# Online Appendix for:

# Strategic or Confused Firms?
# Evidence from "Missing" Transactions in Uganda

Miguel Almunia, Jonas Hjort,
Justine Knebelmann and Lin Tian

July 2021

# A Background on the VAT in Uganda

The Ugandan VAT—introduced in 1996—follows a relatively standard design. A general rate of 18% applies to all sales, with the usual exemptions for necessities and some services.[23] Firms with an annual turnover above 50 million Ugandan Shillings (US$13,700)—a threshold raised to 150 million Ugandan Shillings (US$41,100) in fiscal year 2015-16—are required to be registered for the VAT, while smaller firms can choose to pay a simplified turnover tax.[24] As in other countries, exports are zero-rated, but the VAT applies to imports. The VAT on imports is directly paid at customs, and can be credited as input in the VAT declarations.[25] VAT firms are required to submit monthly VAT declarations to the Uganda Revenue Authority (URA). Payments of positive tax liabilities are due within 30 days of the declaration. Refunds in the case of negative VAT liabilities are restricted. Negative liabilities of less than 5 million Ugandan Shillings (US$1,370) can only be carried over as offset against future VAT liabilities (indefinitely). If the stock of negative liabilities is above this threshold, firms may request a refund but this triggers a desk audit by the URA. The strict regulation of VAT refunds is common practice in low-income countries (Lemgruber *et al.* , 2015).

While the rules regarding VAT declaration and payment are similar across all VAT firms,[26] the URA categorizes firms into three groups for monitoring and enforcement purposes: large taxpayers are handled by a specific Large Taxpayer Office (LTO); medium-size taxpayers are handled by the Medium Taxpayer Office (MTO); and smaller firms are handled by the local URA offices spread out across the country.[27] For further institutional details and descriptive statistics on the VAT system Uganda, see Almunia *et al.* (2017).

[23]For instance, unprocessed agricultural products and medical, educational and financial services are exempted from VAT. Another set of goods and services are zero-rated. A firm producing zero-rated goods may claim input tax credits, while VAT paid on inputs used in the production of exempted goods cannot be recovered (Uganda Revenue Authority, 2016).

[24]This turnover tax replaces both the VAT and the CIT. Firms below the registration threshold may choose to enter the VAT system on a voluntary basis. After the threshold was increased, the majority of firms between the new and the old threshold remained in the VAT system.

[25]Total VAT revenues are divided almost equally between the contributions from the domestic VAT and the VAT on imports.

[26]With the exception that firms with an annual turnover below 200 million Ugandan Shillings (US$55,026) may apply for their VAT to be calculated using cash basis accounting.

[27]LTOs are firms with an annual turnover above 15 billion Ugandan Shillings (US$4.1 million) and/or belonging to specific sectors such as oil and mining, banking, insurance, and government departments. MTOs are firms with a turnover above 2 billion Ugandan Shillings (US$550,260, threshold increased to 5 billion Ugandan Shillings/US$1.3 million in 2015). STOs are firms with an annual turnover lower than the MTO threshold, but above 50 million Ugandan Shillings (13,700 USD, threshold increased to 150 million Ugandan Shillings/US$41,100 in 2015). Below this threshold, which is the same as the mandatory VAT registration threshold, firms are classified as Micro Taxpayers.

# B Fixed-effects analysis

In this section, we present further details for the fixed-effects analysis and results from the robustness checks .

## B.1 Comparison of advantageous and disadvantageous firms

After classifying firms into Advantageous and Disadvantgeous type as described in Section 4, we compare the observable characteristics of each firm-type. Results are shown in Table B.1. We regress a dummy variable for being an Advantageous firm, on a set of firm characteristics. To facilitate comparison, all variables are standardized and have unit standard deviation. We display results for the OLS regression (Columns 1 and 2), and for a LASSO regression (Column 3). The LASSO results show that the characteristics which are significantly different across firm types are the following: Advantageous firms are less likely to belong to the Medium or Large Taxpayers Office (MTO or LTO). This seems consistent with the idea that MTO and LTO firms are under higher scrutiny. Advantageous firms on average have a lower amount of initial offset carried forward, and display lower total output amounts. Advantageous firms are also more downstream, this seems consistent with the idea that VAT compliance is stronger higher up in the production chain. Advantageous firms are more likely to be in the manufacturing, construction, wholesale and retail sectors, and less likely to be in the mining, water/electricity, transportation/accomodation, information/communication, financial, real estate and education sectors.

## B.2 Panel estimation

Exploiting the panel dimension of the data, we investigate if firms that have self-advantageous reporting behaviors in one year tend to be the same ones that have them in the next year. This allows us to verify whether our classification is consistent over time.

We compute the transition matrix by comparing a firm's classifications for different years. That is, we run Equation (1) separately for each year in the sample:

$$d_{ff't} = \delta_c + \delta^b_{fy} + \delta^s_{f'y} + \delta_t + r_{ff't}, \tag{B.1}$$

where $y$ = fiscal year 2013 to fiscal year 2016.

Since the buyer and seller fixed-effects are only identified within a "connected" set (Abowd *et al.* , 1999), we follow Card *et al.* (2013) and restrict the analysis to the largest connected set of buyer-seller network for each year. We also restrict the sample to firms that appear at least in two consecutive years. Table B.2 shows the results as a transition matrix laying out firms' classification in year $t+1$ conditional on their year $t$ classification. We find that 74% of advantageous firms and 65% of disadvantageous firms stay within their classification in the following year.

## TABLE B.1
## COMPARISON OF ADVANTAGEOUS AND DISADVANTAGEOUS FIRMS

| Dep. Variable: Probability of Being Advantageous | | | |
|---|---|---|---|
| | *OLS* | | *LASSO* |
| | Coefficient | P-value | Coefficient |
| in Kampala | 0.02 | 0.00*** | 0.00 |
| Distance to URA office | −0.02 | 0.03** | 0.00 |
| MTO/LTO | −0.04 | 0.00*** | −0.04 |
| VAT Payable | −0.67 | 0.10 | 0.00 |
| VAT Due | −0.02 | 0.21 | 0.00 |
| Offset | −0.36 | 0.10* | 0.00 |
| Initial Offset | −0.01 | 0.01** | −0.01 |
| Total input | −1.10 | 0.10* | 0.00 |
| Total output | 1.43 | 0.10 | −0.01 |
| Ratio of sales to FC | 0.01 | 0.17 | 0.00 |
| Number of clients | 0.02 | 0.00*** | 0.01 |
| Number of suppliers | 0.00 | 0.84 | 0.00 |
| Upstreamness | −0.01 | 0.08* | −0.01 |
| Distinct outputs (all good codes) | 0.01 | 0.80 | 0.00 |
| Distinct outputs (relevant good codes) | 0.00 | 0.92 | 0.00 |
| Distinct inputs (all good codes) | 0.04 | 0.23 | 0.00 |
| Distinct inputs (relevant good codes) | −0.04 | 0.25 | 0.00 |
| Sectors: | | | |
| Agriculture, forestry, fishing | −0.01 | 0.30 | 0.00 |
| Mining, Quarrying | −0.02 | 0.00*** | −0.02 |
| Manufacturing | 0.01 | 0.17 | 0.01 |
| Water, Electricity services | −0.01 | 0.00*** | −0.01 |
| Construction | 0.01 | 0.10 | 0.01 |
| Wholesale and retail | 0.00 | 0.00 | 0.01 |
| Transportation, accomodation services | −0.03 | 0.00*** | −0.02 |
| Information, communication | −0.01 | 0.02** | −0.01 |
| Financial services | −0.03 | 0.00*** | −0.03 |
| Real estate | −0.03 | 0.00*** | −0.02 |
| Professional, Admin, Other Services | −0.01 | 0.38 | 0.00 |
| Public Administration | −0.01 | 0.57 | 0.00 |
| Education | −0.01 | 0.05** | −0.01 |
| Health and social work | 0.00 | 0.74 | 0.00 |
| Arts and Entertainment | −0.01 | 0.21 | 0.00 |

**Notes:** Data source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. This table shows the results of the regression of a firm-type dummy variable – equal to one if the firm is categorized as Advantageous and zero otherwise – on a set of firm characteristics. *Panel A* displays the results from a multivariate regression including all variables listed. *Panel B* display the results from a LASSO regression. All variables are standardized to have unit standard deviation. *in Kampala* is a dummy equal to one if the firm is in Kampala. *Distance* is calculated by assigning each firm to a sub-county and calculating the distance from the center of the sub-county to the closest URA office. *MTO/LTO* is a dummy variable equal to one if the firm is registered in the Medium or Large Taxpayers' Office (as of June 2017). *Vat Payable, Vat Due, Offset, Total inputs and Total Output* are totals over years 2013-2016. *Initial Offset* is the amount of offset carried forward in the first time period of the firm in the time window of the data. *Ratio of sales to FC*, is the ratio of total sales to final consumers over total sales. *Number of clients and Number of suppliers* are the totals over years 2013-2016. *Upstreamness* indicates the firms' distance to final consumption—larger values indicate that the firm is higher up in the production chain. It is computed by creating an input-output matrix, based on firm-to-firm good code transactions. *Distinct outputs and Distinct inputs* are the number of unique good codes within the firm's sales/purchases over the 2013-2016 period. Good codes are based on the universe of transactions from year 2014 and are obtained by applying a machine learning text algorithm to the text descriptions included in the VAT Schedules. Sector is the firm's sector as listed in the tax registry.

