests the declarations withheld from re-randomization that were targeted by collusion schemes were assigned to certain "preferred" inspectors.

7.3 Differential Treatment During Re-Randomization Period

To ascertain the extent to which the IT manipulation during the re-randomization period reflects a continuation of collusion, Table 8 examines whether inspectors treat the manipulated declarations differently. The table replicates the specifications in Table 3 but using different proxies for excess interaction. Panel A shows that for the sub-sample of re-randomized declarations, random excess interaction does not predict how long inspectors take to clear goods, nor whether they will report the declaration as being fraudulent, or change the value or the tax yield. Random excess interaction is negatively correlated with tax revenue losses, suggesting that it is linked to lower, not higher, tax losses, in this sample of re-randomized declarations.

Panel B shows that declarations that were withheld from re-randomization are cleared significantly faster than declarations that were not. The estimates also point to a reduced likelihood of being reported fraudulent and lower value and tax adjustments but these effects are not statistically significant. Declarations withheld from re-randomization exhibit significant and substantial tax revenue losses of 17.5 percentage points on average, relative to other declarations, ceteris paribus.

Panel C suggests that declarations characterized by excess interaction are again associated with significantly accelerated clearance, significantly reduced fraud, lower value and tax adjustments and significantly higher tax revenue losses. However this preferential treatment is driven by the declarations withheld from re-randomization since we did not observe these correlations in the sample of re-randomized declarations analyzed in panel A.

In Panel D we consider the entire sample of declarations and include the excess interaction share, a dummy for being withheld from re-randomization, and the interaction between these two measures. The coefficients on the interaction term are consistently highly statistically significant. Preferential treatment is most pronounced for declarations that are withheld from re-randomization and handled by inspectors who interact excessively frequently with a given broker. Such declarations are especially rapidly cleared, especially less likely to be deemed fraudulent, are subject to significantly lower value and tax adjustments, and, as a result, exhibit higher tax revenue losses.

The preferential treatment by inspectors of declarations characterized by excess interaction was thus driven by manipulation of the IT system. Our placebo tests show clearly that when declarations are truly randomly assigned, there is hardly any excess interaction. Whatever accidental excess interaction nonetheless arises is not correlated with customs outcomes. By contrast, declarations withheld from re-randomization are associated with excess interaction and an increased risk of tax evasion. They receive privileged treatment from inspectors, especially when such inspectors are handling a significantly larger share of a given broker's declarations than would be expected had the assignment of declarations followed official rules. All in all, these results corroborate our methodology to detect collusion and also attest to the difficulties associated with dislodging systemic collusion.

8 How Costly Is Collusion?

How much tax revenue is lost because of collusion? To answer this question we calculate how much more tax revenue would have been collected if there was no excess interaction between inspectors and brokers and, during the re-randomization period, if there was no withholding of declarations from the re-randomization. We calculate separate counterfactual estimates for the period before the re-randomization intervention and for the re-randomization period for two reasons. First, the re-randomization intervention may have had a deterrence effect. Second, the novel IT manipulation uncovered during the re-randomization period arguably facilitates identification of the specific declarations that were the object of collusive schemes, i.e., those that were both withheld from re-randomization and handled by inspectors that were interacting excessively frequently with the broker that registered the declarations.

To calculate the counterfactual tax revenue gains associated with the elimination of collusion during the period before the re-randomization intervention we use the estimates of β_E presented in Panel B of Table 6 and calculate for each declaration the counterfactual tax revenue that would have been collected in the absence of collusion between inspectors and brokers as $T^{NC} = T * exp(\widehat{\beta_E * ES})$, where T is the actual tax yield.[28]

It is important to bear in mind that these estimates reflect the *differential* impact of "collusive evasion" on tax revenues. They do not reflect the gains associated with eliminating tax evasion altogether, but only the gains from eliminating "collusive evasion". We are effectively asking how much more tax revenue would have been collected if declarations subject to "collusive evasion" had been treated like declarations that were not. Our methodology does not address the (rather plausible) possibility that the latter declarations may nonetheless be characterized by "non-collusive evasion" resulting from deals made between randomly assigned inspectors, brokers and/or importers.

The results of this exercise are presented in Table 9 (Panel A). We show estimates both for declarations characterized by significant excess interaction (in the first two columns) as well as for all declarations (in the last two columns). Interestingly, declarations with significant excess interaction yield more tax revenue, 11,538 USD on average, despite being undervalued, than the average declaration with 10,432 USD. This finding reflects the fact that declarations with significant excess interaction are subject to higher tax rates (as was shown in Section 5). In the absence of collusion the average declaration with significant excess interaction would have yielded an additional 993 USD in tax revenue if we valued imports at the median import unit price and an additional 1,552 USD when also correcting for underreporting of quantities. According to our preferred counterfactual estimates, which evaluate hypothetical tax yield using prices reported by exporters and also correct for potential underreporting of quantities, tax yield per declaration would have been 3,135 USD higher. Put differently, the tax yield on declarations likely to be the object of collusive agreements would have been 27 percentage points higher. This number is a lower bound on total tax revenue losses per declaration associated with collusion since the set of declarations characterized by significant excess interaction likely also includes some that were randomly assigned and not the object of collusion schemes.

Given the prevalence of collusion, it had a sizeable impact on overall tax revenues collected in Toamasina. According to our preferred estimates, which use reference prices reported

[28]The details of this calculation are provided in the Appendix.

by exporters and correct for underreporting of quantities, average and thus aggregate tax yield would have been 4.1 percentage points higher each year in the period before the re-randomization intervention.

We quantify tax revenue losses in the re-randomization period by estimating regressions in which we regress measures of potential tax revenue loss on the excess interaction share, a dummy for declarations being withheld and their interaction.[29] The resulting estimates are presented in Appendix Table 21. Counterfactual tax revenue in the absence of collusion is now calculated as $T^{NC} = T * exp(\widehat{\beta_E} * ES + \widehat{\beta_P} * WFR + \widehat{\beta_{EP}} * ES * WFR)$, where WFR is a dummy for declarations whithheld from re-randomization.

The results of these counterfactual calculations are presented in Table 9 (Panel B) both for declarations that were withheld from re-randomization and cleared by an inspector who handled excessively frequently with the broker who registered them (first two columns) as well as for all declarations (last two columns). The latter set of declarations is more likely to have been the object of collusive agreements. Interestingly, collusive declarations now yield less tax revenue (9,435 USD) than the average declaration (10,749 USD). According to our preferred counterfactual estimates which calculate hypothetical tax yield using prices reported by exporters, declarations that were likely the object of collusion would have yielded an *additional* 11,442 USD in tax revenue, which represents a staggering 121.3% increase over actual tax yield. More conservative estimates that calculate revenue losses using median import unit prices still predict a 41.1% gain in tax yield. Both estimates are conservative since we are not able to correct for potential underreporting of quantities. According to our preferred estimates, aggregate tax yield in the re-randomization period would have been 5.1% higher had the re-randomization not been undermined by a new form of IT manipulation. While these back-of-the-envelope estimates are crude and must be interpreted with caution given the difficulties inherent in measuring hypothetical tax yield, they underscore that collusion substantially compromised fiscal performance in Madagascar.

9 Conclusion

Corrupt governance and limited state capacity to raise tax revenue constrain development, yet surprisingly little is known about the extent to which tax evasion is facilitated by (which) bureaucrats. Evidence on effectiveness of reforms to remedy institutionalized corruption is also limited. These questions are especially pertinent for customs agencies in low-income

[29]Note that for this period we are unable to construct estimates of underreporting of quantities as data on the weight of containers measured upon arrival is not available.

countries, which tend to be more reliant on revenues collected at the border than developed countries despite suffering higher levels of evasion.

This paper begins to fill these gaps in the literature by presenting a new methodology to detect collusion between customs inspectors and brokers. Our approach is based on identifying deviations from random assignment of import declarations to inspectors that is prescribed by official rules. Such deviations result in excessively frequent pairing of brokers with the inspector(s) they are colluding with.

Applying this methodology to Madagascar's main port of Toamasina unveiled that 9.7% of declarations were handled by inspectors that were not randomly assigned, plausibly because of manipulation of the IT system that assigns them. Non-randomly assigned declarations were shown to be subject to higher tax rates, have higher risk scores and lower unit prices than those reported for declarations containing the same goods. Non-random assignment is thus associated with higher tax revenue losses. Customs inspectors are shown to provide preferential treatment to these deviant declarations by clearing them faster, being less likely to require value, weight, and tax adjustments, and failing to identify fraud. Such collusion is costly; tax yield for non-randomly assigned declarations would have been 27 percentage points higher in the absence of excess interaction between inspectors and brokers. Overall tax revenues collected in Toamasina would have been 4 percentage points higher in the absence of the collusion schemes unveiled in this paper.

An intervention to curb collusion by having a third party re-randomize inspector assignment validates our methodology because it led to the temporary disappearance of excess interaction between inspectors and brokers. It also triggered a novel form of IT manipulation. While manipulation of inspector assignment was eventually weeded out with the help of better IT infrastructure, our results serve as a reminder that technology is not a panacea in the fight against corruption. Rather, our results illustrate how IT solutions can be captured by bureaucrats and economic operators and serve as a conduit to corruption. The persistence of collusion also points to the importance of altering incentives in order to dislodge systemic corruption.

References

Aizer, A., and J. Doyle (2015): "Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges," *The Quarterly Journal of Economics*, 130, 759–803.

