NBER WORKING PAPER SERIES

STRATEGIC OR CONFUSED FIRMS?
EVIDENCE FROM "MISSING" TRANSACTIONS IN UGANDA

Miguel Almunia
Jonas Hjort
Justine Knebelmann
Lin Tian

Working Paper 29059
http://www.nber.org/papers/w29059

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2021

We thank editor Rema Hanna, three anonymous referees, Michael C. Best, Ray Fisman, Francois Gerard, Jim Hines, Wojciech Kopczuk, David Margolis, Joana Naritomi, Daniel Reck, Floris Zoutman, and many seminar and conference participants for comments and suggestions that substantially improved the paper; Pablo Garcia Guzman, David Henning, and Claude Raisaro for outstanding research assistance; the Uganda Revenue Authority for data sharing and excellent collaboration; and the International Growth Centre, the British Academy, and the Leverhulme Foundation for funding. The views in this paper are those of the authors, and do not necessarily represent those of the Uganda Revenue Authority or the National Bureau of Economic Research.

Strategic or Confused Firms? Evidence from "Missing" Transactions in Uganda
Miguel Almunia, Jonas Hjort, Justine Knebelmann, and Lin Tian
NBER Working Paper No. 29059
July 2021
JEL No. D2,D9,H2,O1

ABSTRACT

Are firms sophisticated maximizers, or do they consistently make errors? Using transaction-level data from Ugandan value-added tax (VAT) returns, we show that sellers and buyers report different amounts 79% of the time, despite invoices being easily cross-checked. We estimate that 25% of firms are disadvantageous misreporters—they systematically misreport own sales and purchases such that their tax liability increases—while 75% are advantageous misreporters. Many firms—especially disadvantageous misreporters—fail to report imported inputs they themselves reported at Customs, increasing their VAT liability. On net, unilateral VAT misreporting cost Uganda about US$384 million in foregone 2013-2016 tax revenue

Miguel Almunia
Department of Economics
CUNEF
Spain
miguel.almunia@gmail.com

Jonas Hjort
Graduate School of Business
Columbia University
3022 Broadway, Uris Hall 622
New York, NY 10027
and NBER
hjort@columbia.edu

Justine Knebelmann
Paris School of Economics
48 Boulevard Jourdan
Paris 75014
France
justine.knebelmann@psemail.eu

Lin Tian
INSEAD
1 Ayer Rajah Avenue Singapore
138676
lin.tian@insead.edu

An online appendix is available at http://www.nber.org/data-appendix/w29059

1 Introduction

In economics, firms are seen as sophisticated organizations—maximizers that make constrained but optimal decisions by carefully assessing the true costs and benefits to themselves. This assumption underlies the models that guide our understanding of how firms behave. Strategic decision-making by firms is by and large taken as self-evident.

There is, however, growing evidence that some firms deviate from optimal behavior.[1] If a significant proportion consistently makes mistakes, the consequences for theory and policy design would be far-reaching. Consider how firms in low-income countries should be taxed—one of the most important questions for economic development (Besley & Persson, 2009; Kleven *et al.*, 2016). The value-added tax (VAT)—now in use in 166 countries around the world—is popular among economists in part because of its enforcement properties. In firm-to-firm transactions, the seller and buyer face asymmetric (mis)reporting incentives and their reports can easily be cross-checked (Ebrill *et al.*, 2001; Kopczuk & Slemrod, 2006; Pomeranz, 2015). This is thought to make the VAT "self-enforcing," but the argument assumes a degree of cross-checking capacity and, more fundamentally, that firms infer the likelihood of such checks and accurately keep track of their sales and purchases.

In this paper, we study the sophistication of firms' decision-making in a low-income country context by analyzing their tax reporting behavior. We use 2013-2016 transaction-level VAT and Customs records on all domestic and international trade involving the 22,388 VAT-registered firms in Uganda. In the first part of our analysis, we document that sellers and buyers report different transacted amounts in 79% of reported firm-pair×month VAT observations. In 60% of mismatch transactions we find a *seller shortfall*, namely the seller reporting the lower value, and in the remaining 40% a *buyer shortfall*. The latter cases are harder to rationalize since the buyer reporting less than the seller raises one or both firms' tax liability, other things equal.

In the second part of our analysis, we develop a fixed-effects methodology that estimates what fraction of each reporting discrepancy can be attributed to the seller vs. the buyer, holding constant each firm's identity and those of its other trade partners. Combining individual firms' estimated reporting discrepancies as buyer and seller in turn allows us to categorize their reporting behavior. Some overreport total purchases and/or underreport total sales such that the firm's overall liability decreases—what we interpret as strategic behavior in a low-enforcement context and label *advantageous* misreporting; and some make systematic *disadvantageous* reporting mistakes that increase the firm's overall

[1] See, among others, Hortacsu & Puller (2008); Cho & Rust (2010); Goldfarb & Xiao (2011); DellaVigna & Gentzkow (2019); Kremer *et al.* (2019); Hjort *et al.* (2020); Dube *et al.* (2020); Tourek (2021).

liability.[2]

We find that 75% of VAT-registered Ugandan firms are advantageous misreporters and 25% are disadvantageous misreporters. Among advantageous misreporters, 10% "look small" by underreporting both sales and purchases and the firm's value-added (a form of fly-under-the-radar behavior first identified by Carrillo *et al.* (2017) in Ecuador). Another 78% are "conspicuous" advantageous misreporters that underreport their sales and overreport their purchases. The remaining 12% "look big" by overreporting both sales and purchases. Over time, 74% (65%) of firms classified as advantageous (disadvantageous) remain in the same category as in the previous year.

In a series of robustness checks, we analyze several ways in which our estimates could under- or overestimate the prevalence of reporting mistakes. We re-estimate our model assuming extensive final sales underreporting, finding that the proportion of disadvantageous firms remains large. When we restrict to firms for which we can reject liability-neutral tax reporting at conventional significance levels, the firm classification is very similar to the baseline, with 23% of disadvantageous firms. Finally, event studies looking at firms switching trade partners strongly substantiate a causal interpretation of the fixed-effects model estimates.

In the third part of our analysis we consider how sophisticated and less sophisticated firms behave in higher state capacity contexts. The case for the VAT assumes some degree of capacity to cross-check firms' tax reports. Our results suggest that low-income countries may not have such capacity. However, like models of firms' response to other public policies, the self-enforcing VAT hypothesis ultimately rests on a more fundamental assumption: that firms behave strategically. Mis-optimizing firms may not respond as anticipated to enforcement incentives.

To investigate, we take advantage of goods being more closely monitored when moving through Customs.[3] We compare an import transaction report at Customs versus the *same firm's* report of the same transaction on the credit side of its domestic VAT records. While, as expected, double reports are more consistent when the same firm makes both reports and one of the two is at Customs, we find discrepancies in a remarkable 48% of such cases. In particular, we again find evidence of firm mistakes. Firms reduce their tax liability by overreporting their imported inputs in VAT returns in 14% of import transactions, while they increase their liability by underreporting in VAT returns in 34% of trans-

[2]We interpret *systematic* underreporting of a firm's liability as strategic behavior and systematic overreporting of a firm's liability as mistakes. By classifying any systematic, self-advantageous reporting errors as strategic behavior, we possibly underestimate the true extent of reporting mistakes.