**TABLE B.2**
**FIRM-TYPE TRANSITION MATRIX**

| | *Firm-pairs observed throughout 2013-2016* | | |
|---|---|---|---|
| | Advantageous (t) | Disadvantageous (t) | Share |
| Advantageous (t+1) | 45.75 | 13.49 | 59.23 |
| | (73.93) | (35.38) | |
| Disadvantageous (t+1) | 16.13 | 24.63 | 40.77 |
| | (26.07) | (64.62) | |
| Share | 61.88 | 38.12 | 100.00 |
| | (100.00) | (100.00) | |

**Notes:** Data source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. This table presents the transition matrix for yearly firm classifications. The sample is restricted to firms that appear at least in two consecutive years and within each year in the largest connected set.

## B.3 Robustness

We perform four robustness checks of our firm classification by varying the sample used for the fixed-effect estimation. First, we re-calculate the estimates of $Q_f$ without replacing the missing fixed-effects by zero, as done in the main analysis. Table B.3 reports the resulting firm-type classification dropping firms for which we do not have both seller and buyer fixed-effects estimated. We show that the classification is very similar to our benchmark fixed-effects model, as reported in Table 1, with a slightly higher share of advantageous misreporters.

Second, we re-run the fixed-effect regression by including controls that affect the propensity of two firms to trade with each other. The objective is that by controlling for these, the likelihood for a seller to trade with a particular buyer is as good as randomly assigned. Specifically, we include two variables, one accounting for geographical proximity, and one accounting for sectoral complementarity. The first one is a dummy variable for whether two firms are located in the same sub-county.[28] The second one is the share of products from the seller's sector that are sold to the buyer's sector. To compute this, we use the official aggregate sector-level Input-Output tables calculated by the Ugandan Bureau of Statistics for financial year 2009. Introducing the controls decreases the sample of firms from 19,137 to 18,629. The results are shown in Panel A of Table B.4. They are similar to what we obtained when running the regression without controls: 77% of firms are classified as Advantageous (against 75% in the main analysis) and 23% are classified as Disadvantageous. Among the Advantageous firms, the respective shares of Conspicuous, Looking-small and Looking-Big are very similar to the ones in the main analysis.

Third, we replicate the analysis on a more consistent sample, as a way to potentially reduce the noise in the estimation, by keeping only firm-pairs with a number of observations larger than ten. The firm classification is displayed in Panel B of Table B.4. The share of Advantageous firms increases to 88%. Among advantageous firms, a larger share are classified as conspicuous—91%, against 78% in the main analysis. The sample is reduced to 12,565 firms.

[28] Uganda is divided up into a total of 1,403 sub-counties (Electoral Commision, 2016).

Fourth, we vary the way the raw transactions data is treated and rounded for the fixed-effects estimation. We first run the estimation on the raw transactions data, without rounding nor adjusting for timing mismatches. Results are shown in Panel A of Table B.5. The shares of Advantageous and Disadvantageous firms are the same as in our main analysis (75 and 25%). Second, we run the estimation after rounding the value of discrepancies at 5% of the transaction value, defined here as the maximum of the values reported by the seller and the buyer. As shown in Panel B of Table B.5, in this case, we find 76% of Advantageous firms and 24% of Disadvantageous firms.

**TABLE B.3**
**FIRM TYPE CLASSIFICATION BASED ON $Q$ STATISTIC (WITHOUT REPLACEMENT)**

| | *Panel A: All firms* | |
|---|---|---|
| | No. of firms | Share of Firms |
| Advantageous | 10,415 | 0.79 |
| Conspicuous | 7,305 | 0.55 |
| Looking small | 1,404 | 0.11 |
| Looking big | 1,706 | 0.13 |
| Disadvantageous | 2,812 | 0.21 |
| Ratio of Adv. to Disadv. | | 3.70 |
| N | 13,227 | |
| | *Panel B: Significant Q's* | |
| | No. of firms | Share of Firms |
| Advantageous | 4,541 | 0.82 |
| Conspicuous | 3,502 | 0.63 |
| Looking small | 474 | 0.09 |
| Looking big | 565 | 0.10 |
| Disadvantageous | 1,016 | 0.18 |
| Ratio of Adv. to Disadv. | | 4.47 |
| N | 5,557 | |

**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 $\hat{\delta}_f^s = \hat{\delta}_f^{s\prime} + \hat{\delta}_c$ and $\hat{\delta}_f^b = \hat{\delta}_f^{b\prime} + \hat{\delta}_c$ where $\hat{\delta}_f^{s\prime}$ and $\hat{\delta}_f^{b\prime}$ are the fixed-effects and $\hat{\delta}_c$ is the constant estimated in equation (1). In this version, we drop firms for which any of the two fixed-effects is missing. 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 B.4**
**FIRM TYPE CLASSIFICATION BASED ON ROBUSTNESS ESTIMATIONS (1/2)**

| | No. of Firms | Share of firms |
|---|---|---|
| | *Panel A: Two-way fixed effect estimation with controls* | |
| Advantageous | 14,318 | 0.77 |
| Conspicuous | 11,355 | 0.61 |
| Looking small | 1,262 | 0.07 |
| Looking big | 1,701 | 0.09 |
| Disadvantageous | 4,311 | 0.23 |
| N | 18,629 | 1.00 |
| | *Panel B: Sample of firm-pairs with $\geq 10$ observations* | |
| Advantageous | 11,002 | 0.88 |
| Conspicuous | 10,052 | 0.80 |
| Looking small | 361 | 0.03 |
| Looking big | 589 | 0.05 |
| Disadvantageous | 1,563 | 0.12 |
| N | 12,565 | 1.00 |

**Notes:** Data Source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. *In Panel A:* We include variables for geographical proximity and for sectoral complementarity as controls in the fixed-effects model (Section B.3 describes how the control variables are computed). *In Panel B:* We run the fixed-effects model on the subset of firm-pairs that appear ten times or more in the initial dataset.

**TABLE B.5**
**FIRM TYPE CLASSIFICATION BASED ON ROBUSTNESS ESTIMATIONS (2/2)**

| | No. of Firms | Share of firms |
|---|---|---|
| | *Panel A: Raw data* | |
| 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 |
| N | 19,137 | 1.00 |
| | *Panel B: Rounding at 5% of transaction value* | |
| Advantageous | 14,453 | 0.76 |
| Conspicuous | 11,354 | 0.59 |
| Looking small | 1,386 | 0.07 |
| Looking big | 1,713 | 0.09 |
| Disadvantageous | 4,684 | 0.24 |
| N | 19,137 | 1.00 |

**Notes:** Data Source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. *In Panel A:* We do not apply any corrections to the transactions data when defining discrepancies, nor rounding nor adjusting for timing mismatches. *In Panel B:* Discrepancies are defined by correcting for timing mismatches in the same way as in our main results, and rounding is done by setting to zero discrepancies that are inferior in absolute value to 5% of transaction size, defined as the maximum between seller declared transaction and buyer declared transaction.

## B.4 Variability analysis

To further assess the robustness of the results and the extent of variability in the firm classification that could come from sampling error, we estimate 95% confidence intervals around the point estimates for $Q_f$. There is no standard methodology in the literature to recover confidence intervals for fixed-effects in AKM-type models.[29] We develop a bootstrap routine to recover the variance of $Q_f$ for each firm $f$, defined as:

$$VAR(Q_f) = w_s^2 \cdot VAR(\hat{\delta}_f^s) + w_b^2 \cdot VAR(\hat{\delta}_f^b) + 2 \cdot w_s \cdot w_b \cdot COV(\hat{\delta}_f^s, \hat{\delta}_f^b)$$

The simple fixed-effects estimation cannot yield the covariance between $\hat{\delta}_f^s$ and $\hat{\delta}_f^b$. We use a bootstrap procedure in which we resample pairs of firms from our main sample and re-estimate the model 100 times, in the spirit of the pairs cluster bootstrap (Cameron *et al.* , 2008). When a seller-buyer pair is drawn all associated observations are drawn, and this is repeated until the sample includes as many pairs as in the main sample (519,111). The sample is then restricted to the largest connected set. The average number of observations in the bootstrap samples is 10,861,701, against 3,373,183 in the main sample. 99.9% of the pairs from our main connected set appear in at least one iteration, and the average (median) number of iterations a given pair appears in is 41 (30).

Then we estimate the standard errors of $\hat{\delta}_f^b$ and $\hat{\delta}_f^s$ by computing their standard deviation over the 100 iterations. We compute the covariance between $\hat{\delta}_f^b$ and $\hat{\delta}_f^s$ using the 100 estimations.

This allows to compute the confidence interval around $Q_f$ as:

$$CI \equiv \left[ Q_f - 1.96 \cdot \sqrt{VAR(Q_f)}\ ;\ Q_f + 1.96 \cdot \sqrt{VAR(Q_f)} \right]$$

Panel B of Table 1 displays the classification of firms when restricting to firms for which the 95% confidence interval around $Q_f$ excludes $0$. 8,012 firms (42% of firms from the main sample) are in this case. The respective shares of Advantageous and Disadvantageous are extremely similar to those in the main sample shown in Panel A, with 77% of Advantageous (against 75% in the main sample). This confirms that the share of Disadvantageous firms estimated in our model does not result from noise in the transactions data.

Taken together, the conclusions from subsections B.2, B.3 and B.4 show that our main results—the classification of firms by type of reporting behavior and the existence of a non trivial fraction of Disadvantageous reporters—are only very moderately sensitive to variations in the definition of discrepancies and in the estimation strategy, and importantly, also hold when addressing sampling error in the data.

[29] In the majority of settings (notably employer-employee data) the dimensionality of the data does not allow to rely on the standard errors of the fixed-effects. Furthermore, to our knowledge, there is no AKM-type paper that restricts the analysis to units for which the estimated fixed-effects are statistically significant. For this reason, our main results incorporate all firms from the largest connected set, and the restriction to firms with a significant $Q_f$ should be taken as a complementary result to assess the robustness of the classification.