Atkin, D., and A. Khandelwal (2019): "How Distortions Alter the Impacts of International Trade in Developing Countries," Discussion paper, National Bureau of Economic Research.

Bardhan, P. (1997): "Corruption and development: a review of issues," *Journal of Economic Literature*, 35(3), 1320–1346.

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

Besley, T., and T. Persson (2009): "The origins of state capacity: Property rights, taxation, and politics," *American Economic Review*, 99(4), 1218–44.

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

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2020): "Incarceration, recidivism, and employment," *Journal of Political Economy*, 128(4), 1269–1324.

Business, D. (2020): "Comparing Business Regulation in 190 Economies," *URL: http://openknowledge. worldbank. org.(date of access: 14.01. 2020).*

Casaburi, L., M. Kremer, and R. Ramrattan (2019): "Crony Capitalism, Collective Action, and ICT: Evidence from Kenyan Contract Farming," Discussion paper, Working paper.

Chalendard, C., G. Raballand, and A. Rakotoarisoa (2019): "The use of detailed statistical data in customs reforms: The case of Madagascar," *Development Policy Review*, 37(4), 546–563.

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).

Cohen, A., and C. S. Yang (2019): "Judicial politics and sentencing decisions," *American Economic Journal: Economic Policy*, 11(1), 160–191.

Di Tella, R., and E. Schargrodsky (2013): "Criminal recidivism after prison and electronic monitoring," *Journal of Political Economy*, 121, 28–73.

Dincecco, M., and N. Ravanilla (2017): "The Importance of Bureaucrats in a Weak State: Evidence from the Philippines," *Available at SSRN 2773884*.

Dutt, P., and D. Traca (2010): "Corruption and bilateral trade flows: extortion or evasion?," *The Review of Economics and Statistics*, 92(4), 843–860.

Ferraz, C., and F. Finan (2008): "Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes," *The Quarterly Journal of Economics*, 123(2), 703–745.

Finan, F., B. A. Olken, and R. Pande (2017): "The personnel economics of the developing state," in *Handbook of Economic Field Experiments*, vol. 2, pp. 467–514. Elsevier.

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.

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.

Jacob, B. A., and S. D. Levitt (2003): “Rotten apples: An investigation of the prevalence and predictors of teacher cheating,” *The Quarterly Journal of Economics*, 118(3), 843–877.

Khan, A. Q., A. I. Khwaja, and B. A. Olken (2015): “Tax farming redux: Experimental evidence on performance pay for tax collectors,” *The Quarterly Journal of Economics*, 131(1), 219–271.

Khan, A. Q., A. I. Khwaja, and B. A. Olken (2019): “Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings,” *American Economic Review*, 109(1), 237–70.

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.

Laajaj, R., M. Eslava, and T. Kinda (2019): “The costs of bureaucracy and corruption at customs: Evidence from the computerization of imports in Colombia,” *Documento CEDE*, (2019-08).

Niehaus, P., and S. Sukhtankar (2013): “Corruption dynamics: The golden goose effect,” *American Economic Journal: Economic Policy*, 5(4), 230–269.

Olken, B. A., and R. Pande (2012): “Corruption in developing countries,” *Annual Review of Economics*, 4(1), 479–509.

Pepinsky, T. B., J. H. Pierskalla, and A. Sacks (2017): “Bureaucracy and service delivery,” *Annual Review of Political Science*, 20, 249–268.

Pomeranz, D., and J. Vila-Belda (2019): “Taking state-capacity research to the field: Insights from collaborations with tax authorities,” *Annual Review of Economics*, 11, 755–781.

Price, J., and J. Wolfers (2010): “Racial discrimination among NBA referees,” *The Quarterly Journal of Economics*, 125, 1859–1887.

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

Rijkers, B., L. Baghdadi, and G. Raballand (2017): “Political connections and tariff evasion evidence from Tunisia,” *The World Bank Economic Review*, 31(2), 459–482.

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.

Shleifer, A., and R. W. Vishny (1993): “Corruption,” *The Quarterly Journal of Economics*, 108(3), 599–617.

——— (2002): *The grabbing hand: Government pathologies and their cures.* Harvard University Press.

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

Wier, L. (2020): "Tax-motivated transfer mispricing in South Africa: Direct evidence using transaction data," *Journal of Public Economics*, 184, 104153.

Xu, G. (2018): "The costs of patronage: Evidence from the british empire," *American Economic Review*, 108(11), 3170–98.

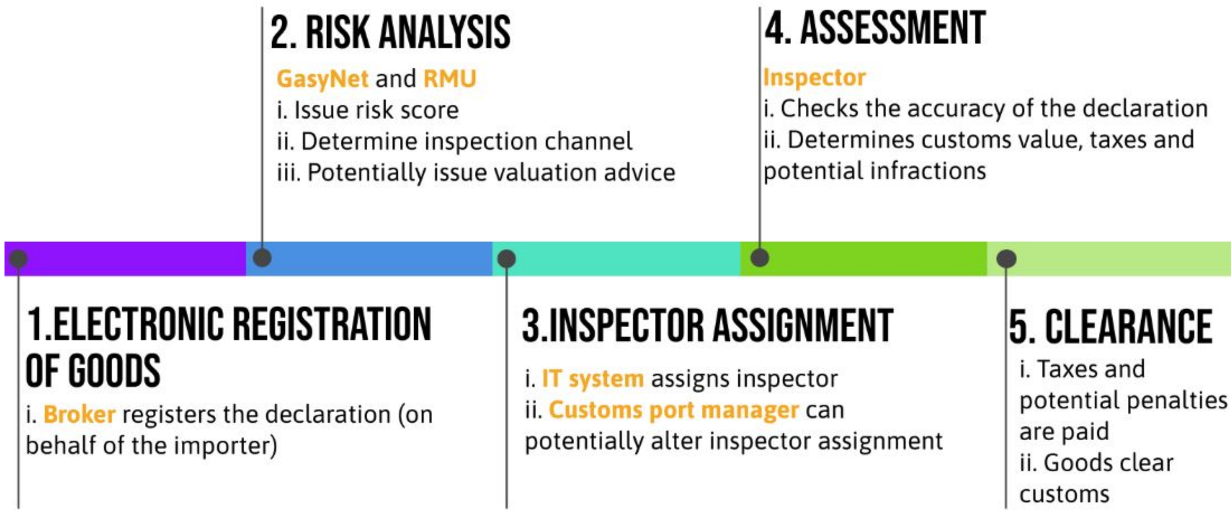
Xu, G., M. Bertrand, and R. Burgess (2018): "Social Proximity and Bureaucrat Performance: Evidence from India," Discussion paper, National Bureau of Economic Research.

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

Zitzewitz, E. (2012): "Forensic economics," *Journal of Economic Literature*, 50(3), 731–69.