[3]It is well documented that tariffs are more stringently enforced than domestic taxes, perhaps because goods have to physically clear Customs (Riezman & Slemrod, 1987; Keen & Lighart, 2002; Emran & Stiglitz, 2005; Keen & Lighart, 2005; Baunsgaard & Keen, 2010; Cagé & Gadenne, 2018).

actions. Importantly, the latter form of disadvantageous behavior is significantly more common among firms classified as disadvantageous misreporters in domestic VAT data.

Overall, our findings suggest that the majority of Ugandan firms are sophisticated enough to respond to weak tax enforcement by considerably underreporting their tax liability, as conventional models of firm behavior assume. However, a non-negligible proportion consistently make costly errors. We quantify the consequences for tax collection, accounting for each firm's misreporting and outstanding VAT liability position. We estimate that the government revenue *gain* due to reporting errors by disadvantageous misreporters is large—around US$138 million during 2013-2016. However, the revenue loss due to misreporting by advantageous misreporters is even larger, at around US$522 million. On net, unilateral VAT misreporting cost the Ugandan government around US$384 million, or 4% of total tax revenue collected, during 2013-2016.

This paper provides what to our knowledge are the first direct estimates of the extent of mistakes in an economy-wide population of firms. The methodology we develop allows us to classify individual firms' behavior as self-advantageous or not, and we observe the entire population of formal, non-micro firms in Uganda's economy. Our analysis builds on an emerging body of evidence of seemingly erroneous firm behavior (see footnote 1).[4]

We also contribute new evidence on how tax evasion responds to the state's enforcement capacity, and in particular how firms characterized by different degrees of sophistication respond. In this sense, our analysis builds most closely—methodologically and thematically—on Fisman & Wei (2004)'s "mirror" data approach to measuring how tariff evasion responds to the tariff rate. However, our focus is on variation in enforcement capacity, linking our analysis with existing work on the causes and consequences of state capacity (Besley & Persson, 2009, 2010; Acemoglu *et al.*, 2015; Page & Pande, 2018; Best *et al.*, 2019). We also build on existing studies of more-vs.-less attentive taxpayers' response to tax rates.[5]

Finally, we show evidence that the VAT is far from self-enforcing in low state capacity settings. This qualifies the common argument that developing countries are especially likely to benefit from use of the VAT (see, e.g., Bird & Gendron, 2007).[6] In doing so, our

[4] Tourek (2021) documents another form of seemingly suboptimal taxpayer behavior—firms reporting identical amounts in their income tax year after year—in neighboring Rwanda.

[5] Chetty *et al.* (2009); Aghion *et al.* (2017); Benzarti (2020); Gillitzer & Skov (2018); Rees-Jones & Taubinsky (2018) provide direct evidence of tax-reporting mistakes by *individuals* (see also Reck (2016)). Like this paper, Aghion *et al.* (2017) show evidence that more sophisticated taxpayers tend to react as theory predicts to tax incentives, while less sophisticated taxpayers do so to a lesser extent.

[6] Tax evasion research has demonstrated the importance of third-party reporting (Slemrod *et al.*, 2001; Kleven *et al.*, 2011; Kleven, 2014), but also its limitations (Pomeranz, 2015; Carrillo *et al.*, 2017; Slemrod *et al.*, 2017; Almunia & Lopez-Rodriguez, 2018; Waseem, 2018). The existing literature shows that in middle-income countries whose enforcement capacity significantly exceeds Uganda's, authorities' ability to cross-

analysis builds on work studying how policy should be tailored to context (see, e.g., Laffont, 2005; Best *et al.*, 2015, 2019; Duflo *et al.*, 2018; Hansman *et al.*, 2019). The massive magnitude of the revenue loss from VAT evasion we document in Uganda—and the corresponding cross-country patterns in Cagé & Gadenne (2018)—suggests that the production efficiency benefits of VATs relative to tariffs are at least in part offset by capacity-constrained governments' ability to raise revenue on domestic transactions.

2 Background

Uganda's tax-to-GDP ratio, at 13% in 2016, is below the African and OECD averages of 18 and 34% (OECD, 2018), while its tax administration costs (2.4% of tax revenues) are similar to other low-income countries (IMF, 2013; Lemgruber *et al.*, 2015).

The VAT was introduced in 1996 and in 2016 contributed 32% of Uganda's total non-tariff tax revenue, similar to elsewhere in Africa (OECD, 2018). Its design is standard, with a general rate of 18%, a credit-invoice system, standard exemptions (e.g., financial services), and zero-rating (e.g., exports). Appendix A provides details.

Since 2012 all VAT-registered firms must file their monthly VAT declarations electronically, within 15 days of the transaction month ending.[7] These must include detailed transaction-level records—spreadsheets listing each sale to and purchase from other VAT-registered firms. This implies that the Uganda Revenue Authority (URA) receives two reports for each transaction between any two VAT-registered firms.

Our analysis exploits the complete administrative data from VAT-registered firms' declarations between 2013 and 2016.[8] The monthly firm-level VAT data include a scrambled Tax Identification Number (TIN), the declaration date, total sales/purchases (amount and VAT charged/paid), total VAT liabilities, and data from the spreadsheets—called VAT "schedules"—detailing each transaction. The schedules include the transaction date, the seller and buyer TINs, the transaction value, and the VAT charged or paid. Schedule 1 (VS1) contains all sales transactions to other VAT-registered firms. Sales to final consumers or non-VAT firms are recorded only as a monthly aggregate. Schedules 2, 3, and 4 contain domestic input purchases, imports, and administrative expenses, respectively. Importantly, the transaction-level records reported in the VAT schedules constitute mean-

check VAT records tends to reduce evasion (Ebrill *et al.*, 2001; Pomeranz, 2015; Carrillo *et al.*, 2017; Mittal & Mahajan, 2017; Waseem, 2020; Naritomi, 2019; Fan *et al.*, 2019). Discrepancies in VAT declarations comparable to what we observe in Uganda are found in Rwanda (Mascagni *et al.*, 2019).

[7] About 80% of VAT returns are reported within 15 days of the return month and another 9% within the next month.

[8] We refer to fiscal year 2013/14 as 2013.

ingful paper trails: they are consistent with the firm-level reports in 97% of cases.

Our dataset contains 22,388 unique VAT-registered firms submitting at least one monthly VAT return between 2013 and 2016, and the transactions data cover 15,569 sellers and 19,421 buyers, leading to 3,373,183 seller-buyer-month observations.[9]

The data on imports comes from Customs declarations submitted to the URA between 2012 and 2016. These are transaction-specific, submitted electronically, and include the value of the goods imported, the type and number of items, and the date of import. The TIN of the importer allows us to match the Customs data to the domestic VAT data. 9,998 VAT-registered firms import at least once.

3 Discrepancies in VAT Declarations

In this section, we document massive VAT reporting discrepancies in Uganda at the seller-buyer-month level.