# C Robustness Checks for the Computation of Revenue Consequences

In this section, we describe our approach to computing the revenue consequences of VAT misreporting by relying on information from the firm-type classification generated through the fixed-effects model.

## C.1 Alternative Methods to Assign Blame to Seller or Buyer

Reporting discrepancy is given by the difference in the declared amounts of monthly transactions between a buyer and a seller. We consider three alternative ways to divide the "blame" for each reporting discrepancy to the two firms involved. The first two approaches use the estimated fixed effects, whereas the third one adopts a "naive" approach.

The baseline method assigns shares of the discrepancy proportionally based on the sign of each firm's fixed effect. Formally, let $s_{it} \in [0,1]$ be the share of the discrepancy assigned to buyer 1 and seller 2, and recall $\hat{\delta}_1^b$ and $\hat{\delta}_2^s$ denote the estimated fixed effects from buyer and seller respectively. Then:

$$s_{1t} = \begin{cases} \frac{\hat{\delta}_1^b}{\hat{\delta}_1^b + \hat{\delta}_2^s} & \text{if } \hat{\delta}_1^b \cdot \hat{\delta}_2^s > 0 \\ 0.5 & \text{if } \hat{\delta}_1^b = \hat{\delta}_1^s = 0 \\ 1 & \text{if } \hat{\delta}_1^b \cdot \hat{\delta}_2^s < 0 \text{ and } \hat{\delta}_1^b \cdot d_{12t} > 0 \end{cases}$$

For example, suppose $\hat{\delta}_1^b = 30$ and $\hat{\delta}_2^s = 10$. For seller shortfall cases ($d_{12t} > 0$), we assign $s_{1t} = 0.75$ and $s_{2t} = 0.25$. In the case of buyer shortfall ($d_{12t} < 0$), we assign $s_{1t} = 0.25$ and $s_{2t} = 0.75$. If the two relevant fixed effects have the opposite signs, e.g., $\hat{\delta}_1^b = 30$ and $\hat{\delta}_2^s = -10$, we assign $s_{1t} = 1$ and $s_{2t} = 0$ in case of seller shortfall, and $s_{1t} = 0$ and $s_{2t} = 1$ in case of buyer shortfall.

The second approach, while also using the estimated fixed effects, focuses on the relative sizes of the two estimated fixed effects. For a given discrepancy $d_{ff't}$ in a given month $t$ between the two firms involved (say, a buyer $f = 1$ and a seller $f' = 2$), we first calculate the difference in the two estimated fixed effects for the two firms involved, i.e., $\hat{\delta}_1^b - \hat{\delta}_2^s$. If the absolute value of $d_{12t}$ is greater than the absolute value of the difference, we allocate the discrepancy between the firm pair such that the assigned discrepancies reflect the difference in the estimated fixed effects.[30] If the absolute value of $d_{12t}$ is less than the absolute value of the difference, we assign all the discrepancy to the more offending firm in the direction of the discrepancy. This means for a seller shortfall case, the entire discrepancy is assigned to the firm with a higher value of the fixed effects; whereas for a buyer shortfall case, the entire discrepancy is assigned to the firm with a lower value of the fixed effects. More formally, we assign the reporting discrepancies, for a given firm $f = 1$ in month $t$, according to the following equation:

[30]For example, if $d_{12t}$ is 60, $\hat{\delta}_1^b$ is 30, and $\hat{\delta}_2^s$ is 20, the assigned discrepancies for the buyer $f = 1$ and the seller $f = 2$ are 35 and 25, respectively. Note that the difference in $\hat{\delta}_1^b$ and $\hat{\delta}_2^s$ of 10 is preserved in the assignment. If $d_{12t}$ is 60, $\hat{\delta}_b$ is 30, and $\hat{\delta}_2^s$ is 30, the assigned discrepancies for the buyer $f = 1$ and the seller $f = 2$ are 30 and 20, respectively. Again, the difference in $\hat{\delta}_1^b$ and $\hat{\delta}_2^s$ of 0 is preserved in the assignment.

$$d_{1t} = \begin{cases} \frac{d_{12t}+(\hat{\delta}_1^b-\hat{\delta}_2^s)}{2}, & \text{if } \mid d_{12t} \mid > \mid \hat{\delta}_1^b - \hat{\delta}_2^s \mid . \\ d_{12t}\frac{\max(\hat{\delta}_1^b-\hat{\delta}_2^s,0)}{\hat{\delta}_1^b-\hat{\delta}_2^s}, & \text{if } \mid d_{12t} \mid \leq \mid \hat{\delta}_1^b - \hat{\delta}_2^s \mid \text{ and } d_{12t} > 0. \\ d_{12t}\frac{\min(\hat{\delta}_1^b-\hat{\delta}_2^s,0)}{\hat{\delta}_1^b-\hat{\delta}_2^s}, & \text{if } \mid d_{12t} \mid \leq \mid \hat{\delta}_1^b - \hat{\delta}_2^s \mid \text{ and } d_{12t} < 0. \end{cases} \tag{C.1}$$

Finally, in the third "naive" approach, we simply assign all seller shortfall to the seller and all buyer shortfall to the buyer.

## C.2 Revenue Consequences with Alternative Methods

Once we assign the firm-pair level misreporting to each of the firms involved, we calculate the revenue consequences of misreporting for all individual firms, defined as the change in the VAT due that would result from correcting each firms' misreporting. In doing this, we take into account the VAT offsets (outstanding tax credits) carried forward. Specifically, we correct VAT declarations for each month in the study period, making sure to update the offsets carried forward in each subsequent month. Given the restrictions on VAT refunds, the impact on the total net VAT due will not be equal to the difference between the total VAT misreported and underreported. In particular, correcting the VAT liability downward in a given month will have no direct impact on the net VAT due in that month if the original VAT liability was already negative given the eligibility threshold for a VAT refund (and the low share of eligible firms asking and getting a refund). We then aggregate the revenue implications at the yearly level, and our main results further aggregate the revenue consequences over the Fiscal Year 2013-2016 period (the fiscal year in Uganda runs from July to June). More details can be found in Almunia *et al.* (2017).

Columns 1-3 of Table C.1 report the revenue consequence calculations using the three approaches described above. The revenue loss due to misreporting remains in the same order of magnitude across the three approaches.

**TABLE C.1**
**SELLER SHORTFALL AND BUYER SHORTFALL IN THE DOMESTIC VAT ADJUSTING FOR FIRM-SPECIFIC CONTRIBUTION TO DISCREPANCIES**

| | (1) Main | (2) Alt. | (3) Naive |
|---|---|---|---|
| No. of distinct firms | 19,137 | 19,137 | 19,137 |
| Percentage of all firms | (100%) | (100%) | (100%) |
| Total net VAT due | 1,553,971 | 1,553,971 | 1,553,971 |
| Seller shortfall | | | |
| Number of distinct firms with seller shortfall | 17,249 | 17,249 | 13,448 |
| Total net VAT due from firms with seller shortfall | 1,275,917 | 1,275,917 | 1,133,456 |
| Total VAT subject to seller shortfall | 899,736 | 899,736 | 899,736 |
| Buyer shortfall | | | |
| Number of distinct firms with buyer shortfall | 17,979 | 17,979 | 17,181 |
| Total net VAT due from firms with buyer shortfall | 1,316,813 | 1,316,813 | 1,262,499 |
| Total VAT subject to buyer shortfall | 727,354 | 727,354 | 727,354 |
| Correcting seller shortfall and buyer shortfall | | | |
| Impact on total net VAT due | 384,154 | 436,152 | 492,844 |
| Percentage of total VAT collected | 28.2% | 32.0% | 36.2% |

**Notes:** Data source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. In this table we display the revenue consequence analysis using three different methods to assign discrepancies to firms (as described in Section C.1). Revenue consequences are calculated by taking the difference between VAT charged in VS1 and VAT paid in VS24, and correcting the VAT liability in the last month of the year for the total VAT under seller shortfall and under buyer shortfall, as explained in Section C.2. In column (1) (main approach), discrepancies are assigned to firms based on the sign of each firm's estimated fixed-effects, in column (2) (alternative approach) discrepancies are assigned to firms based on the relative size of each firm's estimated fixed-effects, and in column (3) (naive approach), we assign all seller shortfall to the seller, and all buyer shortfall to the buyer. All values are in thousands of USD.

# D Simulation

In order to have a better understanding of whether the fixed-effects model from Section 4 correctly captures firm behavior, we conduct a simple simulation exercise. The basic idea is to generate a set of firms that follow advantageous, disadvantageous or neutral reporting behavior, and then check whether the fixed-effects model accurately classifies them.

## D.1 Setup

We model the distribution of true transaction amounts ($Y$) for firm pairs as a chi-squared distribution: $Y \sim \chi^2_3$. We randomly generate 400,000 transactions and then allocate 10,000 firms to be the sellers or buyers in these pairs.[31] Then, we specify three types of firms, whose relative proportion will change across two different scenarios. More concretely, we model advantageous, disadvantageous, and neutral firms.