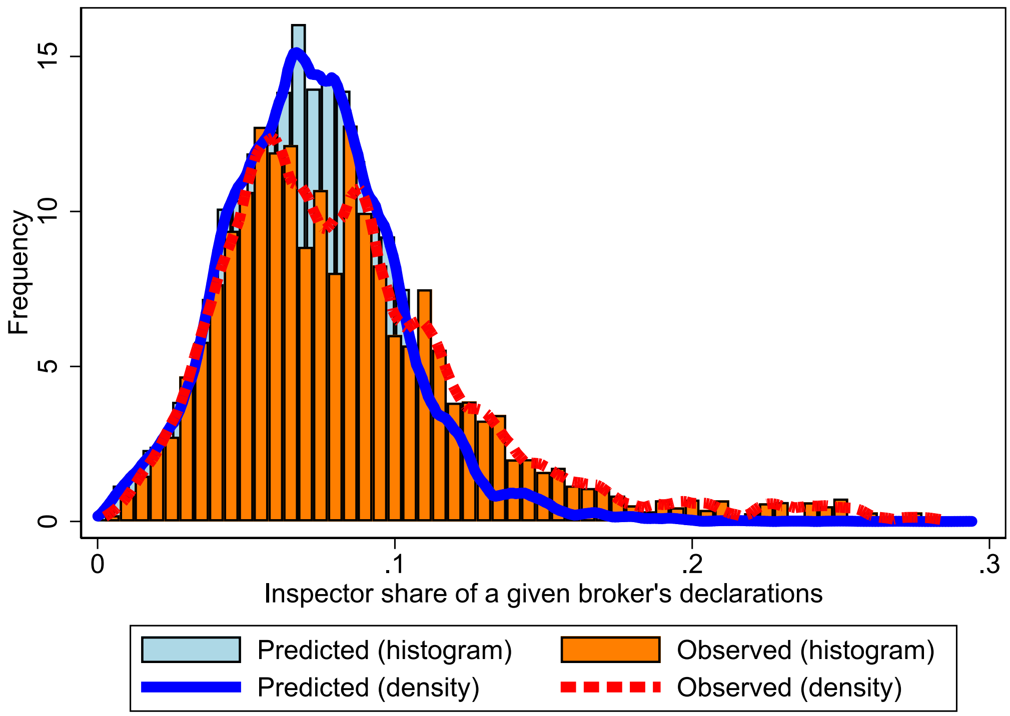
10 Tables and Figures

Figure 1: Stylistic Representation of the Clearance Process



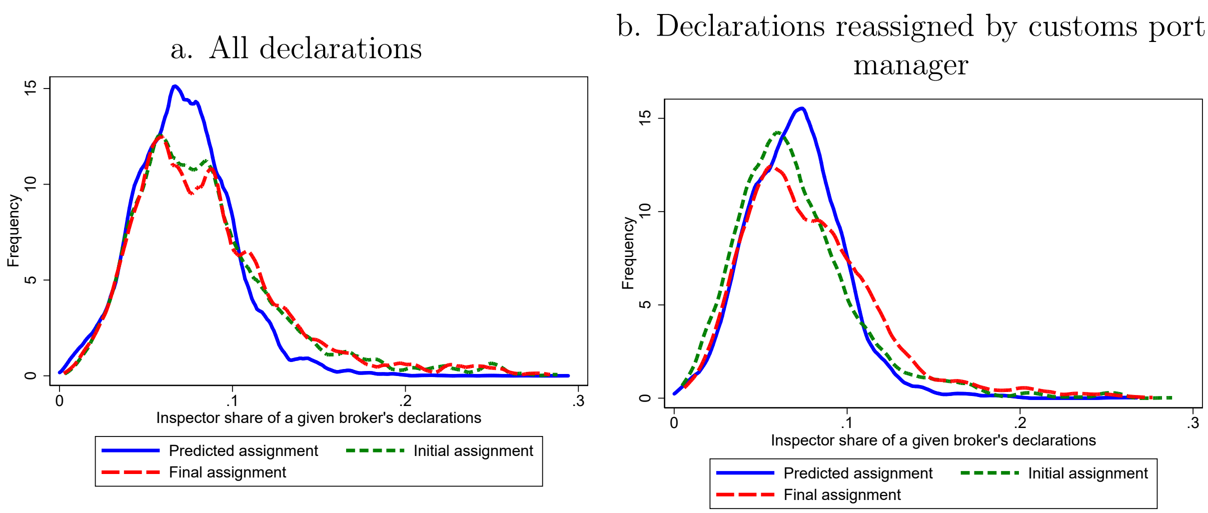
Notes: The figure depicts a stylized representation of the customs clearance process. RMU is the risk management unit of customs. GasyNet is a third-party that assists customs with risk analysis and logistics.

Figure 2: Deviations from Official Rules in Assignment of Declarations to Inspectors



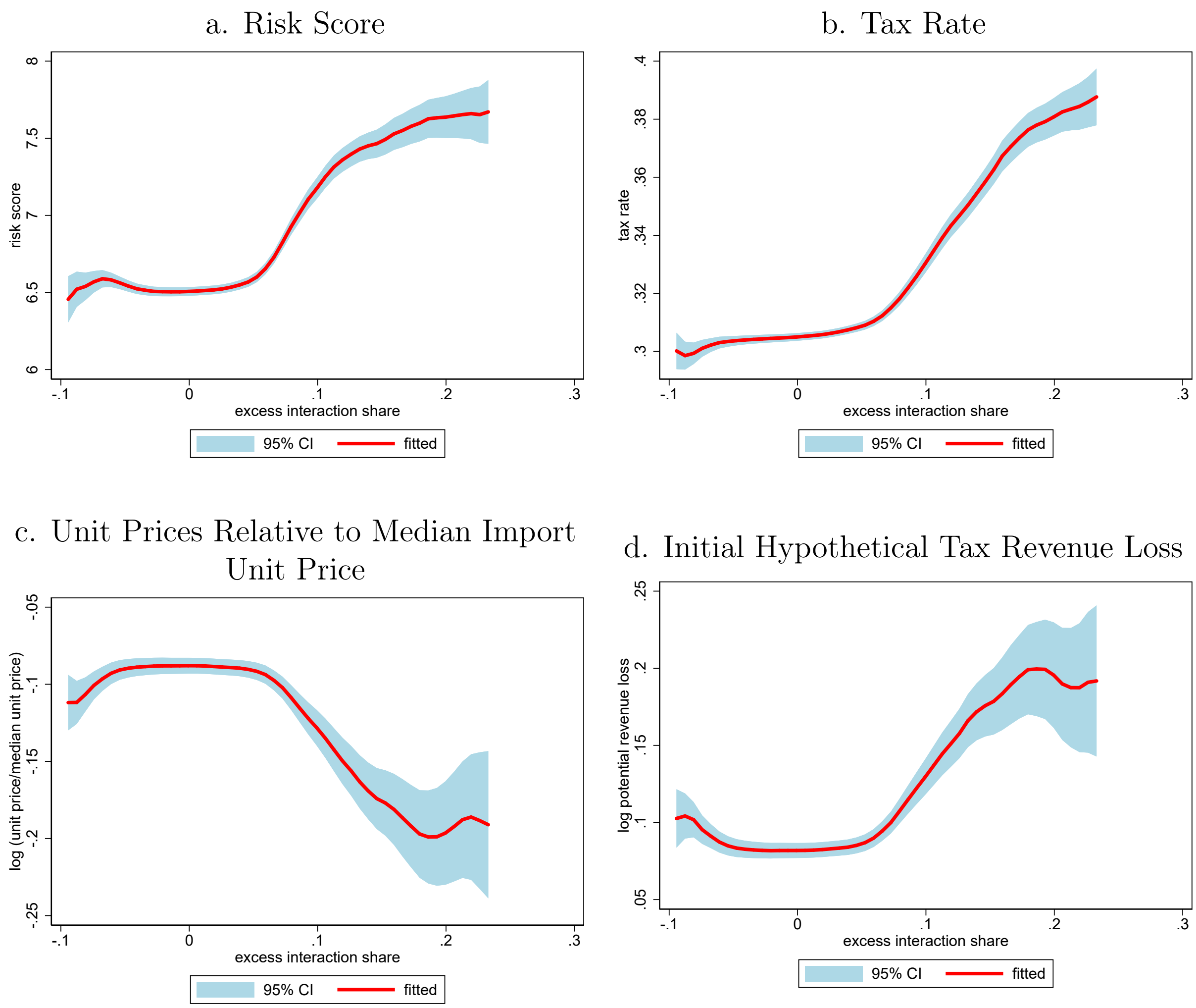
Notes: The figure shows the distribution of the share of declarations of a given broker handled by a given inspector in the period January 1, 2015 to November 17, 2017 (i.e., before the re-randomization intervention). The light blue bars show the histogram of predicted inspection shares if the assignment of declarations followed the official rules (notably random assignment conditional on inspector productivity as explained in Section 2), and the blue solid line the overlaid kernel density plot of such inspection shares. Predicted inspection shares are generated using a multinomial distribution in which inspector productivity is set equal to the share of all import declarations registered in a given semester handled by her. The orange bars indicate the distribution of observed inspection shares, with the red long-dashed line showing the overlaid kernel density plot.

Figure 3: Initial versus Final Inspector Assignment



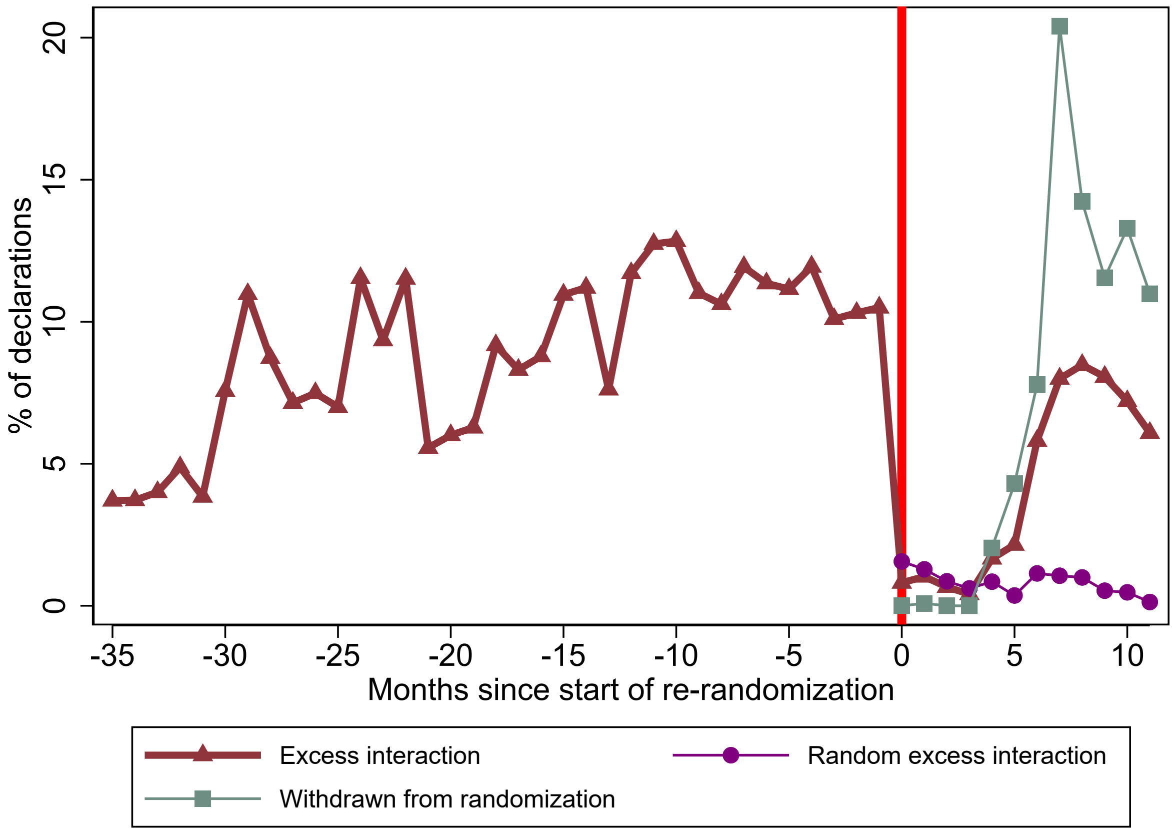
Notes: The figure shows the kernel density distributions of the share of declarations of a given broker handled by a given inspector in the period January 1, 2015 to November 17, 2017 (i.e., before the re-randomization intervention). The blue solid density plot shows the distribution of predicted inspection shares if the assignment of declarations followed the official rules (notably random assignment conditional on inspector productivity as explained in Section 2). Predicted inspection shares are generated using a multinomial distribution in which inspector productivity is set equal to the share of all import declarations registered in a given semester handled by her. The green short-dashed line shows the distribution of the observed initial assignment of a declaration to a given inspector by the IT system (before the customs port manager potentially intervenes). The red long-dashed line shows the distribution of the observed final assignment of a declaration to an inspector after potential re-assignments made by the customs port manager. In the left graph the sample includes all declarations (both those that were re-assigned by the customs port manager and those that were not) while in the right graph the sample includes only declarations that were re-assigned by the customs port manager.