3.1 Conceptual background

For a date j transaction, let y^S_{sbj} and y^B_{sbj} denote the output VAT charged (as reported by the seller s) and the input VAT paid (as reported by the buyer b). We aggregate transactions at the monthly level and define $Y^S_{sbt} \equiv \sum_{j \in J_t} y^S_{sbj}$ and $Y^B_{sbt} \equiv \sum_{j \in J_t} y^B_{sbj}$ where t denotes the transaction month. We define *seller shortfall* as the total VAT charged being *lower* than the total VAT paid, i.e., $Y^S_{sbt} < Y^B_{sbt}$, and *buyer shortfall* as $Y^S_{sbt} > Y^B_{sbt}$.

Seller shortfall may be due to the seller underreporting output VAT or the buyer overreporting input VAT (or both). In either case, it implies a potential financial gain for one or both firms, as the reported tax liability is lower than the true liability. Symmetrically, buyer shortfall may be due to the seller overreporting output VAT or the buyer underreporting input VAT (or both), which implies a potential, eventual financial *loss* for one or both firms.[10]

Other things equal, buyer shortfall points towards mistakes in firms' VAT declarations. However, it might be rational for buyers to understate their purchases if they simultaneously understate their sales, e.g., because this allows them to report a less suspicious (say, nonnegative) VAT liability. Carrillo *et al.* (2017) provide evidence of such "looking small"

[9] Out of 22,388 firms, 19,137 have non-missing firm-as-buyer and/or firm-as-seller fixed-effect estimated as described in Section 4 and therefore make up our main sample of analysis.

[10] This is true also in cases where a firm fully reporting its credits vis-a-vis the URA will not reduce its *current* dues, e.g. because of an (already-) nil or negative liability. Reporting negative VAT liabilities and carrying offsets forward is significantly associated with a lower probability of having a positive VAT liability in the future, both across and within firms.

behavior in Ecuador. Buyer shortfall cases could also be due to sellers engaging in also-liability-reducing "looking big" behavior by overstating both their purchases and sales—perhaps due to beliefs that the tax authority pays more attention to small than big firms (see, e.g., Amodio *et al.*, 2021)—while underreporting their value added. In and of themselves, transaction-pair level discrepancies thus do not allow us to distinguish between sophisticated, self-advantageous tax evasion and reporting mistakes.

3.2 Discrepancies

Ugandan firms' average monthly reported VAT liability for the 2013-2016 period is slightly negative, and the median is zero, as is common in developing countries (Lemgruber *et al.*, 2015; Pomeranz, 2015). While only 15% of firms report negative or zero value added in a full fiscal year, the reported VAT *liability* is zero or negative for 52% of firms (see Table F.1). This proportion is quite similar across firms of different sizes. Many can report positive value added but zero or negative VAT liability. This is because offsets are typically carried over, since refunds are restricted.

We observe seller shortfall in 47% and buyer shortfall in 32% of seller-buyer-month observations, with sellers and buyers reporting the same amount in only 21% of the observations.[11] Figure 1 provides a graphical illustration of these discrepancies. In the left panel, the vertical axis measures the (inverse hyperbolic sine of the) total monthly amounts declared by sellers, and the horizontal axis that of buyers. The data are grouped into a grid where the color of each square represents the number of observations, going from 1 (lightest gray) to more than 50,000 (black). Observations above (below) the 45-degree line correspond to cases of buyer (seller) shortfall. The figure's right panel displays the distribution of reporting discrepancies.

We observe these widespread discrepancies despite taking a number of steps to minimize mismatched transactions. First, we use transaction dates rather than filing dates. Second, we use firms' aggregate monthly records rather than individual transactions, and do not label cases where the seller and buyer declare the same amount, only with a one or two-month lag, as discrepancies. Finally, we allow for rounding errors of 1,000 Ugandan Shillings (about US$0.30).[12]

In Figure 1a, squares on the 45-degree line correspond to observations where seller and buyer-reported amounts match. The dashed curve shows the average amount re-

[11] At the quarterly level, we find discrepancies in 84% of cases, with seller shortfall in 50% of cases and buyer shortfall in 34% of cases.

[12] Alternatively, we consider rounding the value of discrepancies at 5% of the transaction value. The share of discrepancies remains very close to the baseline level with similar proportions of seller and buyer shortfalls.

ported by sellers for different values of the buyer-reported amounts. We see that seller shortfall is quantitatively more important than buyer shortfall in aggregate terms. This is apparent also in the right panel, Figure 1b. The total amount of seller shortfall across all discrepancies is US$906 million, while the total amount of buyer shortfall is US$735 million.

Eighty-four percent of discrepancies are on the extensive margin—one trade partner fails to report transacting in a given month—while 16% are on the intensive margin. Variations in these proportions by firm characteristics are shown in Table G.2: overall these shares are relatively stable across sectors and firm size categories. The share of extensive margin discrepancies decreases with transaction size, but the fraction of the transaction amount unreported is higher for larger transactions.

4 Classifying Firms' Reporting Behavior

In this section we show that most Ugandan firms engage in strategic tax reporting behavior, taking into account the country's low-enforcement environment, as economic theory predicts. We also show that, in contrast, a sizeable minority makes costly reporting mistakes. To do this we evaluate whether firms underreport their value added such that their liability falls, or erroneously overreport value added.

4.1 Assigning the blame: fixed-effects analysis

We allocate a share of the responsibility for each discrepancy to the seller and the buyer based on each firm's aggregate reporting accuracy in their respective transactions. The starting point is a fixed-effects model inspired by Abowd *et al.* (1999, 2002). We define the discrepancy between buyer f, and seller f' in month t as $d_{ff't} \equiv Y^B_{ff't} - Y^S_{ff't}$ such that $d_{ff't} > 0$ implies seller shortfall and $d_{ff't} < 0$ implies buyer shortfall. Then, we estimate:

$$d_{ff't} = \delta_c + \delta^b_f + \delta^s_{f'} + \delta_t + r_{ff't}, \tag{1}$$

where δ^b_f and $\delta^s_{f'}$ denote buyer and seller fixed-effects (defined at the firm level), respectively; δ_t is a month fixed effect; δ_c is a constant, and $r_{ff't}$ is an error term. Since $d_{ff't}$ is the nominal value of the discrepancy, δ^s_f can be interpreted as a firm's average discrepancy as a seller, in monetary terms, controlling for all time-invariant characteristics of its buyers, such as their size and reporting reliability. Similarly, $\delta^b_{f'}$ can be interpreted as a firm's average contribution to discrepancies as a buyer, controlling for all time-invariant

characteristics of its sellers.[13]

As shown in Abowd *et al.* (1999, 2002), the two-dimensional fixed-effects are separately identified only within a "connected set"—firm-pairs that are linked by transaction and all of such firms' trade partners. The largest connected set observed during our 2013-2016 data period covers over 99% of all observations, 90% of sellers, and 94% of buyers. We thus restrict our analysis to this largest connected set of firms.

4.2 Firm-level reporting behavior

We now formalize our classification of firms' reporting behavior. We construct a firm-level discrepancy measure Q_f, adding up the firm's two estimated fixed-effects:

$$Q_f \equiv w_s \cdot \hat{\delta}_f^s + w_b \cdot \hat{\delta}_f^b, \tag{2}$$

where w_s and w_b represent the number of firm-trade partner monthly observations as a seller or buyer, respectively.[14] A firm engages in *advantageous* misreporting behavior if $Q_f > 0$, meaning that it reports in a way that reduces its aggregate VAT liability. Symmetrically, a firm engages in *disadvantageous* misreporting behavior if $Q_f < 0$, which implies that it reports in a way that increases its overall VAT liability.