Specifically, we assume that advantageous firms operating as sellers report the true transaction amount in 50% of their transactions, and they do not report anything in the other 50%. When operating as buyers, advantageous firms report the true amount in 50% of the transactions and overreport their purchases by a factor $b_A \sim U(1,2)$ in the remaining 50%. Formally, the amounts reported by advantageous firms when operating as sellers ($Y_A^S$) and buyers ($Y_A^B$) are specified as follows:

$$Y_A^S = \begin{cases} Y & prob = 0.5 \\ 0 & prob = 0.5 \end{cases} \qquad Y_A^B = \begin{cases} Y & prob = 0.5 \\ Y * b_A & prob = 0.5 \end{cases}$$

Disadvantageous firms are modelled as a mirror image of advantageous ones. We assume that disadvantageous firms operating as sellers report the true amount in 50% of transactions, but they overreport the amount sold by a factor of $b_D \sim U(1,2)$ in the remaining 50%. When operating as buyers, disadvantageous firms report the true amount in 50% of transactions, and they do not report anything in the remaining 50%. Formally, the amounts reported by disadvantageous firms when operating as sellers ($Y_D^S$) and buyers ($Y_D^B$) are specified as follows:

$$Y_D^S = \begin{cases} Y & prob = 0.5 \\ Y * b_D & prob = 0.5 \end{cases} \qquad Y_D^B = \begin{cases} Y & prob = 0.5 \\ 0 & prob = 0.5 \end{cases}$$

The amount reported by neutral firms is set to be always equal to the true amount, such that:

$$Y_C^S = Y \qquad Y_C^B = Y$$

Finally, in order to obtain a distribution of discrepancies that more closely resembles the true data, we incorporate the possibility of symmetric reporting mistakes. We assume

[31] The ratio of firms to transactions corresponds to the average of (i) the median number of times that firms appear as sellers and (ii) the median number that firms appear as buyers in the real data.

these mistakes occur on the extensive margin, meaning that a certain proportion ($p$) of transactions is not reported either by the seller or the buyer, regardless of their firm type.

$$Y^S = \begin{cases} Y^S & prob = 1 - p \\ 0 & prob = p \end{cases} \qquad Y^B = \begin{cases} Y^B & prob = 1 - p \\ 0 & prob = p \end{cases}$$

where $Y^S$ and $Y^B$ represent the amounts reported by sellers and buyers *regardless of their type*, and $p$ is a given probability that will vary across scenarios.

## D.2 Static Simulation

We consider two different cases. In the first case, we assign one third of firms to behave as advantageous misreporters, one third to behave as neutral reporters, and one third to behave as disadvantageous reporters. In the second case, we assign 75% of the firms to behave as advantageous misreporters and the remaining 25% as disadvantageous misreporters, roughly following the proportions found in the analysis presented in Section 4 of the paper.

Table D.1 reports the distribution of outcomes at the firm pair level, which can be either seller shortfall ($Y^S < Y^B$), buyer shortfall ($Y^S > Y^B$), or neutral reporting ($Y^S = Y^B$). Column 1 reports the firm-pair level outcomes in the real data, where 48% of transactions feature seller shortfall, 32% feature buyer shortfall and 21% feature consistent reporting. Columns 2-4 report the distribution of firm-pair level outcomes in the simulated data for three different values of $p$, the share of mistakes. The number of observations declines as we increase $p$ because when there are more extensive-margin mistakes the chance that neither buyer nor seller reports anything ($Y^S = Y^B = 0$) increases, and those firm pairs are treated as unobserved.

In Panel A, where the simulated data includes equally distributed firm types, seller shortfall is less common (28%) than in the real data while consistent reporting is higher (45%). As we increase the share of mistakes in columns 3 and 4, the proportion of pairs with either seller shortfall or buyer shortfall increases, while neutral reporting declines.

In Panel B, where the simulated data includes 75% advantageous and 25% disadvantageous firms, seller shortfall is more common (57%) than in the real data, while the share with buyer shortfall is lower (17%). Consistent reporting is more common than in the real data (26%). As we increase the share of mistakes in columns 3 and 4, the share of seller shortfall remains stable at 58%, while the share of buyer shortfall converges to 29%, similar to that in the real data.

**TABLE D.1**
**SIMULATED FIRM-PAIR OUTCOMES**

| | *Panel A: 1/3 advantageous, 1/3 neutral, 1/3 disadvantageous* | | | |
|---|---|---|---|---|
| | Real data | Simulated data | | |
| | | Mistakes: $p = 0$ | Mistakes: $p = 0.2$ | Mistakes: $p = 0.4$ |
| | (1) | (2) | (3) | (4) |
| Seller shortfall | 0.48 | 0.28 | 0.34 | 0.39 |
| Buyer shortfall | 0.32 | 0.27 | 0.34 | 0.39 |
| Consistent | 0.21 | 0.45 | 0.32 | 0.21 |
| Observations | 3,370,462 | 388,687 | 355,295 | 299,620 |
| | *Panel B: 75% advantageous & 25% disadvantageous* | | | |
| | Real data | Simulated data | | |
| | | Mistakes: $p = 0$ | Mistakes: $p = 0.2$ | Mistakes: $p = 0.4$ |
| | (1) | (2) | (3) | (4) |
| Seller shortfall | 0.48 | 0.58 | 0.59 | 0.59 |
| Buyer shortfall | 0.32 | 0.16 | 0.23 | 0.29 |
| Consistent | 0.21 | 0.26 | 0.19 | 0.13 |
| Observations | 3,370,462 | 382,302 | 340,501 | 281,393 |

**Notes:** This table displays the firm-pair outcomes using simulated data under two different scenarios. *In Panel A*, we consider a scenario in which firms are equally distributed across the three types (advantageous, neutral, and disadvantageous). *In Panel B*, the distribution of firm types matches the proportions obtained in our benchmark fixed-effects models (Table 1). Column (1) shows the distribution of firm-pair outcomes observed in the real data. Column (2) shows the classification with no symmetric mistakes ($p = 0$). Column (3) and (4), respectively, display the classification with 20 ($p = 0.2$) and 40 ($p = 0.4$) percents of symmetric mistakes in reporting.

Table D.2 reports the distribution of firm types obtained by estimating the fixed-effects model described in Section 4 of the paper to the simulated data, for different shares of mistakes. Column 1 reports the fraction of advantageous-type firms that the model identifies as advantageous (and within that broad type as "conspicuous", "looking small" or "looking big") or disadvantageous, for the case in which there is no mistakes ($p = 0$). In turn, column 2 reports the fraction of disadvantageous-type firms that the model identifies with each of the types. Finally, column 3 reports the fraction of neutral firms that are classified with each of the other types. Note that, as in the main analysis presented in the paper, we do not classifiy any firm as neutral because we never obtain the point estimate $Q = 0$.

In Panel A, we find that advantageous-type and disadvantageous-type firms are correctly classified essentially in all cases (see columns 1-2, 4-5, and 7-8). In contrast, neutral firms are split evenly between advantageous and disadvantageous regardless of the amount of mistakes we introduce (see columns 3, 6, and 9). These results suggest that, if we assume firm-types to be equally distributed among the firm population, the model correctly identifies advantageous and disadvantageous behavior, while neutral firms are split between advantageous and disadvantageous with an equal probability.

In Panel B, where 75% of firms are modelled as advantageous and 25% as disadvantageous, all the advantageous-type firms are also correctly classified as advantageous, regardless of the share of mistakes (columns 1, 4, and 7). In the case of disadvantageous firms, 94% are correctly classified, while the remaining 6% are incorrectly classified as advantageous when there are no mistakes ($p = 0$, column 2). This small error rate is due to the fact that a majority of firms in this scenario are advantageous, so disadvantageous firms are disproportionately likely to interact with firms that engage in seller shortfall more frequently. Then, in a small fraction of cases, the seller shortfall discrepancies dominate the buyer shortfall discrepancies, leading to $Q < 0$. As the share of mistakes increases (columns 5 and 8), the success rate in identifying disadvantageous firms falls to 85% and 77%, respectively.

We conclude from this simulation exercise that the fixed-effects model accurately identifies advantageous firms even in the presence of symmetric mistakes, while it does a good job (though not perfect) at identifying disadvantageous firms, with an accuracy rate that declines as the share of mistakes increases. When advantageous firms are more numerous in the population than disadvantageous ones, the fixed-effects model tends to *underestimate* the proportion of disadvantageous firms in the population. This suggests that, if anything, our application of this estimation method to Ugandan firms is likely to underestimate the true share of disadvantageous firms.

**TABLE D.2**
**SIMULATED FIRM-TYPE CLASSIFICATION**

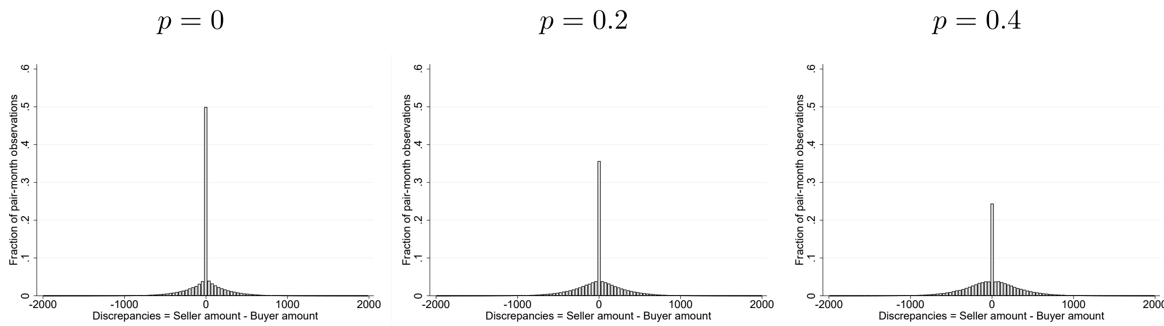
| | $p=0$ | | | $p=0.2$ | | | $p=0.4$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *Panel A: 1/3 advantageous, 1/3 neutral, 1/3 disadvantageous* | | | | | | | | | |
| | % of Adv. | % of Disadv. | % of Neutral | % of Adv. | % of Disadv. | % of Neutral | % of Adv. | % of Disadv. | % of Neutral |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Advantageous | 1.00 | 0.00 | 0.50 | 1.00 | 0.01 | 0.51 | 0.97 | 0.03 | 0.50 |
| Conspicuous | 1.00 | 0.00 | 0.15 | 0.93 | 0.00 | 0.22 | 0.82 | 0.01 | 0.24 |
| Looking small | 0.00 | 0.00 | 0.02 | 0.05 | 0.00 | 0.05 | 0.12 | 0.02 | 0.08 |
| Looking big | 0.00 | 0.00 | 0.33 | 0.02 | 0.00 | 0.24 | 0.04 | 0.00 | 0.18 |
| Disadvantageous | 0.00 | 1.00 | 0.50 | 0.00 | 0.99 | 0.49 | 0.03 | 0.97 | 0.50 |
| *Panel B: 75% advantageous & 25% disadvantageous* | | | | | | | | | |
| | $p=0$ | | | $p=0.2$ | | | $p=0.4$ | | |
| | % of Adv. | % of Disadv. | % of Neutral | % of Adv. | % of Disadv. | % of Neutral | % of Adv. | % of Disadv. | % of Neutral |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Advantageous | 1.00 | 0.06 | n.a. | 1.00 | 0.15 | n.a. | 1.00 | 0.23 | n.a. |
| Conspicuous | 1.00 | 0.02 | n.a. | 0.99 | 0.05 | n.a. | 0.96 | 0.09 | n.a. |
| Looking small | 0.00 | 0.01 | n.a. | 0.00 | 0.06 | n.a. | 0.02 | 0.09 | n.a. |
| Looking big | 0.00 | 0.03 | n.a. | 0.00 | 0.03 | n.a. | 0.01 | 0.05 | n.a. |
| Disadvantageous | 0.00 | 0.94 | n.a. | 0.00 | 0.85 | n.a. | 0.00 | 0.77 | n.a. |