Figure 4: Excess Interaction Share and Evasion Risk



Notes: The graphs show weighted local polynomial plots (using the Epanechnikov kernel function) of excess interaction shares on a selected number of declaration characteristics capturing evasion risk for the period January 1, 2015 to November 17, 2017 (i.e., before the re-randomization intervention). *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). Initial hypothetical tax revenue loss refers to the tax revenue loss estimated using the import value initially submitted by the broker in the IT system. CI stands for confidence interval.

Figure 5: Evolution of Non-Random Assignment



Notes: The line with triangles "Excess interaction" depicts the share of all import declarations that are characterized by significant excess interaction (as defined in Section 4). The red vertical bar denotes the start of the re-randomization intervention in which the assignment of declarations to inspectors was delegated to the third party GasyNet. Soon after this start, the customs IT department managed to withhold several declarations from the re-randomization process. The prevalence of these declarations is shown by the line with squares "Withdrawn from randomization". The line with circles "Random excess interaction" refers to the share of re-randomized declarations that are characterized by significant excess interaction (as defined in Section 4). The sample covers the period January 1, 2015 to November 17, 2018.

Table 1: Descriptive Statistics

Before delegated re-randomization of inspector assignment			
A. Prevalence of significant excess interaction (i.e., non-random assignment)			
Entire period	N Non-randomly assigned	N Total	%
Declarations - after initial assignment	4,183	45,059	9.3%
Declarations - after final assignment	4,353	45,059	9.7%
Average per semester	At least one non-randomly assigned declaration	Total	%
Inspectors	10	17	58.8%
Brokers	15	45	33.3%
Inspector-broker pairs	23	707	3.3%
B. Contributions of declarations with significant excess interaction			
% of final import value			7.9%
% of taxes paid			10.7%

Notes: Declarations are characterized by significant excess interaction if (i) they are handled by an inspector whose *excess interaction share* (the difference between the share of given broker's declarations handled by the inspector in question and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules) is at least 0.05 and (ii) the null hypothesis of conditional random assignment is rejected at the 0.001 significance level (see Section 4 for more information). Initial assignment refers to the assignment originally made by the customs IT system. Final assignment takes into account subsequent potential re-assignment(s) made and therefore corresponds to the last assignment that selected the inspector that cleared the declaration. The final import value corresponds to the import value initially declared by the broker augmented by potential adjustments made by the inspector. The sample covers the period January 1, 2015 to November 17, 2017.

Table 2: Excess Interaction and Tax Evasion Risk

Before delegated re-randomization of inspector assignment					
Dependent variable	**Risk score**	**Tax rate**	**Red channel dummy**	**Mixed shipment dummy**	**Differentiated share**
	(1)	(2)	(3)	(4)	(5)
Excess interaction share	0.306***	5.178***	0.088	0.775**	0.503***
	(0.054)	(0.744)	(0.175)	(0.284)	(0.167)
Observations	45,059	44,522	45,059	45,059	45,059
R-squared	0.010	0.006	0.000	0.005	0.002
	Valuation advice dummy	**Log initial value**	**Log initial weight**	**Log initial unit price (relative to median)**	**Initial hyp. tax revenue loss (%)**
	(6)	(7)	(8)	(9)	(10)
Excess interaction share	0.937***	0.257	1.458***	-0.591***	0.633***
	(0.301)	(0.216)	(0.310)	(0.190)	(0.187)
Observations	45,059	45,059	45,059	45,034	45,034
R-squared	0.019	0.000	0.001	0.002	0.003

Notes: Robust standard errors are presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 3: Differential Treatment by Inspectors

Before delegated re-randomization of inspector assignment					
Dependent variable	**Time**	**Fraud**	**Δlog value**	**Δlog tax**	**Hyp. tax revenue loss (%)**
	(1)	(2)	(3)	(4)	(5)
Excess interaction share	-2.008***	-0.275***	-0.078***	-0.086***	0.389***
	(0.289)	(0.056)	(0.017)	(0.025)	(0.103)
Declaration characteristics	Yes	Yes	Yes	Yes	Yes
Inspector fixed effects	Yes	Yes	Yes	Yes	Yes
Broker fixed effects	Yes	Yes	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	41,121	44,522	44,435	40,471	44,497
R-squared	0.318	0.214	0.152	0.132	0.211

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). Declaration characteristics include the tax rate, the risk score, a dummy for the red channel, the share of value accounted for by differentiated products, a dummy indicating whether the declaration was mixed, and a dummy indicating the declaration was subject to valuation advice. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 4: Addressing Alternative Explanations

Dependent variable	Time	Fraud	Δ log value	Δ log tax	Hyp. tax revenue loss (%)
A. Controlling for familiarity					
	(1)	(2)	(3)	(4)	(5)
Excess interaction share	-1.797***	-0.278***	-0.080***	-0.082***	0.323***
	(0.358)	(0.058)	(0.017)	(0.023)	(0.110)
Log number of interactions past 6 months	-0.042*	-0.000	0.000	-0.001	0.012**
	(0.024)	(0.004)	(0.001)	(0.001)	(0.005)
Observations	40,990	44,359	44,274	40,324	44,335
R-squared	0.321	0.214	0.152	0.133	0.211
B. Controlling for congestion					
	(6)	(7)	(8)	(9)	(10)
Excess interaction share	-2.002***	-0.275***	-0.078***	-0.085***	0.389***
	(0.290)	(0.056)	(0.017)	(0.025)	(0.103)
Log inspector workload	0.098***	-0.004	0.001	0.000	-0.001
	(0.034)	(0.004)	(0.001)	(0.001)	(0.005)
Observations	41,121	44,522	44,435	40,471	44,497
R-squared	0.318	0.214	0.152	0.132	0.211
C. Excluding declarations registered outside regular business hours					
	(11)	(12)	(13)	(14)	(15)
Excess interaction share	-2.033***	-0.270***	-0.079***	-0.088***	0.385***
	(0.290)	(0.053)	(0.016)	(0.023)	(0.110)
Observations	40,285	43,497	43,410	39,534	43,473
R-squared	0.316	0.220	0.156	0.136	0.210
D. Adding importer fixed effects					
	(16)	(17)	(18)	(19)	(20)
Excess interaction share	-2.051***	-0.170***	-0.069***	-0.081***	0.242***
	(0.232)	(0.054)	(0.016)	(0.020)	(0.065)
Observations	40,311	43,691	43,601	39,678	43,669
R-squared	0.393	0.327	0.292	0.297	0.429
E. Adding importer fixed effects & excess interaction share with importers					
	(21)	(22)	(23)	(24)	(25)
Excess interaction share	-2.123***	-0.228***	-0.040	-0.067*	0.168
	(0.560)	(0.080)	(0.031)	(0.038)	(0.123)
Excess interaction share with importer	-0.125	0.016	-0.006	0.000	0.078
	(0.297)	(0.080)	(0.015)	(0.024)	(0.077)
Observations	9,391	10,133	10,115	9,036	10,130
R-squared	0.365	0.300	0.231	0.219	0.229
F. Dropping top 3 inspectors with largest share of declarations with significant excess interaction each semester					
	(26)	(27)	(28)	(29)	(30)
Excess interaction share	-1.928***	-0.232***	-0.064***	-0.069***	0.395**
	(0.375)	(0.061)	(0.015)	(0.023)	(0.152)
Observations	33,712	36,387	36,313	32,982	36,366
R-squared	0.319	0.219	0.155	0.137	0.205

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). All specifications include the tax rate, the risk score, a dummy for the red channel, the share of value accounted for by differentiated products, a dummy indicating whether the declaration was mixed, and a dummy indicating the declaration was subject to valuation advice, as well as inspector fixed effects, broker fixed effects, source country fixed effects, HS 2-digit product fixed effects, and month-year fixed effects. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 5: Item-Level Differential Treatment Regressions

	Before delegated re-randomization of inspector assignment				
Dependent variable	**Log initial unit price**	**Δ log unit price**	**Log final unit price**	**Δ log weight**	**Hyp. tax revenue loss (%)**
	(1)	(2)	(3)	(4)	(5)
Excess interaction share	-1.248***	-0.038*	-1.285***	-0.007*	0.383**
	(0.315)	(0.020)	(0.305)	(0.004)	(0.148)
Inspector-HS 8-digit product FE	Yes	Yes	Yes	Yes	Yes
Declaration characteristics	Yes	Yes	Yes	Yes	Yes
Item characteristics	Yes	Yes	Yes	Yes	Yes
Broker FE	Yes	Yes	Yes	Yes	Yes
Source country FE	Yes	Yes	Yes	Yes	Yes
Month-year FE	Yes	Yes	Yes	Yes	Yes
Observations	186,823	186,823	186,823	186,823	186,519
R-squared	0.569	0.150	0.566	0.136	0.326