We further classify advantageous misreporters into three subcategories. First, a firm engaging in *conspicuous* advantageous misreporting is one for which $\hat{\delta}_f^s \geq 0$ and $\hat{\delta}_f^b \geq 0$. This implies that the firm both underreports its sales and overreports its purchases. Second, a firm engaging in *looking-small* advantageous misreporting is one for which $\hat{\delta}_f^s \geq 0$ and $\hat{\delta}_f^b < 0$. This implies that the firm underreports its sales and underreports its purchases. Finally, a firm engaging in *looking-big* advantageous misreporting is one for which $\hat{\delta}_f^s < 0$ and $\hat{\delta}_f^b \geq 0$, thus overreporting its sales and its purchases.

Panel A of Table 1 shows the resulting classification of firms. We find that 14,358 of the 19,137 Ugandan VAT-eligible firms (75%) are *advantageous* misreporters. This suggests that when the VAT is implemented in a low-state capacity context without systematic cross-checks, the majority of firms misreport to lower their VAT liability.

Of the firms that misreport in an advantageous way, 78% are conspicuous advanta-

[13] In Table B.4 we show results from running (1) with various controls that affect the probability of two firms trading with each other. The results are very similar.

[14] More precisely, $\hat{\delta}_f^s = \hat{\delta}_f^{s'} + \hat{\delta}_c$ and $\hat{\delta}_f^b = \hat{\delta}_f^{b'} + \hat{\delta}_c$ where $\hat{\delta}_f^{s'}$ and $\hat{\delta}_f^{b'}$ are the fixed-effects estimated in (1). By adding the mean discrepancy ($\hat{\delta}_c$) to the deviations from the mean, $\hat{\delta}_f^s$ and $\hat{\delta}_f^b$ give us each firm's reporting discrepancies as a seller (respectively, a buyer) controlling for trade partners' effect and time variations. We replace missing buyer- or seller-FE estimates with zero. In Table B.3, we show that the classification is very similar when we drop firms with missing FEs from our analysis.

geous misreporters, only 10% are looking-small advantageous misreporters, and the remaining 12% are looking-big advantageous misreporters. The high proportion of conspicuous advantageous misreporters suggests that the majority of Ugandan firms believe that the tax authority is unlikely to detect evasion by monitoring firms' reported value added.

We also find that 4,779 firms (25%) misreport in a *disadvantageous* way. A substantial share of firms thus make systematic reporting errors. Such errors can take many different concrete forms, but are asymmetric in nature: *on net*, disadvantageous misreporting behavior raises a firm's tax liability. Our terminology thus labels a firm as "confused" if the systematic component of its (mis)reporting behavior increases the firm's tax liability, and vice versa for "strategic".[15]

Advantageous and disadvantageous misreporting occurs with comparable frequency among smaller, medium-sized, and somewhat larger VAT-registered firms, as shown in Figure E.1. However, the figure also shows that the average Q_f measure markedly increases among the largest firms, suggesting that they are more sophisticated tax (mis)reporters than other firms. A more detailed comparison of the two types of firms is in Appendix Table B.1.

4.3 Interpretation and robustness

We conjecture that the methodology we develop sheds new light on firms' decision-making. A first concern to consider is the potential influence of sampling error on the fixed-effect estimates used to construct Q_f (Lancaster, 2000). Fortunately, our sample is large in the relevant dimensions. Within our 3,373,183 observations, sellers appear 240 times and sell to 37 buyers on average; buyers appear 184 times and buy from 28 sellers on average; and seller-buyer pairs appear 21 times on average. This "connectedness" distinguishes the network we study from those in traditional applications of the Abowd *et al.* (1999, 2002) methodology to employer-employee data (see also Fontaine *et al.*, 2020).

Additionally, each additional firm yields more observations in both of the two fixed-effects dimensions in our setting, since each firm is itself both a seller and buyer. Therefore the estimated fixed-effects are arguably asymptotic both in N and T, instead of only in T,

[15]Our methodology cannot detect misreporting of individual firm-pair×month transaction values, and "nets out" any *symmetric* misreporting across a firm's various transaction partners. The advantageous and disadvantageous misreporting we capture is thus systematic. Given that negative liabilities can be carried over to later months, one example of the latter is not bothering to include all input purchases in the firm's tax declaration when its liability is in any case negative. We find, in fact, that firms classified as disadvantageous misreporters—especially those with a negative buyer fixed-effect—are 20% less likely to file a VAT return with a negative liability, but 18% more likely to file a null return (Table G.3). This is just one example of (systematic) disadvantageous misreporting behavior.

as is usually the case.[16] Cluster-bootstrapping to estimate standard errors on $\hat{\delta}_f^s$ and $\hat{\delta}_f^b$ (see Appendix B), we thus report the classification that results if we restrict attention to the 42% of firms for which we can reject $Q_f = 0$ at conventional significance levels.

For this subsample we find very similar proportions of advantageous (77%) and disadvantageous (23%) firms as in the full sample. We show this in Panel B of Table 1. Also as in the full sample, the majority of advantageous misreporters are "conspicuous" ones (83%), with a smaller share of "looking-small" (8%) and "looking-big" (9%) advantageous misreporters.

We next re-estimate (1) and classify firms via (2) separately for each year in our sample. We find that 74% (resp., 65%) of firms classified as advantageous (disadvantageous) misreporters in year t stay within that classification also in the subsequent year, as shown in Table B.2. Both these results and those in Panel B of Table 1 suggest that the fixed-effects model captures persistent forms of firm behavior (see also the simulation in Appendix G). However, disadvantageous behavior appears to be somewhat less persistent over time than advantageous behavior.

A second concern to consider is whether buyer-seller matching could bias our estimates of δ_f^b and δ_f^s. To investigate, we depict events in which a firm switches trade partners (see also Card *et al.*, 2013). We classify a firm's "old" and "new" trade partner into quartiles using the average discrepancy they each incur in their trade with *other* firms during periods around such a switch. As seen in Figures D.1 and D.2, the firm's reporting discrepancies do not appear to be trending up or down, nor dip or spike, before a switch in trade partner (type). However, its discrepancies change abruptly—and in the direction the change in trading-partner type predicts—when the switch happens. Finally, the discrepancy changes associated with switching trade partners appear symmetric: firms switching from a partner in the top quartile of average discrepancies to a partner in the bottom quartile experience a reduction of similar (absolute) magnitude to those switching in the opposite direction. These observations indicate that sellers and buyers do not sort into trade relationships based on unmodelled match effects in unilateral VAT (mis)reporting.