**Notes:** This table displays the firm-type classification using simulated data under two different scenarios. *In Panel A*, we consider a scenario in which firms are equally distributed across the three types (advantageous, neutral, and disadvantageous). *In Panel B*, the distribution of firm types matches the proportions obtained in our benchmark fixed-effects models (Table 1). Columns (1) to (3) show the classification with no symmetric mistakes ($p = 0$). Columns (4) to (6) and (7) to (9), respectively, display the classification with 20 ($p = 0.2$) and 40 ($p = 0.4$) percents of symmetric mistakes in reporting.

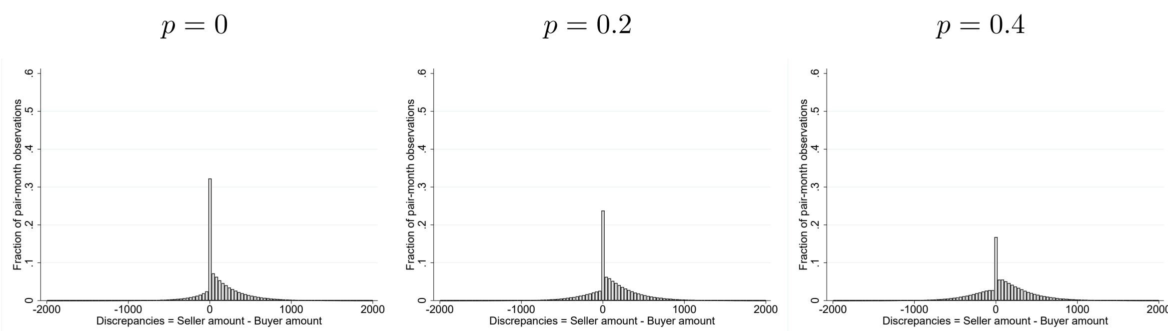
Figure D.1 depicts the distribution of the raw discrepancy in the amounts reported by seller and buyer in the simulated data, for different proportions of mistakes. In Panel A, the distribution when there are no mistakes ($p = 0$) is clearly centered around 0, while it becomes more spread out as we introduce more mistakes ($p = 0.2$ and $p = 0.4$). In Panel B, the distribution is also concentrated around zero but shows a small tilt towards positive values (where a positive value implies seller shortfall) similarly to the real firm-pair level data. As the share of mistakes increases, the distribution is also more spread out, as expected.

**FIGURE D.1**
**SIMULATED DISTRIBUTION OF DISCREPANCIES**

**PANEL A: 1/3 ADVANTAGEOUS, 1/3 NEUTRAL, 1/3 DISADVANTAGEOUS**




**PANEL B: 75% ADVANTAGEOUS & 25% DISADVANTAGEOUS**

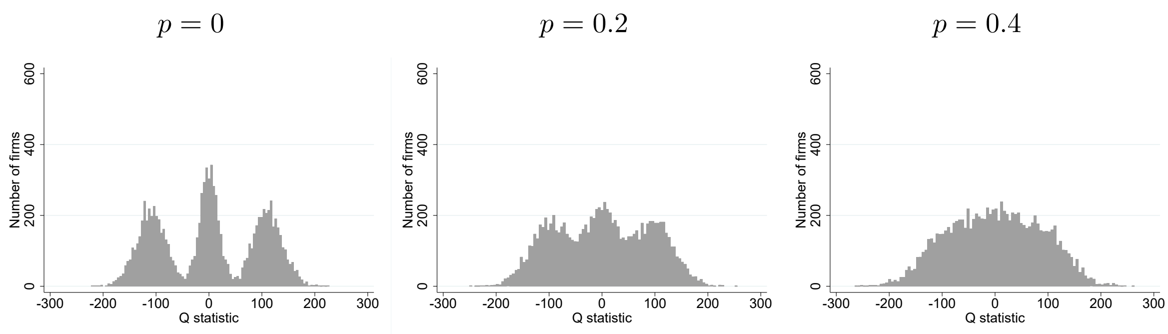



**Notes:** This figure displays the simulated distribution of discrepancies in the amounts reported by seller and buyer, under two different scenarios. *In Panel A*, we consider a scenario in which firms are equally distributed across the three types (advantageous, neutral, and disadvantageous). *In Panel B*, the distribution of firm types matches the proportions obtained in our benchmark fixed-effects models (Table 1). In both Panels, the figure to the left shows the classification with no symmetric mistakes ($p = 0$). The center and right figures, respectively, display the classification with 20 ($p = 0.2$) and 40 ($p = 0.4$) percents of symmetric mistakes in reporting.

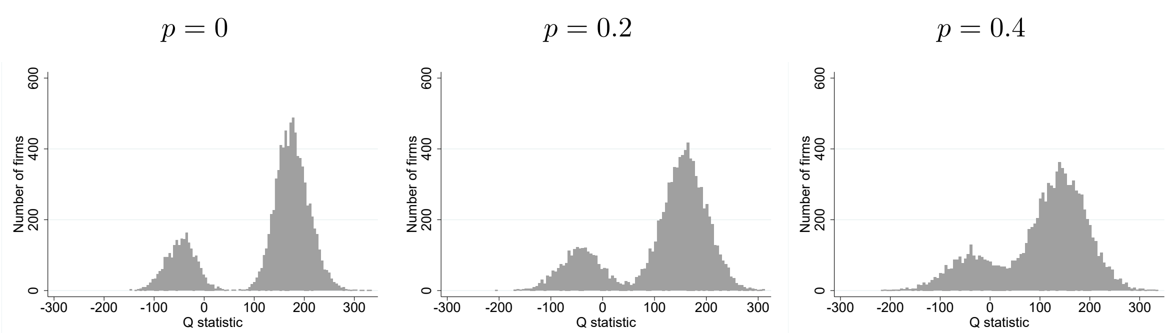
Figure D.2 depicts the distribution of the Q-statistic, which is a weighted sum of the estimated seller and buyer fixed effects at the firm level (see Section 4 for details on the estimation of Q). In Panel A, when $p = 0$ we observe a trimodal distribution, where there is a clear clustering of estimates around three groups: advantageous-type firms have Q estimates distributed around a positive mean, whereas the opposite is true for disadvantageous-type firms. Neutral firms have Q estimates distributed around 0, showing both positive and negative values. This three-group clustering fades as we increase the share of mistakes, such that the mean Q estimate for advantageous firms declines, and the mean Q estimate for disadvantageous firms increases. In Panel B, when there is no mistakes we observe two groups of estimates: disadvantageous-type firms tend to have estimates of Q below zero, with a few estimates just above zero, while the more numerous advantageous-type firms have Q estimates distributed around a positive mean, and none of them fall below zero. As the share of mistakes increases (to $p = 0.2$ and $p = 0.4$), the two distributions become flatter and they have some overlap. Consistent with the results presented in Table D.2, this leads mainly to misclassification of disadvantageous firms as advantageous, but not the other way around.

**FIGURE D.2**
**SIMULATED DISTRIBUTION OF ESTIMATED Q-STATISTICS**

**PANEL A: 1/3 ADVANTAGEOUS, 1/3 NEUTRAL 1/3 DISADVANTAGEOUS**




**PANEL B: 75% ADVANTAGEOUS & 25% DISADVANTAGEOUS**




**Notes:** This figure displays the simulated distribution of $Q(f)$, under two different scenarios. *In Panel A*, we consider a scenario in which firms are equally distributed across the three types (advantageous, neutral, and disadvantageous). *In Panel B*, the distribution of firm types matches the proportions obtained in our benchmark fixed-effects models (Table 1). In both Panels, the figure to the left shows the classification with no symmetric mistakes ($p = 0$). The center and right figures, respectively, display the classification with 20 ($p = 0.2$) and 40 ($p = 0.4$) percents of symmetric mistakes in reporting.

## D.3 Panel Simulation

In Section 4.3 of the paper, we report that a firm identified as advantageous in year $t$ has a 74% chance of being identified as advantageous in year $t + 1$, while the equivalent proportion is 65% for disadvantageous firms (see Table B.2). One question that arises from those results is whether the proportions are below 100% because firms switch types over time, or because of noise in the data due to (potentially symmetric) reporting mistakes. In order to explore this question, we incorporate a time dimension to our simulation. In particular, we follow the procedure described in Section D.2 twice to generate a two-period dataset of simulated firm pairs. We also allow firms to change their deterministic type across periods, denoting with $c$ the share of firms that change their type between $t$ and $t+1$ due to a real behavior change.