Notes: FE stands for fixed effects. Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). All specifications include as declaration characteristics the risk score, a dummy for the red channel, a dummy indicating whether the declaration was mixed, and a dummy indicating the declaration was subject to valuation advice and as item characteristics the tax rate and a dummy for whether the product is differentiated. Initial unit price refers to unit value initially declared by the broker. Final unit price refers to the initial value augmented by potential revisions made by customs. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 6: Excess Interaction and Tax Revenue Losses

Before delegated re-randomization of inspector assignment						
Dependent variable	**Weight gap**	**Hypothetical tax revenue loss (%)**				
Reference weight	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Reference price	*No*	*Importers*	*Importers*	*Exporters*	*Exporters*	*Third-party*
A. Differential treatment by inspectors						
	(1)	(2)	(3)	(4)	(5)	(6)
Excess interaction share	0.155**	0.389***	0.575***	0.760***	1.126***	0.294***
	(0.058)	(0.103)	(0.095)	(0.148)	(0.233)	(0.058)
Declaration characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Inspector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Broker fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,750	44,497	23,937	31,103	16,457	4,254
R-squared	0.100	0.211	0.250	0.571	0.454	0.431
B. Overall tax revenue losses associated with collusion						
	(7)	(8)	(9)	(10)	(11)	(12)
Excess interaction share	0.088	0.732***	1.112***	1.659***	2.085***	0.851***
	(0.055)	(0.201)	(0.175)	(0.234)	(0.196)	(0.095)
Exog. decl. characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,916	45,033	23,965	31,402	16,475	4,258
R-squared	0.095	0.181	0.222	0.532	0.420	0.342

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). Declaration characteristics include the tax rate, the risk score, a dummy for the red channel, the share of value accounted for by differentiated products, a dummy indicating whether the declaration was mixed, and a dummy indicating the declaration was subject to valuation advice. Exogenous declarations characteristics include the tax rate, the share of value accounted for by differentiated products, and a dummy indicating whether the declaration was mixed. "Importers", "Exporters" and "Third-party" refer, respectively, to median import unit prices, unit prices reported by countries exporting to Madagascar, and transaction-specific valuation advice provided by the third-party GasyNet based on its own proprietary data. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 7: Excess Interaction and Tax Evasion Risk During Re-Randomization Period

	After delegation of the re-randomization of inspector assignment					
Dependent variable	**Excess interaction share**	**Tax rate**	**Risk score**	**Log initial weight**	**Log initial unit price**	**Initial hyp. tax revenue loss (%)**
A. Random excess interaction share (re-randomized declarations only)						
		(1)	(2)	(3)	(4)	(5)
Random excess interaction share		0.037	-0.693	0.702	0.196	-0.196
		(0.195)	(2.796)	(1.453)	(0.185)	(0.173)
Observations		16,461	15,925	16,461	16,454	16,454
R-squared		0.000	0.000	0.000	0.000	0.000
B. Withheld from randomization (WFR)						
	(6)	(7)	(8)	(9)	(10)	(11)
WFR	0.064**	0.088***	1.173***	0.160**	-0.197***	0.199***
	(0.024)	(0.005)	(0.084)	(0.053)	(0.045)	(0.045)
Observations	17,736	17,738	17,169	17,738	17,728	17,728
R-squared	0.153	0.026	0.011	0.001	0.011	0.012
C. Excess interaction						
		(12)	(13)	(14)	(15)	(16)
Excess interaction share		0.307***	4.212***	0.392	-0.840***	0.836***
		(0.068)	(1.325)	(0.841)	(0.243)	(0.242)
Observations		17,736	17,167	17,736	17,726	17,726
R-squared		0.008	0.004	0.000	0.005	0.006
D. Combined measures						
		(17)	(18)	(19)	(20)	(21)
WFR		0.079***	1.015***	0.177*	-0.110***	0.114***
		(0.005)	(0.131)	(0.088)	(0.026)	(0.025)
Excess interaction share		0.093	1.196	0.147	0.032	-0.032
		(0.111)	(2.013)	(1.512)	(0.110)	(0.102)
WFR*Excess interaction share		0.051	1.169	-0.386	-1.286***	1.262***
		(0.113)	(2.359)	(1.790)	(0.138)	(0.142)
Observations		17,736	17,167	17,736	17,726	17,726
R-squared		0.027	0.011	0.001	0.015	0.016

Notes: WFR stands for withheld from randomization. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). *Random excess interaction share* is the excess interaction share calculated using only the set of declarations that were not withheld from randomization. Robust standard errors are presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. Initial refers to the value initially submitted in the customs IT system by the broker. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period November 18, 2017 to November 17, 2018.

Table 8: Differential Treatment During Re-Randomization Period

	After delegation of the re-randomization of inspector assignment				
Dependent variable	**Time**	**Fraud**	**Δlog value**	**Δlog tax**	**Hyp. tax revenue loss (%)**
A. Random excess interaction share (re-randomized declarations only)					
	(1)	(2)	(3)	(4)	(5)
Random excess interaction share	-0.264	-0.030	0.001	-0.007	-0.320***
	(0.531)	(0.114)	(0.036)	(0.026)	(0.094)
R-squared	15,899	15,925	13,963	15,907	15,918
Observations	0.227	0.394	0.254	0.275	0.164
B. Withheld from randomization					
	(6)	(7)	(8)	(9)	(10)
Withheld from randomization (WFR)	-0.853***	-0.015	-0.003	-0.003	0.175***
	(0.101)	(0.018)	(0.005)	(0.005)	(0.023)
Observations	16,455	17,169	15,189	17,147	17,159
R-squared	0.239	0.389	0.246	0.271	0.191
C. Excess interaction share					
	(11)	(12)	(13)	(14)	(15)
Excess interaction share	-2.352***	-0.187***	-0.037*	-0.051**	0.421***
	(0.464)	(0.042)	(0.017)	(0.020)	(0.051)
Observations	16,453	17,167	15,187	17,145	17,157
R-squared	0.232	0.390	0.246	0.271	0.185
D. Excess interaction share and declarations withheld from re-randomization (and their interaction)					
	(16)	(17)	(18)	(19)	(20)
Excess interaction share	-0.639***	0.007	0.002	0.004	0.129***
	(0.087)	(0.016)	(0.005)	(0.004)	(0.022)
Withheld from randomization (WFR)	-1.060**	-0.075	-0.011	-0.014	-0.097
	(0.416)	(0.053)	(0.016)	(0.015)	(0.058)
WFR*Excess interaction share	-2.542***	-0.370***	-0.084*	-0.128***	0.956***
	(0.788)	(0.102)	(0.040)	(0.037)	(0.144)
Observations	16,453	17,167	15,187	17,145	17,157
R-squared	0.241	0.390	0.246	0.272	0.192

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). *Random excess interaction share* is the excess interaction share calculated using only the set of declarations that were not withheld from randomization. All specifications include the tax rate, the risk score, a dummy for the red channel, the share of value accounted for by differentiated products, a dummy indicating whether the declaration was mixed, and a dummy indicating the declaration was subject to valuation advice, as well as inspector fixed effects, broker fixed effects, source country fixed effects, HS 2-digit product fixed effects, and month-year fixed effects. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period November 18, 2017 to November 17, 2018.

Table 9: The Costs of Collusion

Estimates of counterfactual additional revenue yield in the absence of collusion

A. Before delegated re-randomization of inspector assignment

	Declarations with significant excess interaction		All declarations	
	Average per declaration		Average per declaration	
	USD	% of actual yield	USD	% of actual yield
Actual tax yield	$11,538		$10,432	
Additional counterfactual tax yield based on:				
Median import prices	$993	8.6%	$144	1.4%
Median import prices & port authority weight	$1,552	13.5%	$217	2.1%
Prices reported by countries exporting to Madagascar	$2,413	20.9%	$330	3.2%
Prices reported by countries exporting to Madagascar & port authority weight	$3,135	27.2%	$423	4.1%
Third-party valuation advice	$1,165	10.1%	$167	1.6%

B. After delegated re-randomization of inspector assignment

	Declarations with significant excess interaction		All declarations	
	Average per declaration		Average per declaration	
	USD	% of actual yield	USD	% of actual yield
Actual tax yield	$9,435		$10,749	
Additional counterfactual tax yield based on:				
Median import prices	$3,879	41.1%	$219	2.0%
Prices reported by countries exporting to Madagascar	$11,442	121.3%	$545	5.1%
Third-party valuation advice	$1,227	13.0%	$67	0.6%

Notes: Before delegated re-randomization of inspector assignment refers to the period January 1, 2015 to November 17, 2017. After delegated re-randomization of inspector assignment refers to the period November 18, 2017 to November 17, 2018.