A third concern to consider is that we do not observe misreporting of sales to final consumers. If firms we classify as disadvantageous misreporters—those that overreport their firm-to-firm sales or underreport their inputs—also underreport a large enough share of sales to final consumers, their total misreporting may in principle be advantageous. To investigate, we re-estimate our model assuming that all firms underreport a given proportion of their sales to final consumers. As seen in Table 2, the proportion of advantageous

[16] This also distinguishes our setting from employer-employee data, where the two Abowd *et al.* (1999, 2002) fixed-effects dimensions are units of different nature.

firms increases to 77% when we assume that all firms underreport final sales by 10%. Even assuming an implausibly high degree of misreporting of sales to final consumers—50%—the share of disadvantageous firms remains high at about 19%.[17]

We conclude that the results in Table 1—a majority of strategic misreporters, but a notable minority of persistently confused firms—likely reflect true variation in firm type and unilateral VAT misreporting in Uganda, underscoring the importance of accounting for heterogeneity in firm sophistication in theory and policy design.

4.4 Revenue consequences

The results in Subsection 3.2 suggest that there may be significant positive revenue consequences for the Ugandan government of disadvantageous VAT misreporting, but also that, in aggregate, VAT misreporting likely decreases government revenues substantially. However, revenue consequences of VAT misreporting are not a simple sum of seller and buyer shortfalls: an increased (or decreased) liability attributed to one firm may have different revenue consequences from one attributed to the firm's trade partner because of rules for refunding negative VAT liabilities (see Appendix A and Almunia *et al.* (2017)).

We divide up each reporting discrepancy $d_{ff't}$ between the two firms using the seller and buyer fixed-effects estimated in Subsection 4.1. If the two fixed-effects have the same sign, we assign shares of the discrepancy in proportion to these. If they have opposite signs, we assign the entire discrepancy to the firm whose fixed effect matches the sign of the discrepancy. Details are in Appendix C.

Our estimates imply that the Ugandan government would have lost US$137 million in tax revenues during 2013-2016 if (only) disadvantageous misreporting were eliminated, as seen in the bottom rows of Table 3. If (only) advantageous misreporting were eliminated, our estimates imply a revenue *gain* of about US$522 million (assuming that liabilities can be collected). If both forms of misreporting were eliminated, our estimates imply a revenue gain of US$384 million, or about 28% of the total VAT collected.[18] These estimates are very similar if we use an alternative way to apportion discrepancies based on the estimated fixed-effects, and also if we naively assume that all instances of seller shortfall

[17] Assuming that the entire VAT compliance gap estimated for Uganda is due to evasion on sales to final consumers—which this paper shows is far from the case—would imply that firms misreport sales to final consumers by 50% (IMF, 2014).

[18] Many Ugandan firms have positive outstanding balances with the URA. This helps explain why the revenue consequences of eliminating disadvantageous misreporting are proportionally smaller (in absolute value) than those of eliminating advantageous misreporting. This, in combination with the correlation between individual firms' buyer and seller shortfalls (see Subsection 4.2), also helps explain why the revenue gain from eliminating all VAT misreporting is smaller than the sum of the gain from eliminating respectively disadvantageous and advantageous misreporting.

are entirely due to sellers and all instances of buyer shortfall due to buyers, as shown in Appendix C.

5 Enhanced Enforcement Capacity and VAT Evasion by Strategic and Confused Firms

We now show evidence that firms misreport less when the state's tax enforcement capacity is greater, but that less sophisticated firms adjust their behavior to a lesser extent. We leverage the fact that imports are subject to greater oversight than domestic transactions.

When Ugandan firms file for Customs clearance of an import transaction, they are required to pay the VAT on the imported goods plus tariffs. To later obtain the corresponding tax credit, they declare the input VAT paid on imports on their VAT "schedules". We thus compare, in firm-month observations, *a given firm's* double reports of the same transaction.[19]

The same amount is reported at Customs and in the firm's VAT declaration in 53% of observations. In 14% of cases, the firm claims a larger amount in VAT credit than what it reported at Customs, thus reducing its VAT liability. This self-advantageous misreporting is less frequent than occurrences of seller shortfall in domestic transactions, in line with the intuition that many firms adjust their behavior to the state's enforcement capacity.

In the remaining 34% of observations, firms report a lower amount in their VAT declaration than at Customs, thus leaving input tax credit unclaimed. This behavior, which we label *seemingly anomalous*, is analogous to buyer shortfall discrepancies in domestic VAT transactions, with the difference that here, the same firm makes both tax declarations.

Seemingly anomalous underclaiming of input tax credit from imported goods may reflect disadvantageous behavior. This appears to be part of the explanation. First, monthly VAT returns reporting a null tax liability are 22 percentage points more likely to display seemingly anomalous import reporting than returns with a positive VAT liability, perhaps because some firms with a null VAT liability do not bother claiming input VAT credits from imports (see Table 4 and footnote 15). Second, seemingly anomalous reporting is less frequent in the early and final months of each fiscal year, when tax matters may be more salient to taxpayers (see Table G.4). However, seemingly anomalous reporting may also represent strategic behavior. There is for example anecdotal evidence that some goods are imported into Uganda by businesses even though they are destined for consumption by

[19] We do so for the 9,318 firms that import and for which we estimate seller and buyer fixed effects in Section 4.

individuals. Because these are not actual business inputs, they do not generate input VAT credits and are legitimately not reported as such.

To investigate, we compare transaction amounts reported at Customs and in domestic VAT declarations separately for firms classified as advantageous and disadvantageous misreporters based on *domestic* VAT transactions in Section 4. In Table 4 the outcome variable is a dummy variable that is equal to one for monthly observations with seemingly anomalous reporting as defined above.[20]

We find that disadvantageous misreporters and firms with a negative buyer fixed-effect are respectively 4.4 and 8.3 percentage points (13% and 24%) more likely to engage in seemingly anomalous reporting of imports than other firms.[21] These estimates, shown in columns 1 and 3 of Table 4, point towards financially irrational behavior by (some) firms and help validate the classification procedure in Section 4.

In contrast, we find no statistically significant difference between advantageous and disadvantageous misreporters' propensity to engage in self-advantageous misreporting of imports (see Table G.5). Both types of firms appear to adjust their behavior to the verifiable nature of imported inputs and engage in less self-advantageous misreporting of imports than of domestic transactions.

Overall, the results in this section indicate that strategic firms misreport less when the state's tax enforcement capacity is greater, while confused firms do so to a lesser extent.

6 Conclusion

In this paper we analyze the extent to which firms make decisions that benefit themselves. The context is tax reporting in a low-enforcement setting: Uganda. We document widespread discrepancies between seller and buyer VAT reports, with dramatic consequences for tax revenue collected. By comparing a given firm's misreporting of sales and purchases over time, we show that, while a majority of firms misreport in a way that reduces their tax liability, a non-negligible fraction—about a quarter—consistently misreports such that their tax liability increases.

In the second part of the paper, we show that firms classified as strategic and confused—advantageous and disadvantageous misreporters—appear to respond differently to the state's tax enforcement capacity. All firms misreport less at customs where goods are subject to greater monitoring, but confused firms are more likely to underreport their input

[20] We allow for rounding errors and pure timing mismatches, as in Section 4. We also control for firm size (deciles of reported annual turnover) and sector in all specifications.

[21] These estimates remain of the same order of magnitude when we control for null VAT reported or include dummies for the type of goods being imported, as seen in columns 2 and 4.

tax credit for imported goods on their VAT returns.[22]

These findings suggest that (i) the proportion of firms that do not engage in sophisticated optimization as usually assumed is high—with important implications for theory and policy—but (ii) the majority of firms nevertheless respond to low state capacity by evading taxes. Together, these two conclusions call into question the self-enforcement properties of the VAT in limited enforcement contexts.