The results of this exercise are reported in Table D.3, which shows the proportion of firms that are classified as advantageous (and, in parentheses, disadvantageous) in period $t$ and also in period $t+1$. The parameter $c$ denotes the share of firms that are assigned ex-ante to switch their type between the two periods. The parameter $p$ indicates the share of extensive-margin mistakes, as described in Section D.2.

Panel A of Table D.3 reports the results for the simulation with one-third of firms of each type: advantageous, neutral and disadvantageous. When there are no mistakes ($p = 0$) and we allow no type switches ($c = 0$), the estimated firm type (advantageous or disadvantageous) is the same in both periods for about 83-84% of firms. This percentage masks considerable heterogeneity: firms whose true type is advantageous are correctly classified 100% of the time in both periods, as well as firms whose true type is disadvantageous. However, firms whose reporting behavior is neutral (i.e., those that always report the true amount) are evenly split in period $t$, and then again in period $t+1$. Therefore, we should expect only half of these neutral firms to be classified with the same type across periods. Since 33.3% of firms are neutral, about half of them (16.6%) are classified with the same type in both periods. Adding that to the 66.6% of firms that are consistently classified as advantageous or disadvantageous yields the 83-84% obtained in the aggregate. As we increase the share of firms who switch types ($c$) or the share of reporting mistakes ($p$), the percentage of firms that receive the same classification in both periods declines, but less than proportionally. For example, with $c = 0.1$ and $p = 0$, the percentages go down to 81% and 80% (for advantageous and disadvantageous, respectively). With $c = 0$ and $p = 0.2$, they decline to 79% for both types. When we assume that $c = 0.3$ and $p = 0.4$, the percentages decline to 65% and 63%, respectively.

Panel B of Table D.3 reports the results for the version of the simulation with 75% advantageous and 25% disadvantageous firms. In this case, the estimated firm types are highly consistent over time, though not perfectly so: 98% of firms identified as advantageous in period $t$ have the same type in period $t+1$, while the same is true for 96% of firms classified as disadvantageous in period $t$. This is consistent with the results obtained in the static simulation: since 6% of disadvantageous firms are incorrectly classified as advantageous in period $t$, we would expect some of them to be correctly classified as disadvantageous in period $t+1$. As we increase the proportion of type switchers ($c$), these percentages decline mechanically both for firms initially classified as advantageous and disadvantageous. When $c = 0.3$ and $p = 0$, we find that 70% of firms classified as advantageous (and 67% of those classified as disadvantageous) in period $t$ receive the same label in $t+1$. At the other extreme, when $c = 0$ and $p = 0.4$, the percentages are 94% and 77%, respectively. Finally, when $c = 0.3$ and $p = 0.4$, the percentages decline to 70% and 58%, respectively. These results imply that the increase in the proportion of switchers ($c$) has a mechanical (negative) effect on the proportion of firms classified with the same type in both periods. Meanwhile, an increase in the share of mistakes has a small effect on the consistency of advantageous classification, but a larger effect on the consistency of disadvantageous classification.

The overall conclusion from this panel simulation is that the estimated transitions across types observed in Section 4.3 is consistent with a situation where there is a substantial share of symmetric reporting mistakes, as well as a non-negligible fraction of type switchers across periods. That said, the fixed-effects analysis is able to capture a systematic component of firm behavior that is broadly persistent across periods.

**TABLE D.3**
**SIMULATED FIRM-TYPE CLASSIFICATION: CONSISTENCY ACROSS PERIODS**

| | | *Panel A: 1/3 advantageous, 1/3 neutral, 1/3 disadvantageous* | | |
|---|---|---|---|---|
| | | Mistakes | | |
| | | $p = 0$ | $p = 0.2$ | $p = 0.4$ |
| Switchers | $c = 0$ | 0.84 (0.83) | 0.83 (0.83) | 0.80 (0.79) |
| | $c = 0.1$ | 0.81 (0.80) | 0.80 (0.79) | 0.76 (0.75) |
| | $c = 0.2$ | 0.77 (0.77) | 0.77 (0.75) | 0.74 (0.72) |
| | $c = 0.3$ | 0.73 (0.74) | 0.72 (0.72) | 0.70 (0.68) |
| | | *Panel B: 75% advantageous & 25% disadvantageous* | | |
| | | Mistakes | | |
| | | $p = 0$ | $p = 0.2$ | $p = 0.4$ |
| Switchers | $c = 0$ | 0.98 (0.96) | 0.96 (0.87) | 0.94 (0.77) |
| | $c = 0.1$ | 0.90 (0.84) | 0.90 (0.78) | 0.89 (0.72) |
| | $c = 0.2$ | 0.81 (0.76) | 0.81 (0.70) | 0.80 (0.64) |
| | $c = 0.3$ | 0.70 (0.67) | 0.71 (0.62) | 0.70 (0.58) |

**Notes:** This table displays the percentage of firms labelled as advantageous (disadvantageous) in period $t$ which are also classified as advantageous (disadvantageous) in period $t + 1$. The parameter $p$ represents the proportion of symmetric mistakes within transactions, whereas $c$ stands for the share of type-changing firms.

# E Switchers' Graph

We use an event study approach to analyze reporting discrepancies incurred by firms that switch trade partners. We define a switch as chronological pairs of trading spells involving the a given buyer (resp. seller) but two different sellers (resp. buyers). Figures E.1 and E.2 show how reporting discrepancies change around such events. An old trading spell is defined as a sequence of at least two consecutive months in which a buyer-seller pair trades with each other, and subsequently stops trading for at least two months. The second to last period is labeled as time $t-2$, hence the last period before they stop trading is labeled as $t-1$. At time $t$, the firm stops trading with the set of old trade partners by construction. Starting from $t$, we similarly identify a set of new trade partners for each buyer (seller), composed of new sellers (buyers) which a given buyer (seller) has not traded with in at least the last two periods, but maintains the trading relationship for at least two consecutive months following the switch.

We classify the two trade partners (old and new) involved in the event into quartiles using the average detrended discrepancies they incur in the VAT reporting with *other* firms during the two months that the spell ends (for the earlier spell) or starts (or the later spell)—-analogous to firm's coworker wages in Card *et al.* (2013). In the figures, the horizontal axis displays event time, i.e., trading months. The vertical axis displays the average detrended discrepancies the firms of interest (sellers in Figure E.1 and buyers in Figure E.2) have in a given month.[32]

The results are shown in Figures E.1 and E.2. We make two initial observations. First, discrepancies change sharply, and in the expected direction when a firm trades with different partners associated with different degrees of reporting discrepancies: For example, sellers switching from a lowest quartile-buyer to a highest quartile-buyer experience a significant increase in reporting discrepancies; whereas sellers switching from a highest quartile-buyer to a lowest quartile-buyer experience a significant decrease in reporting discrepancies. Second, the figures show no systematic pattern that discrepancy is rising for firms that subsequently switch to a higher quartile-partner, and vice versa, thus suggesting that drift in discrepancies in switches are uncorrelated.

More importantly, the discrepancy changes associated with switching trade partners appear symmetric: firms switching from a partner in the highest quartile to another partner in the lowest quartile experience an increase in the average reporting discrepancies of similar magnitude to those switching in the other direction. This provides strong evidence against selection of trade partners based on discrepancies—a concern potentially invalidating the estimates from the fixed-effects model.[33] The striking symmetry in the switchers' graphs indicates that the selection effect is unlikely to bias our estimates.

[32]We detrend raw discrepancies by regressing them on month-year fixed effects.

[33]This selection or sorting would imply that firms switching from partners associated with large discrepancies to partners associated with small discrepancies experience greater decreases than the increase experienced from moving in the opposite direction: firms trading with a small-discrepancy partner enjoy both a small average discrepancy and an improved match effect, and firms trading with a large-discrepancy partner is hit with a large average discrepancy but an offsetting improved match effect.