11 Appendix (for Online Publication Only)

Table 10: Variable definitions, 1/3

Variable name	Variable definition and data source(s)
A. Collusion proxies	
Excess interaction share	Difference between the observed share of a given broker's declarations handled by an inspector in a given semester and the hypothetical share the inspector would be expected to handle if the declarations were conditionally randomly assigned (predicted using a multinomial distribution that should govern the assignment of import declarations to inspectors if official assignment procedures were abided by as explained in section 4). Defined at the inspector-broker-semester level. Source: Madagascar customs.
Significant excess interaction indicator	Indicator variable equal to 1 if the following two conditions are met: (i) the excess interaction share exceeds 0.05 and (ii) the probability that the allocation of import declarations to inspectors is random does not exceed 0.001. Defined at the inspector-broker-semester level. Source: Madagascar customs.
Withheld from randomization	Indicator variable equal to 1 if the random assignment of the declaration was declaration was not performed by third-party Gasynet even though it was supposed to, and 0 if it was. Source: GasyNet.
B. *Ex-ante* risk characteristics & other characteristics of declarations	
Tax rate	Sum of taxes (including tariffs as well as Value Added Taxes) that have to be paid divided by the import value retained by customs. Defined at the declaration level. Source: Madagascar customs.
Risk score	Score calculated by Gasynet that indicates the risk of tax evasion for the import declaration ranging from 1 (very low risk) to 9 (very high risk). Defined at the declaration level. Source: GasyNet.
Red channel dummy	Dummy variable equal to 1 if the customs risk management system routed the declaration to the frontline inspection channel (red channel) and 0 otherwise. Defined at the declaration level. Source: Madagascar customs.
Mixed shipment dummy	Dummy variable equal to 1 if the import declaration includes more than 1 HS 6-digit product and 0 otherwise. Defined at the declaration level. Source: Madagascar customs.
Differentiated share	Share of HS 6-digit products in the import declaration that are defined as differentiated according to the conservative classification by Rauch (1999). Defined at the declaration level. Source: Rauch (1999) and a concordance between HS 6-digit revision 2012 classification and SITC revision 2 classification from UN COMTRADE.
Valuation advice dummy	Indicator variable equal to 1 if GasyNet provided a valuation advice for this import declaration and 0 otherwise. Defined at the declaration level. Source: GasyNet.
Log initial value (usd)	Log of the initially declared import value in USD (converted from Ariary using monthly exchange rates calculated as an average of daily exchange rates from the Central Bank of Madagascar). Defined at the declaration level. Source: Madagascar customs.

Table 11: Variable definitions, 2/3

Variable name	Variable definition and data source(s)
B. *Ex-ante* risk characteristics & other charact. of declarations, con't	
Log initial weight (kg)	Log of the initially declared total weight (in kilograms). Defined at the declaration level. Source: Madagascar customs.
Log port authority weight (kg)	Log of the sum of the weight of all containers used to ship the goods included in the import declaration measured at the port upon arrival (in kilograms). Defined at the declaration level (for containerized declarations). Source: Madagascar International Container Terminal Services Limited (MICTSL).
Log initial unit price (relative to median import unit price)	Difference between the log of the initially submitted unit price by the importer or his representative (defined as the weighted average of the unit prices (values divided by weights) for all items included in the import declaration, with weights being the initially submitted weights for each item) and the log of the initial average internal reference price of the declaration. Defined at the declaration level. Source: Madagascar customs.
Initial hypothetical tax revenue loss (relative to median import prices)	Computed as log(1+ (total taxation rate $\times$ initial average internal reference price $\times$ initially submitted weight by the importer or his representative))-log(1 + (total taxation rate $\times$ initially submitted value by the importer or his representative)). Defined at the declaration level. Source: Madagascar customs.
C. Customs outcomes	
Log clearance time (hours)	Log of the difference in time (measured in hours) between the date of assessment of the declaration by the inspector and the date of assignment of the declaration to the inspector that cleared the declaration. Defined at the declaration level. Source: Madagascar customs.
Fraud	Dummy variable equal to 1 if the customs inspector identifies fraud in the import declaration and 0 otherwise. Defined at the declaration level. Source: Madagascar customs.
Δ log value	Difference between the log of the declaration value retained by customs and the log of the initially submitted value by the importer or his representative. Defined at the declaration level. Source: Madagascar customs.
Δ log tax	Difference between the log of the taxes paid (including tariffs, VAT) on the declaration and the log of taxes that should have been paid in the absence of customs controls (which equal taxes paid minus tax adjustment by the customs inspector). Variable is measured in percent. Defined at the declaration level. Source: Madagascar customs.
Hypothetical tax revenue loss (relative to median import prices)	Computed as log (1+ (tax rate $\times$ final average internal reference price $\times$ final weight retained by customs)) - log (1 + (tax rate $\times$ final value retained by customs)). Defined at the declaration level. Source: Madagascar customs.

Table 12: Variable definitions, 3/3

Variable name	Variable definition and data source(s)
C. Customs outcomes, con't	
Δ log weight	Difference between the log of the final weight retained by customs and the log of the initially submitted weight by the importer or his representative. Variable is measured in percent. Defined at the declaration level. Source: Madagascar customs.
Weight gap (relative to port authority weight)	Difference between the log of the port authority weight and the log of the initially submitted weight by the importer or his representative. Defined at the declaration level. Source: Madagascar customs and Madagascar International Container Terminal Services Limited (MICTSL).
Hypothetical tax revenue loss (relative to median import prices & port authority weight)	Computed as log (1+ (tax rate $\times$ final average internal reference price $\times$ port authority weight)) - log (1 + (tax rate $\times$ final value retained by customs)). Defined at the declaration level. Source: Madagascar customs.
Hypothetical tax revenue loss (relative to prices reported by countries exporting to Madagascar)	Computed as log (1+ (tax rate $\times$ average final external reference price $\times$ final weight retained by customs)) - log (1 + (tax rate $\times$ final value retained by customs)). Defined at the declaration level. Sources: Madagascar customs and UN COMTRADE.
Hypothetical tax revenue loss (relative to prices reported by countries exporting to Madagascar & port authority weight)	Computed as log (1+ (tax rate $\times$ average final external reference price $\times$ port authority weight)) - log (1 + (tax rate $\times$ final value retained by customs)). Defined at the declaration level. Sources: Madagascar customs and UN COMTRADE.
Hypothetical tax revenue loss (relative to price advised by the third-party)	Computed as log (1+ (tax rate $\times$ reference value F.O.B. suggested by third-party GasyNet)) - log (1 + (tax rate $\times$ final value retained by customs)). Defined at the declaration level. Sources: Madagascar customs and GasyNet.
D. Additional control variables	
Familiarity	Log of the number of declarations registered by a given broker and cleared by a given inspector in the six preceding months. Source: Madagascar customs.
Log Workload	Log of the number of declarations assigned to the inspector in a given month. Source: Madagascar customs.
E. Auxiliary variables	
Internal reference price (IRP)	Median of unit price (ratio of value to quantity in kilograms) of a given HS 6-digit product from a given country of origin computed across all import declarations in Madagascar customs data in each year. Defined at the country-HS 6-digit product-year level. Source: Madagascar customs.
Initial [final] average internal reference price	Weighted average of the IRPs for all items included in the import declaration with weights being the initially submitted weights by the importer or his representative [final weights retained by customs] for each item. Defined at the declaration level. Source: Madagascar customs.
External reference price (ERP)	Unit price (ratio of value to quantity in kilograms) of a given product being exported by a given trading partner to Madasgascar. Defined at the country-HS 6-digit product-year level. Source: UN COMTRADE.
Initial [final] average external reference price	Weighted average of the ERPs for all items included in the import declaration with weights being the initially submitted weights by the importer or his representative [final weight retained by customs] for each item. Defined at the declaration level. Sources: Madagascar customs and UN COMTRADE.

Calculation of counterfactual tax revenue in the absence of collusion

To assess how much tax revenue is lost because of the collusion schemes we detect, we conduct a back of the envelope calculation of how much higher tax revenues would have been in the absence of excess interaction between brokers and inspectors.

Recall that our measure of hypothetical tax revenue loss, denoted *loss*, is constructed as the difference between log hypothetical tax yield (based on some reference price) and log actual tax yield: $loss = log(T^H) - log(T)$. Analogously, we can define hypothetical tax revenue loss *in the absence of collusion* as the difference between hypothetical tax yield (based on some reference price) and tax yield *in the absence of* collusion (which is the unknown variable we are interested in measuring): $loss^{NC} = log(T^H) - log(T^{NC})$. These two definitions in turn imply that we can write the log tax yield in the absence of collusion as:

$$log(T^{NC}) = log(T) - (loss^{NC} - loss) \tag{5}$$

Focusing on the period before the re-randomization intervention, we use the estimates of β_E presented in Panel B of Table 6 to obtain predicted hypothetical tax revenue losses *in the presence of* collusion as:

$$\widehat{loss} = \widehat{\beta_E} \times ES + \widehat{\beta_Z} Z \tag{6}$$

where the vector Z includes all explanatory variables other than ES.

We can use these same estimates to predict counterfactual tax revenue losses that would have materialized *in the absence* of collusion as:[30]

$$\widehat{loss^{NC}} = \widehat{\beta_Z} Z \tag{7}$$

Subtracting Equation (6) from Equation (7) we have $\widehat{loss^{NC}} - \widehat{loss} = -\widehat{\beta_E} ES$ and we can now compute counterfactual tax yield in the absence of excess interaction by plugging

[30] Note that we are simply recalculating predicted tax revenue losses while assuming excess interaction ES is 0 for each declaration.