[22]This analysis alone does not imply that the overarching reason for widespread misreporting is low enforcement capacity—there could be additional important contributors. To investigate, future research could leverage local tax enforcement shocks.

References

ABOWD, JOHN M., KRAMARZ, FRANCIS, & MARGOLIS, DAVID N. 1999. High Wage Workers and High Wage Firms. *Econometrica*, **67**(2), 251–333.

ABOWD, JOHN M, CREECY, R, & KRAMARZ, FRANCIS. 2002. *Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data*. Census Bureau Technical paper No. TP 2002-06.

ACEMOGLU, DARON, GARCIA-JIMENO, CAMILO, & ROBINSON, JAMES A. 2015. State Capacity and Economic Development: A Network Approach. *American Economic Review*, **105**(8), 2364–2409.

AGHION, PHILIPPE, AKCIGIT, UFUK, , LEQUIEN, MATHIEU, & STANTCHEVA, STEFANIE. 2017. *Tax Simplicity and Heterogeneous Learning*. NBER Working Paper 24049.

ALMUNIA, MIGUEL, & LOPEZ-RODRIGUEZ, DAVID. 2018. Under the Radar: The Effects of Monitoring Firms on Tax Compliance. *American Economic Journal: Economic Policy*, **10**(1), 1–38.

ALMUNIA, MIGUEL, GERARD, FRANCOIS, HJORT, JONAS, KNEBELMANN, JUSTINE, NAKYAMBADDE, DOROTHY, RAISARO, CLAUDE, & TIAN, LIN. 2017. *An Analysis of Discrepancies in Tax Declarations submitted Under Value-Added Tax in Uganda*. International Growth Centre Project Report.

AMODIO, FRANCESCO, CHOI, JIEUN, DE GIORGI, GIACOMO, & RAHMAN, AMINUR. 2021. *Bribes vs. Taxes: Market Structure and Incentives*. working paper, McGill University.

BAUNSGAARD, THOMAS, & KEEN, MICHAEL. 2010. Tax revenue and (or?) trade liberalization. *Journal of Public Economics*, **94**(9-10), 563–577.

BENZARTI, YOUSSEF. 2020. How Taxing Is Tax Filing? Using Revealed Preferences to Estimate Compliance Costs. *American Economic Journal: Economic Policy*, **12**(4), 38–57.

BESLEY, T., & PERSSON, T. 2010. State capacity, conflict and development. *Econometrica*, **78**, 1–34.

BESLEY, TIMOTHY, & PERSSON, TORSTEN. 2009. The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, **99**, 1218–44.

BEST, MICHAEL, BROCKMEYER, ANNE, KLEVEN, HENRIK, SPINNEWIJN, JOHANNES, & WASEEM, MAZHAR. 2015. Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan. *Journal of Political Economy*, **123**(6), 1311–1355.

BEST, MICHAEL CARLOS, HJORT, JONAS, & SZAKONYI, DAVID. 2019. *Individuals and Organizations as Sources of State Effectiveness*. NBER Working Paper 23350.

BIRD, RICHARD, & GENDRON, PIERRE-PASCAL. 2007. *The VAT in Developing and Transitional Countries*. Cambridge and New York: Cambridge University Press.

CAGÉ, JULIA, & GADENNE, LUCIE. 2018. Tax Revenues, Development, and the Fiscal Cost of Trade Liberalization, 1792-2006. *Explorations in Economic History*, **70**(C), 1–24.

CARD, DAVID, HEINING, JORG, & KLINE, PATRICK. 2013. Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, **128**(3), 967–1015.

CARRILLO, PAUL, POMERANZ, DINA, & SINGHAL, MONICA. 2017. Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics*, **9**(2), 144–164.

CHETTY, RAJ, KROFT, KORY, & LOONEY, ADAM. 2009. Salience and Taxation: Theory and Evidence. *American Economic Review*, **99**(4), 1145–1177.

CHO, SUNGJIN, & RUST, JOHN. 2010. The flat rental puzzle. *Review of Economic Studies*, **77**(2), 560–594.

DELLAVIGNA, STEFANO, & GENTZKOW, MATTHEW. 2019. Uniform Pricing in US Retail Chains. *Quarterly Journal of Economics*, **134**(4), 2011–2084.

DUBE, ARINDRAJIT, MANNING, ALAN, & NAIDU, SURESH. 2020. *Monopsony and Employer Mis-Optimization Explain Why Wages Bunch at Round Numbers*. NBER Working Paper 24991.

DUFLO, ESTHER, GREENSTONE, MICHAEL, PANDE, ROHINI, & RYAN, NICHOLAS. 2018. The Value of Discretion in the Enforcement of Regulation: Experimental Evidence and Structural Estimates from Environmental Inspections in India. *Econometrica*, **86**(6), 2123–2160.

EBRILL, LIAM, BODIN, JEAN-PAUL, KEEN, MICHAEL, & SUMMERS, VICTORIA. 2001. *The Modern VAT*. Washington: International Monetary Fund.

EMRAN, M. SHAHE, & STIGLITZ, JOSEPH E. 2005. On Selective Indirect Tax Reform in Developing Countries. *Journal of Public Economics*, **89**(4), 599–623.

FAN, HAICHAO, QIAN, NANCY, WEN, JAYA, & YU, LIU. 2019. *The Dynamic Effects of Computerized VAT Invoices on Chinese Manufacturing Firms*. NBER Working Paper 24414.

FISMAN, RAYMOND, & WEI, SHANG-JIN. 2004. Tax Rates and Tax Evasion: Evidence from "Missing Imports" in China. *Journal of Political Economy*, **112**(2), 471–496.

FONTAINE, FRANÇOIS, MARTIN, JULIEN, & MEJEAN, ISABELLE. 2020. Price discrimination within and across EMU markets: Evidence from French exporters. *Journal of International Economics*, **124**, 103300.

GILLITZER, CHRISTIAN, & SKOV, PEER EBBESEN. 2018. The use of third-party information reporting for tax deductions: evidence and implications from charitable deductions in Denmark. *Oxford Economic Papers*, **70**(3), 892–916.

GOLDFARB, AVI, & XIAO, MO. 2011. Who thinks about the competition? Managerial ability and strategic entry in US local telephone markets. *American Economic Review*, **101**(7), 3130–3161.

HANSMAN, CHRISTOPHER, HJORT, JONAS, & LEÓN, GIANMARCO. 2019. Interlinked Firms and the Consequences of Piecemeal Regulation. *Journal of the European Economic Association*, **17**(3), 876–916.

HJORT, JONAS, LI, XUAN, & SARSONS, HEATHER. 2020. *Across-Country Wage Compression in Multinationals*. NBER working paper no. 26788.

HORTACSU, ALI, & PULLER, STEVEN L. 2008. Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market. *The RAND Journal of Economics*,, **39**(1), 86–114.

IMF. 2013. *Uganda - Staff Report for 2013 Article IV Consultation*. International Monetary

Fund Country Report, volume 13/215.

IMF. 2014. *Revenue Administration Gap Analysis Program-The Value-Added Tax Gap Uganda*. International Monetary Fund Fiscal Affairs Department.