FIGURE E.1
SWITCHERS' GRAPH: SELLERS

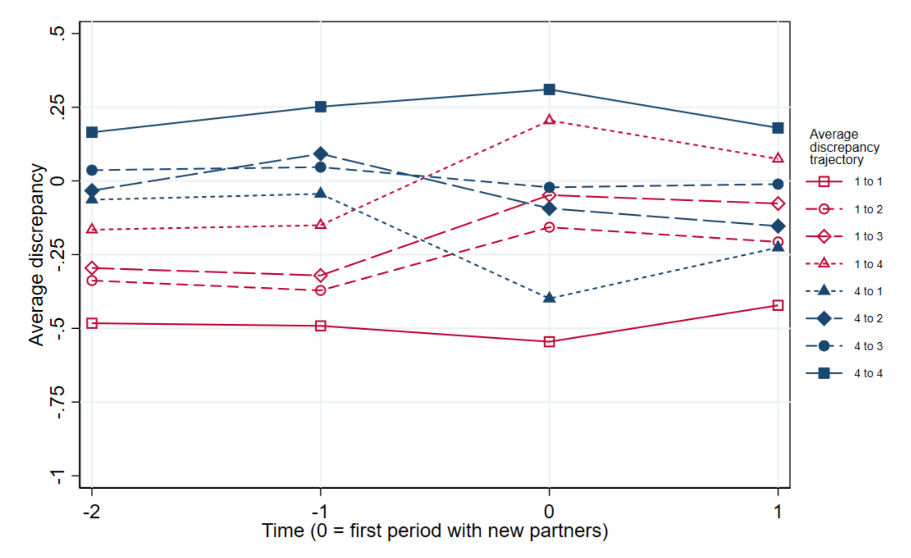
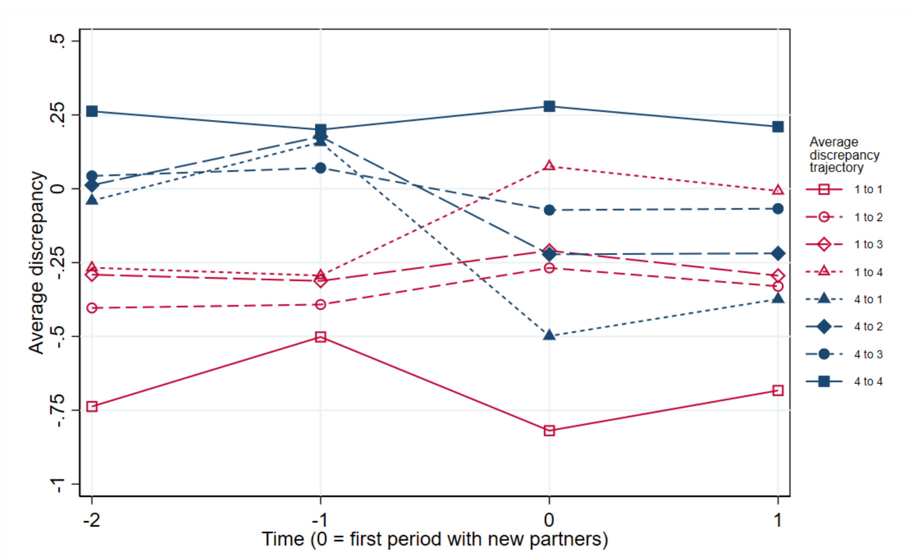



FIGURE E.2
SWITCHERS' GRAPH: BUYERS



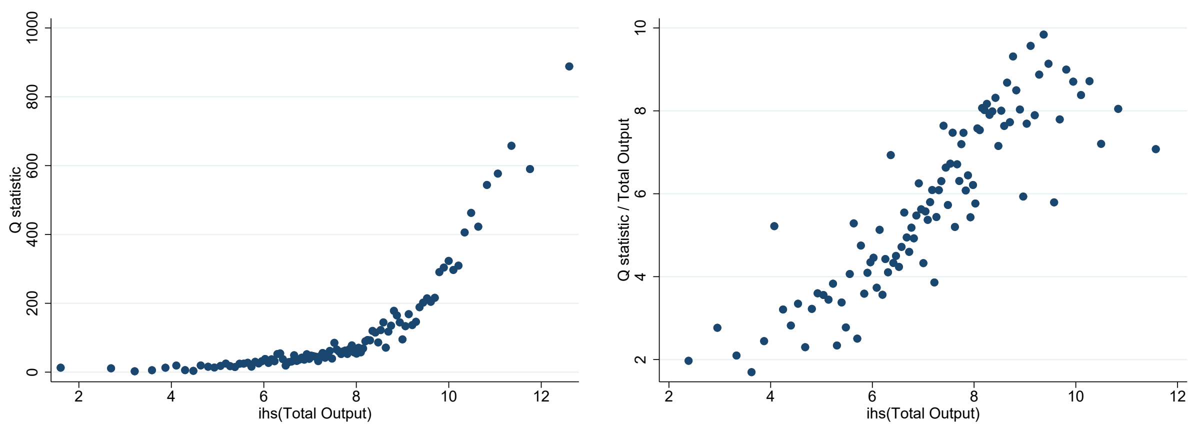

**Notes:** The figures show time trends in discrepancies around the time that firms switch trade partners, for sellers in the top panel and buyers in the bottom panel. We define an "old trade partner" as a firm that has at least two consecutive months of trade with the firm under consideration and subsequently stops trading for at least two months; whereas a "new trade partner" is one that the firm has not traded with previously. The figures plot the average discrepancy on the vertical axis and event time (i.e., trading months) on the horizontal axis, for different types of quartile-to-quartile switches.

# F Additional figures

In this section, we present additional figures mentioned in the paper. Figure F.1 shows how firms' estimated Q-statistic correlates with firm size. The average $Q_f$ measure is similarly distributed across most of the distribution of firm size. However, the figure also shows that the average $Q_f$ measure markedly increases among the largest firms, suggesting that the largest firms are more sophisticated tax (mis)reporters than other firms.

**FIGURE F.1**
**Q STATISTIC OVER FIRM SIZE**

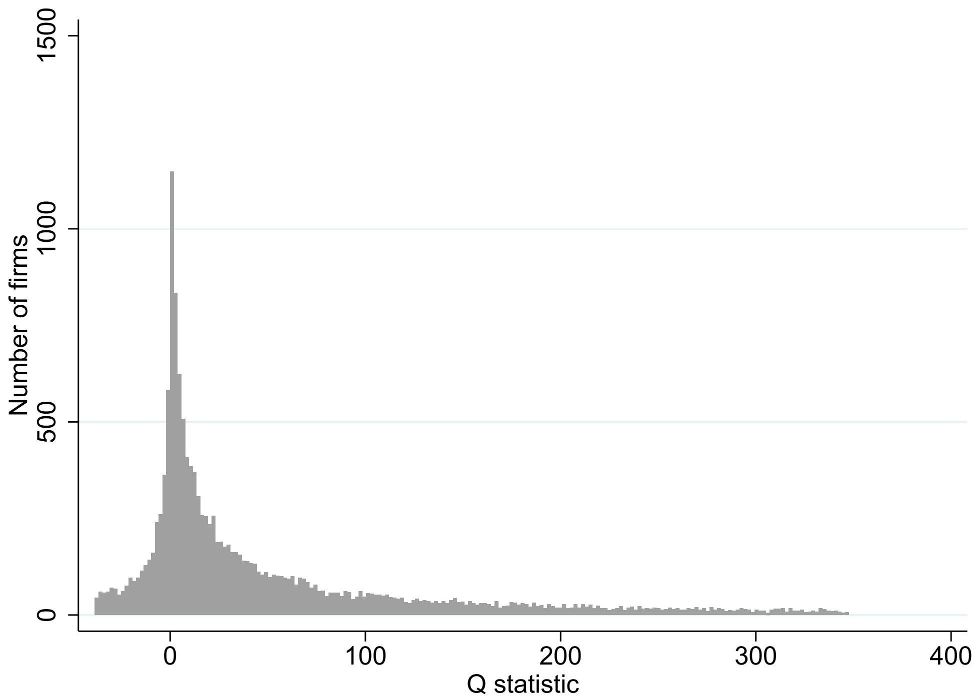
**(A) Q STATISTIC (ABSOLUTE)** **(B) Q STATISTIC (RELATIVE TO FIRM SIZE)**




**Notes:** *In Panel A*, we plot firms' estimated Q statistic — $Q_f$ in equation (2) — over the inverse hyperbolic sine transformation of firms' total output in the study period. *In Panel B*, we first normalize Q estimates by firm sizes, and then plot them against firm size. Data source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016.

Figure F.2 shows the distribution of Q-statistic across firms. The histogram is concentrated around 0, and clearly skewed to the right. This illustrates the fact that a majority of firms have a positive Q-statistic and are labeled as advantageous reporters.

**FIGURE F.2**
**DISTRIBUTION OF Q STATISTIC.**




**Notes:** In this Figure, we plot the distribution of firms' estimated Q statistic ($Q_f$ in Equation (2)). Data source: VAT Schedules data for fiscal years 2013-2016.

# G Additional tables

Table G.1 shows the distribution of value-added and VAT liability by firm size for fiscal years 2013-2016. While only 15% of firms report negative or zero value-added over the four fiscal years (the difference between total output VAT and total input VAT is also low at 22%), the reported VAT liability is zero or negative for 52% of the firms. Value-added proportions are similar for LTO and MTO firms, while the share with a value-added equal or less than zero is higher for other VAT firms. The proportions for the difference between total output VAT and total input VAT as well as for VAT liability are generally similar across firm size.

**TABLE G.1**
**DISTRIBUTION OF VALUE-ADDED AND VAT LIABILITY BY FIRM SIZE**

| | | (1)<br>Value added | (2)<br>Output-Input VAT | (3)<br>VAT liability |
|---|---|---|---|---|
| | Share $> 0$ | 84.33% | 77.36% | 48.26% |
| All VAT firms | Share $= 0$ | 5.12% | 7.43% | 6.47% |
| (N = 22,388) | Share $< 0$ | 10.55% | 15.21% | 45.27% |
| | Share $> 0$ | 93.08% | 77.75% | 48.64% |
| LTO firms | Share $= 0$ | 0.81% | 0.77% | 1.28% |
| (N = 738) | Share $< 0$ | 6.11% | 21.49% | 50.07% |
| | Share $> 0$ | 91.85% | 79.94% | 50.69% |
| MTO firms | Share $= 0$ | 0.71% | 1.39% | 1.41% |
| (N = 1,635) | Share $< 0$ | 7.43% | 18.66% | 47.91% |
| | Share $> 0$ | 82.82% | 77.00% | 47.92% |
| Other VAT firms | Share $= 0$ | 5.95% | 8.62% | 7.44% |
| (N = 20,015) | Share $< 0$ | 11.22% | 14.39% | 44.63% |

**Notes:** Data source: VAT Monthly Summary data for fiscal years 2013-2016. Column (1) shows total value added, including goods that are VAT-exempt. Column (2) shows the difference between total output VAT and total input VAT. Column (3) shows total tax liability, taking into account VAT credits carried over from previous periods. Firms can display a positive Output-Input VAT, but a nil or negative VAT liability once offsets are subtracted. LTOs are firms with an annual turnover above 15 billion Ugandan Shillings (USD 4.1 million) and/or belonging to specific sectors such as oil and mining, banking, insurance, and government departments. MTOs are firms with a turnover above 2 billion Ugandan Shillings (USD 550,260, threshold increased to 5 billion Ugandan Shillings/USD 1.3 million in 2015). Other VAT firms refer to VAT-paying firms with an annual turnover lower than the MTO threshold.