$\widehat{\beta_E} ES$ into Equation (5) and taking exponents:

$$\widehat{T^{NC}} = T \times exp(\widehat{\beta_E} \times ES) \tag{8}$$

We construct measures of counterfactual tax yield $\widehat{T^{NC}}$ *in the absence of* collusion for each declaration using alternative estimates of β_E. We show the averages across declarations with significant excess interaction in the first two columns of Table 9 and the averages across all declarations in the last two columns of Table 9. Comparing these measures of counterfactual tax yield $\widehat{T}^{NC}$ *in the absence of* collusion to the actual tax yield provides an estimate of how much tax revenue would have been collected if excess interaction between inspectors and brokers was eliminated.

Focusing on the re-randomization period, we use the estimates of β_E, β_P, and β_{EP} presented in Appendix Table 21 and we follow the same logic as above to obtain the counterfactual tax yield *in the absence of* collusion analogously to what is done in Equation (8) as:

$$\widehat{T^{NC}} = T \times exp(\widehat{\beta_E} \times ES + \widehat{\beta_P} \times PM + \widehat{\beta_{EP}} \times ES \times PM) \tag{9}$$

Table 13: Descriptive Statistics

Before delegated re-randomization of inspector assignment						
	Other declarations			**Excess interaction declarations**		
	Mean	**SD**	**N**	**Mean**	**SD**	**N**
Excess interaction						
Excess interaction share	0.00	0.02	40,706	0.11	0.06	4,353
Risk characteristics						
Controls						
Tax rate	0.30	0.14	40,706	0.36	0.11	4,353
Risk score	6.51	2.87	40,189	7.51	2.01	4,333
Red channel dummy	0.28	0.45	40,706	0.31	0.46	4,353
Mixed shipment dummy	0.34	0.47	40,706	0.48	0.50	4,353
Differentiated share	0.68	0.46	40,706	0.79	0.39	4,353
Valuation advice dummy	0.08	0.27	40,706	0.22	0.42	4,353
Additional characteristics						
Log initial weight (kg)	10.00	1.73	40,706	10.24	1.10	4,353
Log initial value (usd)	10.08	1.18	40,706	10.15	0.85	4,353
Log port authority weight (kg)	10.55	1.08	23,530	10.50	0.92	2,766
Log initial unit price (relative to median import unit price)	-0.09	0.53	40,689	-0.17	0.53	4,345
Initial hypothetical tax revenue loss (relative to median import prices)	0.08	0.50	40,689	0.17	0.52	4,345
Customs outcomes						
Main outcomes						
Log clearance time (hours)	2.99	1.66	37,640	3.08	1.77	3,969
Fraud	0.08	0.27	40,706	0.14	0.34	4,353
Δ log value	0.02	0.08	40,631	0.03	0.09	4,340
Δ log tax	0.02	0.09	36,496	0.03	0.10	4,129
Hypothetical tax revenue loss (relative to median import prices)	0.08	0.49	40,689	0.16	0.52	4,345
Additional outcomes						
Weight gap (relative to port authority weight)	0.01	0.30	23,193	0.04	0.36	2,708
Hypothetical tax revenue loss (relative to median import prices & port authority weight)	0.09	0.55	40,700	0.23	0.63	2,593
Hypothetical tax revenue loss (relative to prices reported by countries exporting to Madagascar)	0.57	0.95	28,766	1.14	1.09	2,636
Hypothetical tax revenue loss (relative to prices reported by countries exporting to Madagascar & port authority weight)	0.36	1.14	14,882	1.00	1.15	1,593
Hypothetical tax revenue loss (relative to third-party valuation advice)	0.10	0.18	3,291	0.24	0.26	967

Notes: Excess interaction declarations are those that (i) are handled by an inspector whose excess interaction share (the difference between the share of given broker's declarations handled by the inspector in question and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors was conditionally random as prescribed by official rules) is at least 0.05 and (ii) for which we can reject the null hypothesis of conditional random assignment at the 0.001 significance level. The sample covers the period January 1, 2015 to November 17, 2017.

Table 14: Descriptive Statistics

After delegated re-randomization of inspector assignment						
	Re-randomized declarations			**Declarations withheld from randomization**		
	Mean	**SD**	**N**	**Mean**	**SD**	**N**
Excess interaction						
Excess interaction share	0.01	0.03	16,461	0.07	0.09	1,275
Risk characteristics						
Controls						
Tax rate	0.29	0.14	16,463	0.38	0.09	1,275
Risk score	6.14	2.97	15,927	7.31	2.11	1,242
Red channel dummy	0.17	0.38	16,463	0.04	0.20	1,275
Mixed shipment dummy	0.30	0.46	16,463	0.52	0.50	1,275
Differentiated share	0.68	0.46	16,463	0.77	0.39	1,275
Valuation advice dummy	0.08	0.27	16,463	0.17	0.38	1,275
Additional characteristics						
Log initial weight (kg)	9.99	1.79	16,463	10.15	1.00	1,275
Log initial value (usd)	10.66	1.92	6,723	10.06	1.61	515
Log initial unit price (relative to median import unit price)	-0.05	0.47	16,456	-0.24	0.53	1,272
Initial hypothetical tax revenue loss (relative to median import prices)	0.05	0.46	16,456	0.25	0.52	1,272
Customs outcomes						
Main outcomes						
Log clearance time (hours)	3.65	1.30	16,434	2.79	1.68	563
Fraud	0.09	0.28	16,463	0.12	0.32	1,275
Δ log value	0.02	0.08	16,440	0.02	0.09	1,271
Δ log tax	0.02	0.10	14,389	0.03	0.10	1,254
Hypothetical tax revenue loss (relative to median import prices)	0.04	0.45	16,456	0.24	0.52	1,272
Additional outcomes						
Hypothetical tax revenue loss (relative to prices reported by countries exporting to Madagascar)	0.49	0.91	12,025	1.40	1.04	730
Hypothetical tax revenue loss (relative to third-party valuation advice)	0.12	0.16	1,288	0.22	0.20	220

Notes: Re-randomized declarations are those for whom the assignment of the initial inspector was (re-)randomized by GasyNet. Declarations withheld from randomization are those that were not re-randomized because they were withheld from randomization by the customs IT department. The sample covers the period November 18, 2017 to November 17, 2018.

Table 15: Determinants of Excess Interaction Before Delegated Re-Randomization of Inspector Assignment

Dependent variable	**Excess interaction share**			
	(1)	(2)	(3)	(4)
Tax rate	0.025**	0.020***	0.020***	0.011***
	(0.009)	(0.005)	(0.005)	(0.003)
Risk score dummy	0.001***	0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Mixed shipment dummy	-0.000	-0.000	-0.000	0.001**
	(0.002)	(0.001)	(0.001)	(0.000)
Differentiated share	0.002	-0.000	-0.000	0.001
	(0.002)	(0.001)	(0.001)	(0.001)
Valuation advice dummy	0.014**	0.009***	0.009***	0.005***
	(0.006)	(0.003)	(0.003)	(0.001)
Red channel dummy	0.000	-0.001	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)
Log initial value (usd)			0.000	-0.000
			(0.000)	(0.000)
Log initial unit price (relative to median import unit price)			-0.002**	-0.001**
			(0.001)	(0.001)
Month-year fixed effects	Yes	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes	Yes
Inspector fixed effects	No	Yes	Yes	Yes
Broker fixed effects	No	Yes	Yes	Yes
Importer fixed effects	No	No	No	Yes
Observations	44,522	44,522	44,497	43,669
R-squared	0.072	0.225	0.226	0.377

Notes: Robust standard errors are presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 16: Differential Treatment by Inspectors - Additional Robustness Tests

	Before delegated re-randomization of inspector assignment				
Dependent variable	**Time**	**Fraud**	**Δ log value**	**Δ log tax**	**Hyp. tax revenue loss (%)**
A. Controlling for inspector-month and broker-month fixed effects					
	(1)	(2)	(3)	(4)	(5)
Excess interaction share	-1.869***	-0.186***	-0.057***	-0.068***	0.254**
	(0.215)	(0.052)	(0.014)	(0.020)	(0.099)
Observations	41,098	44,510	44,423	40,449	44,485
R-squared	0.344	0.280	0.218	0.207	0.258
B. Using an alternative measure of excess interaction					
	(6)	(7)	(8)	(9)	(10)
Excess interaction dummy	-0.246***	-0.028**	-0.009***	-0.010**	0.030**
	(0.044)	(0.011)	(0.003)	(0.004)	(0.012)
Observations	41,121	44,522	44,435	40,471	44,497
R-squared	0.265	0.211	0.150	0.130	0.210
C. Controlling for importer-broker fixed effects					
	(11)	(12)	(13)	(14)	(15)
Excess interaction share	-2.006***	-0.134**	-0.054***	-0.062***	0.235***
	(0.276)	(0.058)	(0.017)	(0.021)	(0.064)
Observations	39,422	42,761	42,674	38,803	42,738
R-squared	0.379	0.348	0.325	0.326	0.447

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). All specifications include the tax rate, the risk score, a dummy for the red channel, the share of value accounted for by differentiated products, a dummy indicating whether the declaration was mixed, and a dummy indicating the declaration was subject to valuation advice, as well as source country fixed effects and HS 2-digit product fixed effects. The specifications in Panel A include additionally inspector-month and broker-month fixed effects, those in Panel B include additionally inspector, broker and month-year fixed effects, and those in Panel C include additionally inspector and month-year fixed effects. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 17: Differential Treatment by Inspector - Heterogeneity