KEEN, MICHAEL, & LIGHART, JENNY. 2002. Coordinating Tariff Reduction and Domestic Tax Reform. *Journal of International Economics*, **56**(2), 489–507.

KEEN, MICHAEL, & LIGHART, JENNY. 2005. Coordinating Tariff Reduction and Domestic Tax Reform under Imperfect Competition. *Review of International Economics*, **13**(2), 385–390.

KLEVEN, HENRIK. 2014. How Can Scandinavians Tax So Much? *Journal of Economic Perspectives*, **28**(4), 77–98.

KLEVEN, HENRIK, KNUDSEN, MARTIN, KREINER, CLAUS, PEDERSEN, SOREN, & SAEZ, EMMANUEL. 2011. Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, **79**(3), 651–692.

KLEVEN, HENRIK, KREINER, CLAUS, & SAEZ, EMMANUEL. 2016. Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries. *Economica*, **83**, 219–246.

KOPCZUK, WOJCIECH, & SLEMROD, JOEL. 2006. Putting Firms into Optimal Tax Theory. *American Economic Review, Papers and Proceedings*, **96**, 130–134.

KREMER, MICHAEL, RAO, GAUTAM, & SCHILBACH, FRANK. 2019. Behavioral Development Economics. *In:* BERNHEIM, DOUGLAS, DELLAVIGNA, STEFANO, & LAIBSON, DAVID (eds), *Handbook of Behavioral Economics*, vol. 2. Elsevier.

LAFFONT, JEAN-JAQUES. 2005. *Regulation and Development*. Cambridge University Press.

LANCASTER, TONY. 2000. The Incidental Parameter Problem since 1948. *Journal of Econometrics*, **95**, 391–413.

LEMGRUBER, ANDREA, MASTERS, ANDREW, & CLEARY, DUNCAN. 2015. *Understanding Revenue Administration: An Initial Data Analysis Using the Revenue Administration Fiscal Information Tool*. International Monetary Fund Fiscal Affairs Department Paper Series.

MASCAGNI, GIULIA, MUKAMA, DENIS, & SANTORO, FABRIZIO. 2019. An Analysis of Discrepancies in Taxpayers? VAT Declarations in Rwanda. *International Centre for Taxation and Development Working Paper 92*.

MITTAL, SHEKHAR, & MAHAJAN, APRAJIT. 2017. *VAT in Emerging Economies: Does Third Party Verification Matter*. Mimeo UCLA.

NARITOMI, JOANA. 2019. Consumers as Tax Auditors. *American Economic Review*, **109**(9), 3031–3072.

OECD. 2018. *Revenue Statistics in Africa 2018*. OECD Centre for Tax Policy and Administration, African Union Commission (AUC) and African Tax Administration Forum (ATAF).

PAGE, LUCY, & PANDE, ROHINI. 2018. Ending Global Poverty: Why Money Isn't Enough. *Journal of Economic Perspectives*, **32**, 173–200.

POMERANZ, DINA. 2015. No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. *American Economic Review*, **105**(8), 2539–69.

RECK, DANIEL. 2016. *Taxes and Mistakes: What?s in a Sufficient Statistic?* Mimeo LSE.

REES-JONES, ALEX, & TAUBINSKY, DMITRY. 2018. Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment. *Review of Economic Studies*, **85**(4), 317–330.

RIEZMAN, RAYMOND, & SLEMROD, JOEL. 1987. Tariffs and collection costs. *Review of World Economics*, **123**(3), 545–549.

SLEMROD, JOEL, BLUMENTHAL, MARSHA, & CHRISTIAN, CHARLES. 2001. Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota. *Journal of Public Economics*, **79**, 455–483.

SLEMROD, JOEL, COLLINS, BRETT, HOOPES, JEFFREY, RECK, DANIEL, & SEBASTIANI, MICHAEL. 2017. Does Credit-card Information Reporting Improve Small-business Tax Compliance? *Journal of Public Economics*, **149**, 1–19.

TOUREK, GABRIEL. 2021. *Targeting in Tax Behavior: Evidence from Rwandan Firms*. working paper, Massachusetts Institute of Technology.

WASEEM, MAZHAR. 2018. Taxes, informality and income shifting: Evidence from a recent Pakistani tax reform. *Journal of Public Economics*, **157**, 41 – 77.

WASEEM, MAZHAR. 2020. The Role of Withholding in the Self-Enforcement of a Value-Added Tax: Evidence from Pakistan. *Review of Economics and Statistics, forthcoming*.

Figures

FIGURE 1
DISCREPANCIES IN THE DOMESTIC VAT DATA

(A) VAT AMOUNTS DECLARED BY SELLERS VS BUYERS

(B) DISTRIBUTION OF REPORTING DISCREPANCIES

Notes: Data source: VAT Schedules for fiscal years 2013-2016. Panel (A) plots the inverse hyperbolic sine (ihs) transformation of amounts reported by sellers over that by buyers for all monthly transaction data in fiscal years 2013-2016. The data are grouped into a 0.05×0.05 grid and the color represents the number of observations in each square, going from 1 (lightest gray) to more than 50,000 (black). Squares on the 45-degree line correspond to observations where seller and buyer-reported amounts match. Observations above that line correspond to cases of buyer shortfall, while those below indicate cases of seller shortfall. The dashed line represents the conditional mean of ihs(Amount reported by sellers) for the values of ihs(Amount reported by buyers). In Panel (B), we show the distribution of discrepancies in the reporting of transactions by sellers and buyers for fiscal years 2013-2016, calculated by taking the difference between VAT charged in VS1 and VAT paid in VS24. We use the inverse hyperbolic sine transformation of VS1 and VS24. Share ≥ 1: 0.028; Share ≤ -1: 0.031.

Tables

TABLE 1
FIRM TYPE CLASSIFICATION BASED ON Q STATISTIC

	Panel A: All firms	
	No. of firms	Share of Firms
Advantageous	14,358	0.75
Conspicuous	11,248	0.59
Looking small	1,404	0.07
Looking big	1,706	0.09
Disadvantageous	4,779	0.25
Ratio of Adv. to Disadv.		3.00
N	19,137	
	Panel B: Significant Q's	
	No. of firms	Share of Firms
Advantageous	6,150	0.77
Conspicuous	5,111	0.64
Looking small	474	0.06
Looking big	565	0.07
Disadvantageous	1,862	0.23
Ratio of Adv. to Disadv.		3.30
N	8,012	

Notes: Data Source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. Firm types are defined based on Q_f, calculated as the weighted sum of the estimated firm-as-buyer fixed-effect and firm-as-seller fixed-effect, i.e., : $Q_f = w_s \cdot \hat{\delta}_f^s + w_b \cdot \hat{\delta}_f^b$. w_s (respectively, w_b) is the number of firm-trade partner monthly observations as a seller (resp., as a buyer), and $\hat{\delta}_f^s = \hat{\delta}_f^{s'} + \hat{\delta}_c$ and $\hat{\delta}_f^b = \hat{\delta}_f^{b'} + \hat{\delta}_c$ where $\hat{\delta}_f^{s'}$ and $\hat{\delta}_f^{b'}$ are the fixed-effects and $\hat{\delta}_c$ is the constant estimated in equation (1). Firm classifications are defined as: (1) **Advantageous**: $Q_f > 0$. Advantageous firms are further categorized into: (1a) Conspicuous Advantageous: $w_s \cdot \hat{\delta}_f^s \geq 0$ and $w_b \cdot \hat{\delta}_f^b \geq 0$; (1b) Looking small Advantageous: $w_s \cdot \hat{\delta}_f^s \geq 0$ and $w_b \cdot \hat{\delta}_f^b < 0$; and (1c) Looking big Advantageous: $w_s \cdot \hat{\delta}_f^s < 0$ and $w_b \cdot \hat{\delta}_f^b \geq 0$. (2) **Disadvantageous**: $Q_f < 0$. *In Panel B*, the sample is restricted to firms for which the confidence interval around Q_f excludes 0. To compute the variance of Q_f, we use a pairs cluster bootstrap approach, details are in Appendix B.4.