Table G.2 shows variation in extensive (i.e., either trade partner fails to report any transaction with the trade partner in a given month) vs. intensive margin (i.e., conditioning on reporting, the reported amounts are different between sellers and buyers) proportions by firm characteristics. These shares are relatively stable, across sectors, and across firm size categories. Regarding transaction sizes, the share of extensive margin discrepancies decreases with transaction size.

## Table G.2
## Extensive margin versus Intensive margin discrepancies by firm characteristics

| | Share of transactions with... | | |
|---|---|---|---|
| Firm characteristics | No Discrepancy | Extensive Margin discr. | Intensive Margin discr |
| MTO/LTO | 0.21 | 0.63 | 0.16 |
| STO | 0.20 | 0.70 | 0.11 |
| | | | |
| Transaction size: Large | 0.21 | 0.53 | 0.26 |
| Transaction size: Medium | 0.21 | 0.68 | 0.11 |
| Transaction size: Small | 0.19 | 0.78 | 0.03 |
| | | | |
| Agriculture, forestry, fishing | 0.27 | 0.59 | 0.13 |
| Mining, Quarrying | 0.24 | 0.64 | 0.12 |
| Manufacturing | 0.25 | 0.59 | 0.16 |
| Water, Electricity services | 0.09 | 0.84 | 0.07 |
| Construction | 0.26 | 0.58 | 0.16 |
| Wholesale and retail | 0.22 | 0.64 | 0.14 |
| Transportation, accomodation services | 0.15 | 0.74 | 0.11 |
| Information, communication | 0.11 | 0.74 | 0.15 |
| Financial services | 0.08 | 0.84 | 0.08 |
| Real estate | 0.18 | 0.70 | 0.12 |
| Professional, Admin, Other Services | 0.20 | 0.68 | 0.11 |
| Public Administration | 0.17 | 0.73 | 0.08 |
| Education | 0.09 | 0.85 | 0.05 |
| Health and social work | 0.26 | 0.67 | 0.06 |
| Arts and Entertainment | 0.15 | 0.75 | 0.10 |
| Total | 0.21 | 0.66 | 0.13 |

**Notes:** Data source: VAT Monthly Summary and VAT Schedules data for fiscal years 2013-2016. This table displays the share of pair-month transactions that display no discrepancy (the seller and the buyer declare the same amount, we allow for rounding of 1,000 UGX and for pure timing mismatches.), a discrepancy on the extensive margin (either the seller or the buyer doesn't declare the transaction at all), and a discrepancy on the intensive margin (the seller and the buyer declare different positive amounts), by firm characteristics. Observations are at the firm-month level, and are associated to firms' characteristics irrespective of whether the firm is the buyer or the seller. Firms are categorized either as MTO/LTO (Medium Taxpayer Office, Large Taxpayer Office), or STO (Small Taxpayer Office). Transaction size is defined by tercile of the maximum amount declared by either trade partner. The sector categories correspond to the firm's sector as listed in the tax registry.

Table G.3 shows how VAT liability is related to firm types. In particular, firms with null or positive VAT liability are more likely to be disadvantageous misreporters, whereas firms with negative VAT liabilities correlate with advantageous misreporting.

## TABLE G.3
## FIRM-TYPE AND VAT MONTHLY LIABILITY

| | Dep. Var.: VAT Liability | | | | | |
|---|---|---|---|---|---|---|
| | Null | Null | Positive | Positive | Negative | Negative |
| Firm Type | (1) | (2) | (3) | (4) | (5) | (6) |
| Disadvantageous | 0.034*** | | 0.035*** | | -0.069*** | |
| | (0.005) | | (0.007) | | (0.006) | |
| Negative Buyer FE | | 0.030*** | | 0.039*** | | -0.069*** |
| | | (0.005) | | (0.007) | | (0.006) |
| Negative Seller FE | | -0.035*** | | 0.025*** | | 0.010 |
| | | (0.005) | | (0.007) | | (0.007) |
| Month-Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Size | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 712927 | 712927 | 712927 | 712927 | 712927 | 712927 |
| R-squared | 0.02 | 0.02 | 0.00 | 0.01 | 0.01 | 0.02 |
| Mean of dep. | 0.19 | 0.19 | 0.46 | 0.46 | 0.35 | 0.35 |

**Notes:** Data source: VAT Schedules and Monthly Summary data for fiscal years 2013-2016. This table shows the results of the regression of monthly VAT liability on firm-type. In Columns (1) and (2) (resp., Columns (3) and (4), resp. Columns (5) and (6)), the dependent variable is a dummy equal to one if the VAT liability is null (resp., positive, resp. negative). In Columns (1), (3), (5), the regressor of interest is $Disadvantageous$, a dummy equal to one if the firm is categorized as Disadvantageous and zero otherwise. In Columns (2), (4), (6), the regressors of interest are two dummies, *Negative Buyer FE* and *Negative Seller FE*, equal to one if the firm's buyer (resp., seller) fixed-effect is equal to zero. We control for firm size in all specifications, with a categorical variable indicating whether a firm is classified as medium taxpayer (MTO), large taxpayer (LTO), or none (STO). Standard errors, clustered at the firm level, are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table G.4 analyzes whether the likelihood of seemingly anomalous beneficial behavior at customs varies across months within the year, and with monthly VAT liability as reported in the MVR. The 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.

## TABLE G.4
## SEEMINGLY ANOMALOUS CUSTOMS REPORTING BY VAT LIABILITY AND BY MONTH

| Dependent variable: | SA | SA Extensive | SA Intensive | SA | SA Extensive | SA Intensive |
|---|---|---|---|---|---|---|
| Customs behavior | (1) | (2) | (3) | (4) | (5) | (6) |
| July | -0.013** | -0.020*** | 0.001 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| August | -0.015** | -0.016*** | -0.003 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| September | -0.010* | -0.011** | -0.002 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| October | -0.006 | -0.007 | -0.003 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| November | -0.002 | -0.003 | -0.001 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| January | -0.011** | -0.007 | -0.007 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| February | -0.040*** | -0.017*** | -0.032*** | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| March | -0.019*** | -0.019*** | -0.005 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| April | -0.024*** | -0.021*** | -0.010 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| May | -0.026*** | -0.028*** | -0.007 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| June | -0.021*** | -0.027*** | -0.002 | | | |
| | (0.006) | (0.005) | (0.007) | | | |
| Null VAT | | | | 0.220*** | 0.295*** | 0.102*** |
| | | | | (0.014) | (0.015) | (0.018) |
| Year FE | Yes | Yes | Yes | | | |
| Month-Year FE | | | | Yes | Yes | Yes |
| Size and Sector FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 123304 | 123304 | 76510 | 123304 | 123304 | 76510 |
| R-squared | 0.02 | 0.06 | 0.01 | 0.04 | 0.09 | 0.01 |
| Mean of dep. | 0.34 | 0.20 | 0.23 | 0.34 | 0.20 | 0.23 |

**Notes:** Data source: VAT Schedule 3, MVR and Customs data for fiscal years 2013-2016. Observations are at the firm-month level. The dependent variable in Columns (1) and (4) is a dummy equal to one if the firm claims lower VAT amounts incurred on imports in VS3 than VAT paid on imports as recorded in the Customs data for the same month. We allow for rounding of 1,000 UGX and for pure timing mismatches. In Columns (2) and (5), the outcome variable indicates seemingly anomalous reporting on the extensive margin, equal to one if the firm reports nothing in VS3 for a month in which VAT paid on imports at customs is non-zero. In Columns (3) and (6), we restrict the sample to firm-month observations where a positive amount is reported both at Customs and in VS3, and the dependent variable is a dummy indicating seemingly anomalous behavior on the intensive margin, equal to one if the VAT claimed in VS3 is lower than the VAT paid on imports as reported in Customs. In Columns (1) to (3), the explanatory variables are dummies for each month. The reference is December. Note that the fiscal year in Uganda runs from July to June. Months are based on invoice dates. In Columns (4) to (6), the explanatory variable of interest is a dummy equal to one if the VAT liability reported in the MVR is zero. In all specifications, we control for firm size as measure by annual decile of reported turnover, and for firm sector. Standard errors, clustered at the firm level, are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table G.5 analyzes whether disadvantageous firms are more likely to behave in a self-beneficial way at customs. Specifically, we find that self-advantageous misreporting of imports is unrelated to firm-types.

**TABLE G.5**
**SELF-ADVANTAGEOUS MISREPORTING OF IMPORTS**

| | **Dep.Var.: Self-advantageous misreporting** | |
|---|---|---|
| Firm Type | (1) | (2) |
| Disadvantageous | 0.001 | 0.005 |
| | (0.007) | (0.007) |
| Null VAT | | -0.081*** |
| | | (0.006) |
| Month-Year FE | Yes | Yes |
| Size and Sector FE | Yes | Yes |
| HS Share of Import | No | Yes |
| Observations | 123303 | 123303 |
| R-squared | 0.05 | 0.06 |
| Mean of dep. | 0.14 | 0.14 |

**Notes:** Data source: VAT Schedule 3, MVR and Customs data for fiscal years 2013-2016. Observations are at the firm-month level. The dependent variable is a dummy equal to one if the firm claims higher VAT amounts incurred on imports in VS3 than 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. 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 $Q_f$ from equation (2). In all specifications, we control for firm size as measure by annual decile of reported turnover, and for firm sector. In Column (2), we additionally control for null monthly VAT liability as reported in MVR, 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.