Tax Evasion in Relational Contracts*

Brian Dillon, Cornell University

Jihye Jang, Cornell University

John Mulenga, Zambia Revenue Authority

Twivwe Siwale, International Growth Centre

April 2025

Abstract

Do firm-to-firm relationships involve tacit collusion to evade taxes? We examine this question in Zambia, where fiscal capacity is low and many firm-to-firm transactions do not generate a paper trail. For six months we randomly incentivize retail firms to request formal tax invoices from their suppliers, which makes the transaction more transparent to the tax authority. We find that both high and low levels of financial incentives induce retail firms to collect more tax invoices from their suppliers. Firms with strong supplier relationships at baseline are less responsive to the incentives, consistent with the presence of relational contracts under which there is an agreement that transactions will not be recorded. Using administrative data from the revenue authority we find that our interventions had a small but detectable impact on total VAT revenues.

*We are grateful to the International Growth Centre for funding that supported this project, to Taonga Mwale for outstanding leadership of the field team, to the full team of enumerators, and to Laurel Krovetz, Dana Smith, and Ben Norton for research assistance. This work has benefited from conversations with Pierre Bachas, Anne Brockmeyer, Anders Jensen, Rocco Machiavello, Joana Naritomi, and seminar participants at Cornell, IGC, the Zambia Revenue Authority, and the 8th Zurich Conference on Public Finance in Developing Countries. The project is registered at the AEA RCT Registry as trial number AEARCTR-0011009.

1 Introduction

Informal relationships with suppliers are critical for many small and medium enterprises (SMEs). These relationships are often sustained not by formal contracts, but by *relational contracts*, i.e., by the future value of the relationship (Baker et al., 2002). Relational contracts are especially important in low-income countries, where firms have limited recourse to contract enforcement, and many straddle the boundary between formality and informality (Macchiavello, 2022). Recent empirical work has emphasized the importance of these arrangements for providing supply assurance, demand assurance, and trade credit in atmospheres of substantial uncertainty (Macchiavello and Morjaria, 2015,0; Ghani and Reed, 2017; Rudder and Dillon, 2023).

The role of mutually beneficial tax evasion has received less attention in the literature on firm-to-firm relationships in low-income countries. Yet, two factors raise the possibility that tax evasion may be a central feature of firm relationships. The first is that these firms operate in information-scarce environments. Governments rely on information generated by third parties to raise the real or perceived cost of evasion (Kleven et al., 2011; Jensen, 2019). If the "third party" is a firm in a long-running relationship with another firm, and the environment is unlikely to generate other information that would draw the attention of tax authorities, then the structure of firm-to-firm relationships could undercut the effectiveness of tax policy. The second factor is that many small and medium enterprises (SMEs) operate with razor thin profit margins. For these firms, the financial savings from tax evasion could be the difference between positive and negative profits.

In this paper we use two randomized control trials (RCTs) to understand how tax policy and collusion to evade taxes shape the relational contracts between firms. The setting is the retail sector in Lusaka, Zambia. In Zambia, all firms with revenue above a threshold are required to collect and remit value-added taxes (VAT).[1] The VAT is a popular tax tool in low-income countries because it creates opposing incentives for VAT-registered firms along the supply chain (the buyer is incentivized to provide information about the seller's tax liability). The VAT mechanism breaks down at the "last mile," because the final consumer (who is not VAT-registered) has no reason to request a tax invoice or report the purchase (Pomeranz, 2015).

For the first RCT (the lottery RCT), we began with a census of 36 markets and commercial areas in Lusaka. Our target population consisted of firms that were too small to be registered to pay VAT, but established enough to have a fixed physical presence (a shop). These small retailers represent an important bloc of buyers from VAT-registered wholesalers. We randomized these firms to Control, Low Treatment, and High Treatment, stratifying on the strength of supplier relationships at baseline. Treated firms could win cash during monthly follow-up visits, by playing a game in which the win probability increased in the total value of VAT invoices the firm had retained since our previous visit (with higher value prizes for the High group than the Low group). The Control group did not play the lottery.

[1] The threshold is 800,000 kwacha per year, or about USD 35,000.

The aim of the lottery RCT was to increase the value to buyers of requesting formal VAT invoices from their suppliers. Requesting an invoice would increase the probability that the seller would be required to pay taxes on the transaction, and hence would be costly for the seller (in expectation). We are interested in whether firms with strong supplier relationships are less responsive to the incentive. If so, that would suggest that their relationships were in part premised on not creating a paper trail of their transactions, i.e., on tax evasion. We also examine whether firms that request invoices experience a reduction in the relational services they receive from suppliers (e.g., goods on credit, or favorable delivery terms).

In the second RCT (the messaging RCT), which we ran immediately after the lottery RCT, we randomly assigned firms to receive either a Treatment message or a Control message. The Control group received a placebo message. The Treatment group received a script incorporating key elements from prior work on tax morale messaging: one passage primed peer effects, another emphasized reciprocity, a third primed fairness. The messages were read aloud to respondents during survey visits, and also provided as a printed letter. The goal of the messaging RCT was to study whether the financial incentives in the lottery RCT led to habit formation in regard to invoice-requesting behavior, and whether the extent of habit formation was influenced by tax morale messaging. The main outcomes in both RCTs are the value and share of VAT invoices retained from wholesale purchasing. From the seller's perspective, issuing a VAT invoice increases the probability that the ZRA learns about the transaction, which may raise the cost of under-reporting sales.