TABLE 2
FIRM TYPES ASSUMING UNDERREPORTING OF SALES TO FINAL CONSUMERS

	Panel A 10% of sales to FC		*Panel B* 30% of sales to FC		*Panel C* 50% of sales to FC	
	No. of Firms	Share of firms	No. of Firms	Share of firms	No. of Firms	Share of firms
Disadvantageous	4,187	0.22	3,649	0.19	3,324	0.17
Advantageous	14,950	0.78	15,488	0.81	15,813	0.83
Conspicuous	12,080	0.63	12,607	0.66	12,934	0.68
Looking small	1,742	0.09	2,086	0.11	2,320	0.12
Looking big	1,128	0.06	795	0.04	559	0.03

Notes: Data source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. This table presents robustness of firm-type classification, assuming various percentages of sales to final consumers are subject to seller shortfall. Firm types are defined based on Q_f, which is calculated as the weighted sum of the estimated firm-as-buyer fixed-effect and firm-as-seller fixed-effect, i.e., : $Q_f = w_s \cdot (\hat{\delta}_f^s + FC) + w_b \cdot \hat{\delta}_f^b$. w_s (respectively, w_b) is the number of firm-trade partner monthly observations as a seller (resp., as a buyer), and $\hat{\delta}_f^s = \hat{\delta}_f^{s'} + \hat{\delta}_c$ and $\hat{\delta}_f^b = \hat{\delta}_f^{b'} + \hat{\delta}_c$ where $\hat{\delta}_f^{s'}$ and $\hat{\delta}_f^{b'}$ are the fixed-effects and $\hat{\delta}_c$ is the constant estimated in equation (1). FC indicates average monthly unreported sales to final consumers: in Panel A, we consider that sellers do not report 10% of their sales to final consumers, in Panel B, 30%, in Panel C, 50%. Firm classifications are defined as: (1) **Advantageous**: $Q_f > 0$. Advantageous firms are further categorized into: (1a) Conspicuous Advantageous: $w_s \cdot (\hat{\delta}_f^s + FC) \geq 0$ and $w_b \cdot \hat{\delta}_f^b \geq 0$; (1b) Looking small Advantageous: $w_s \cdot (\hat{\delta}_f^s + FC) \geq 0$ and $w_b \cdot \hat{\delta}_f^b < 0$; and (1c) Looking big Advantageous: $w_s \cdot (\hat{\delta}_f^s + FC) < 0$ and $w_b \cdot \hat{\delta}_f^b \geq 0$. (2) **Disadvantageous**: $Q_f < 0$.

TABLE 3
REVENUE CONSEQUENCES BY FIRM TYPE

	All	(1) Disadv.	(2) Adv.	(2a) Conspic.	(2b) Looking Small	(2c) Looking Big
No. of distinct firms	19,137	4,779	14,358	11,248	1,404	1,706
Percentage of all firms	(100%)	(25%)	(75%)	(59%)	(7%)	(9%)
Total net VAT due	1,553,971	672,052	881,919	562,235	107,358	212,326
Seller shortfall						
Number of distinct firms with seller shortfall	17,249	3,999	13,250	10,178	1,391	1,681
Total net VAT due from firms with seller shortfall	1,275,917	575,655	700,262	438,417	89,462	172,382
Total VAT subject to seller shortfall	899,736	101,959	797,776	351,397	396,986	49,393
Buyer shortfall						
Number of distinct firms with buyer shortfall	17,979	4,490	13,489	10,416	1,381	1,692
Total net VAT due from firms with buyer shortfall	1,316,813	614,770	702,043	439,842	89,107	173,094
Total VAT subject to buyer shortfall	727,354	419,675	307,679	147,921	51,920	107,838
Correcting seller shortfall and buyer shortfall						
Impact on total net VAT due	384,154	−138,442	522,597	207,688	326,193	−11,285
Percentage of total VAT collected	28.2%	−10.2%	38.4%	15.2%	24.0%	−0.8%

Notes: Data Source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. Revenue consequences are calculated by correcting the VAT liability in the last month of the year for the total VAT under seller shortfall and under buyer shortfall. Shortfall is assigned using firms' estimated fixed-effects, see Appendix C for details. The first column shows results for the whole sample, while Columns (1) to (2c), firms are divided into sub-types based on their Q_f statistic. All values are in thousands of USD.

TABLE 4
SEEMINGLY ANOMALOUS REPORTING AT CUSTOMS AND FIRM TYPE

	Dep.Var.: seemingly anomalous reporting			
Firm Type	(1)	(2)	(3)	(4)
Disadvantageous	0.049***	0.043***		
	(0.010)	(0.009)		
Null VAT		0.220***		0.211***
		(0.012)		(0.012)
Negative Buyer FE			0.091***	0.078***
			(0.010)	(0.009)
Negative Seller FE			-0.009	-0.001
			(0.009)	(0.008)
Month-Year FE	Yes	Yes	Yes	Yes
Size and Sector FE	Yes	Yes	Yes	Yes
HS Share of Import	No	Yes	No	Yes
Observations	123303	123303	123303	123303
R-squared	0.03	0.07	0.03	0.07
Mean of dep.	0.34	0.34	0.34	0.34

Notes: Data source: VAT Schedule 3, MVR and Customs data for fiscal years 2013-2016. This regression analyzes whether disadvantageous firms, and firms which have a negative seller (buyer) fixed-effect are more likely to behave in a seemingly anomalous way at Customs. Observations are at the firm-month level. The dependent variable is a dummy equal to one if the VAT amounts on imports claimed in VS3 are lower than the VAT paid on imports recorded in the Customs data in the same month. We allow for 1,000 UGX rounding and for pure timing mismatches. In Columns (1) and (2), the explanatory variable of interest is a (time invariant) dummy for firm type, equal to one if the firm is classified as Disadvantageous, based on the value of Q_f, as explained in Section 4.2. In Columns (3) and (4), the explanatory variables of interest are dummies equal to one if the buyer (resp. seller) fixed-effect estimated for the firm as described in Section 4.2 is negative. In all specifications, we control for firm size as measure by annual decile of reported turnover, and for firm sector. In Columns (2) and (4), we additionally control for a dummy indicating null monthly VAT liability reported and for the type of goods imported as measured by dummies for each of the 21 HS Good Code Sections (equal to one if the firm imports at least one good from the corresponding section). Standard errors, clustered at the firm level, are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.