Before delegated re-randomization of inspector assignment					
Dependent variable	**Time**	**Fraud**	**Δ log value**	**Δ log tax**	**Hyp. tax revenue loss (%)**
A. Interaction with tax rate					
	(1)	(2)	(3)	(4)	(5)
Excess interaction share	-1.974**	0.023	-0.100	-0.010	-0.199
	(0.725)	(0.067)	(0.067)	(0.036)	(0.174)
Tax rate	0.913***	0.186***	0.040***	0.048***	0.095*
	(0.144)	(0.020)	(0.010)	(0.008)	(0.047)
Excess interaction share × Tax rate	-0.100	-0.862***	0.040	-0.196	1.698***
	(2.275)	(0.183)	(0.220)	(0.131)	(0.413)
Declaration characteristics	Yes	Yes	Yes	Yes	Yes
Inspector fixed effects	Yes	Yes	Yes	Yes	Yes
Broker fixed effects	Yes	Yes	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	41,121	44,522	40,471	44,435	44,497
R-squared	0.318	0.214	0.132	0.152	0.211
B. Interaction with initial hypothetical tax revenue loss					
	(6)	(7)	(8)	(9)	(10)
Excess interaction share	-2.081***	-0.239***	-0.071***	-0.083***	0.081***
	(0.261)	(0.072)	(0.017)	(0.024)	(0.016)
Initial hypothetical tax revenue loss	0.137***	0.024***	0.015***	0.016***	0.970***
(IHTRL)	(0.017)	(0.004)	(0.002)	(0.002)	(0.003)
Excess interaction share × IHTRL	0.232	-0.238*	-0.061*	-0.032	0.085
	(0.379)	(0.119)	(0.034)	(0.037)	(0.073)
Declaration characteristics	Yes	Yes	Yes	Yes	Yes
Inspector fixed effects	Yes	Yes	Yes	Yes	Yes
Broker fixed effects	Yes	Yes	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	41,098	44,497	44,421	40,452	44,497
R-squared	0.319	0.214	0.157	0.137	0.969

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). Declaration characteristics include the tax rate (whose coefficient is shown in Panel A), the risk score, a dummy for the red channel, the share of value accounted for by differentiated products, a dummy indicating whether the declaration was mixed, and a dummy indicating the declaration was subject to valuation advice. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 18: Prevalence of Reassignments across Inspectors

Before delegated re-randomization of inspector assignment			
	% of all declara-tions	**% of all reassign-ments**	**Probability of reassignment conditional on initial state**
Declarations without initial excess interaction	**90.9%**		
Not reassigned	84.8%		
Reassigned to inspector without excess int. (RNN)	5.8%	90.0%	6.4%
Reassigned to inspector with excess int. (RNE)	0.3%	5.2%	0.4%
Declarations with initial excess interaction	**9.0%**		
Not reassigned	8.7%		
Reassigned to inspector without excess int. (REN)	0.3%	4.1%	3.0%
Reassigned to inspector with excess int. (REE)	0.0%	0.7%	0.4%
Any reassignment	6.4%		

Notes: The table considers reassignments across inspectors made by the customs port manager. Reassignments No Excess → No Excess (RNN) are those in which a declaration is taken from an inspector who did not act excessively frequently with the broker handling the declaration and is assigned to another inspector who did not interact excessively frequently with the broker either. Reassignments Excess → No Excess (REN) are those in which a declaration is taken away from an inspector who was interacting excessively frequently with the broker in question and is assigned to an inspector who was not. Reassignments RNE, and REE are defined analogously. The sample covers the period January 1, 2015 to November 17, 2017.

Table 19: Impact of Reassignments on Customs Outcomes

	Before delegated re-randomization of inspector assignment				
Dependent variable	**Time**	**Fraud**	**Δ log value**	**Δ log tax**	**Hyp. tax revenue loss (%)**
	(1)	(2)	(3)	(4)	(5)
Excess interaction share	-2.038***	-0.264***	-0.074***	-0.080***	0.384***
	(0.291)	(0.057)	(0.018)	(0.026)	(0.106)
Reassignments					
No Excess $\rightarrow$ No Excess (RNN)	-0.339***	0.041***	0.010***	0.014***	0.003
	(0.059)	(0.009)	(0.002)	(0.003)	(0.010)
No Excess $\rightarrow$ Excess (RNE)	-0.913**	0.012	0.006	0.004	-0.007
	(0.406)	(0.029)	(0.011)	(0.014)	(0.059)
Excess $\rightarrow$ No Excess (REN)	-0.269	0.109***	0.054***	0.051***	-0.084
	(0.202)	(0.032)	(0.014)	(0.014)	(0.053)
Excess $\rightarrow$ Excess (REE)	0.456**	0.172**	0.044	0.054*	0.059
	(0.184)	(0.084)	(0.027)	(0.027)	(0.112)
Declaration characteristics	Yes	Yes	Yes	Yes	Yes
Inspector fixed effects	Yes	Yes	Yes	Yes	Yes
Broker fixed effects	Yes	Yes	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes
P-values					
Test for difference (RNN)=(RNE)	0.145	0.287	0.703	0.454	0.859
Test for difference (REN)=(RNN)	0.735	0.023	0.003	0.015	0.126
Test for difference (RNE)=(REE)	0.004	0.092	0.238	0.135	0.520
Observations	41,121	44,522	44,435	40,471	44,497
R-squared	0.269	0.214	0.152	0.133	0.211

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. The table considers reassignments across inspectors made by the customs port manager. Reassignments No Excess→ No Excess (RNN) are those in which a declaration is taken from an inspector who did not act excessively frequently with the broker handling the declaration and is assigned to another inspector who did not interact excessively frequently with the broker either. Reassignments Excess → No Excess (REN) are those in which a declaration is taken away from an inspector who was interacting excessively frequently with the broker in question and is assigned to an inspector who was not. Reassignments RNE, and REE are defined analogously. "Observations" refers to the number of non-singleton observations. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 20: Excess Interaction and Tax Revenue Losses - Assessing Selection Bias

	Before delegated re-randomization of inspector assignment					
Dependent variable	**Hyp. tax revenue loss (%) (Relative to median import unit price - port authority weight not considered)**					
Sample						
Reference price	—	Importers		Exporters		Third-party
Port authority weight	Yes	No	Yes	No	Yes	—
	(1)	(2)	(3)	(4)	(5)	(6)
Excess interaction share	0.953***	0.730***	1.020***	0.883***	1.210***	1.232***
	(0.199)	(0.199)	(0.193)	(0.198)	(0.183)	(0.198)
Exogenous decl. characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,900	45,033	23,965	31,401	16,475	4,249
R-squared	0.227	0.180	0.235	0.102	0.155	0.240

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). Declaration characteristics include the tax rate, the risk score, a dummy for the red channel, the share of value accounted for by differentiated products, a dummy indicating whether the declaration was mixed, and a dummy indicating the declaration was subject to valuation advice. Exogenous declaration characteristics include the tax rate, the share of value accounted for by differentiated products, and a dummy indicating whether the declaration was mixed. "Importers", "Exporters" and "Third-party" refer, respectively, to median import unit prices, prices reported by countries exporting to Madagascar, and transaction-specific valuation advice provided by the third-party GasyNet based on its own proprietary data. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period January 1, 2015 to November 17, 2017.

Table 21: Excess Interaction and Tax Revenue Losses

After the delegated re-randomization of inspector assignment			
Dependent variable	**Hypothetical revenue loss (%)**		
Reference price	Importers	Exporters	Third party
	(1)	(2)	(3)
Excess interaction share	0.146***	0.267***	0.024**
	(0.022)	(0.051)	(0.010)
Withheld from randomization (WFR)	0.043	0.021	0.184
	(0.140)	(0.230)	(0.178)
WFR × Excess interaction share	1.279***	3.328***	0.479*
	(0.185)	(0.601)	(0.242)
Exogenous declaration characteristics	Yes	Yes	Yes
Source country fixed effects	Yes	Yes	Yes
HS 2-digit product fixed effects	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes
Observations	17,726	12,753	1,508
R-squared	0.149	0.544	0.436

Notes: Standard errors are clustered by inspector and presented in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. *Excess interaction share* is the difference between the share of given broker's declarations handled by an inspector in a given semester and the hypothetical share she would be expected to handle if the allocation of declarations to inspectors were random conditional on their productivity, as prescribed by the official assignment rules (as explained in Section 4). Declaration characteristics include the tax rate, the share of value accounted for by differentiated products, a dummy indicating whether the declaration was mixed. "Importers", "Exporters" and "Third-party" refer, respectively, to median unit import prices, prices reported by countries exporting to Madagascar, and transaction-specific valuation advice provided by the third-party GasyNet based on its own proprietary data. "Observations" refers to the number of non-singleton observations. OLS estimation is used. The sample covers the period November 18, 2017 to November 17, 2018.

Table 22: Persistence of Collusion

Correlation Matrix						
		Withheld from randomization (WFR)	Lagged WFR	Excess interaction share	Lagged excess interaction share	Random excess interaction share
Lagged WFR	ρ	0.342***				
	N	415				
Excess interaction share	ρ	0.327***	0.287***			
	N	987	415			
Lagged excess interaction share	ρ	0.088**	0.392***	0.343***		
	N	816	415	3339		
Random excess interaction share	ρ	0.002	-0.017	0.824***	0.127***	
	N	985	415	985	816	
Lagged random excess inter. share	ρ	0.047	0.140**	0.073	0.934***	-0.014
	N	415	415	415	415	415

Notes: ***, **, and * indicate significance at 1%, 5%, and 10% levels, respectively. The unit of observation used for the calculation of the correlations is the average across declarations handled by an inspector-broker pair in a given semester. The sample covers the period January 1, 2015 to November 17, 2018.