Our analysis leads to five main sets of findings. First, the financial incentives in the lottery RCT are impactful. On average across five follow-up rounds, the total value of VAT invoices retained by High treatment firms is 624 ZWK greater than the Control group, representing a 54% increase over the Control mean of 1162 ZWK ($p < 0.01$). The Low treatment effect is smaller in magnitude, at 392 ZWK ($p < 0.10$). These effects are consistent in magnitude over the five rounds, except for a drop in the Low group effect in the final round.

Second, firms with strong supplier relationships at baseline are less responsive to the financial incentives. We define the baseline strength of supplier relationships in two ways. The first is the percentage of stock bought from regular suppliers; the second is an index based on benefits received from suppliers, such as credit, favorable delivery terms, and others. With either definition, treatment effects are 37-45% larger, and confidence intervals tighter, for firms that have weaker supplier relationships at baseline. This pattern is consistent with the idea that firm-to-firm relationships reduce the effectiveness of tax-related incentives. A similar heterogeneity analysis with the control firms is instructive for interpreting magnitudes. In the Control group, firms with strong supplier relationships retain an order of magnitude *more* VAT invoices than firms with weaker relationships. We interpret this as indicating a role for compliance costs (anecdotes from data collection support this interpretation). Within the context of some retailer-supplier relationships, VAT invoices may already be exchanged, e.g., because the buyer places large orders. Suppliers may also be more willing to provide invoices to regular customers when they are asked. For these reasons,

relationships may be correlated with lower hassle costs for requesting invoices, even if they also serve to reduce the buyer's willingness to make such requests. If treated firms with strong relationships face lower compliance costs than treated firms with weaker relationships, then our estimated effects are a lower bound on the direct effect of firm-to-firm relationships in deterring tax compliance.

In our third set of findings we examine mechanisms. Treated firms make adjustments on multiple margins. Firms in the High treatment group are more likely to stop purchasing from a supplier that is not VAT-registered, and more likely to inquire about the VAT status of a new supplier. Both High and Low treatment firms see an increase in the share of inventory they acquire from regular suppliers, indicating an increase in attachment to sellers that provide invoices.

The fourth set of results deals with habit formation and the effects of tax morale messaging. To study habit formation, we examine whether firms that were exogenously induced to request more invoices in rounds 1-5, during the Lottery RCT), continue to do so once the incentives are removed. In the reduced form version of this analysis we find that the High treatment from the Lottery RCT remains positive and statistically significant across follow-up rounds 6-8, even as the average value of retained VAT invoices falls across all firms. The related IV analysis—in which we use Lottery treatment assignment as an instrument for the value (or share) of VAT invoices retained over the first five rounds, and then test whether firms that retain more invoices in the early rounds continue to do so in the later rounds—leads to similar conclusions.

Tax morale messaging has at most a modest effect on invoice requesting and retention. The ATE of the Messaging treatment is statistically significant when pooled across rounds 6-8, but the magnitude is small, and the effect is driven primarily by round 6 (the first round after the removal of financial incentives). In the IV analysis of the persistence of financial incentives, the messaging treatment has no additional effect on invoice retention.

The final set of results relates to suppliers' responses to the interventions targeting our sample firms. Surprisingly, VAT-registered suppliers issued more formal invoices to treated retailer firms and increased reporting of these invoices to the tax authority. This finding highlights the broader impact of firm-to-firm interactions on VAT compliance within supply chains. It suggests that interventions targeting retailers can create significant ripple effects upstream, influencing supplier behavior in ways that enhance transparency and compliance at multiple levels.

This collection of findings makes contributions on a number of fronts. Recent work has shown that incentivizing third-party agents to request VAT invoices can lead to sustained increases in tax remittances (Naritomi 2018). Programs such as these are part of a broader set of recent innovations to improve tax compliance, including providing incentives to tax collectors (Khan et al, 2018) and rewarding compliant taxpayers with public goods (Carrillo et al. 2017). The tax rebate lottery that we study differs from prior work, in that the program is available only to retail firms that meet specific criteria. Individual consumers, market vendors with no fixed stall, and non-retail shops are excluded from the program, as are any firms large enough to be VAT-registered. The target subgroup studied here has some

important characteristics: firms are easy to identify, easy to locate, likely to already have a taxpayer ID, likely to already use mobile money, and due to their collective buying power, able to exert substantial upward pressure on VAT-registered suppliers. Such a program has the potential to deliver net increases in public revenues in settings where an economywide program would be infeasible.

Second, our study adds to recent work on the importance of relational contracts for firms in low-income countries. Prior work has shown that relational contracts are important predictors of how firm-to-firm relationships navigate economic crises (Machiavello and Morjaria 2015), that competition may reduce relational contracting (Macchiavello and Morjaria, 2021), and that relational contracts can evolve in the face of entry by new suppliers (Ghani and Reed, 2021). To date, the emphasis in this literature has been on establishing the existence of these contracts and understanding the market failures that they help overcome, not on the way that their presence may undercut policy. We find evidence that firms with stronger supplier relationships are more resistant to requesting invoices, which is strongly suggestive of a link between the receipt of relationships services as a form of compensation for foregoing tax-related paperwork.

The third contribution is policy-related. These experiments were designed in full collaboration with the ZRA. The interventions represent policy options that are under active consideration in Zambia, or, in the Messaging RCT, variations on programs that already exist. The findings of this project are shared regularly within ZRA, and will be a direct input to decisions about new tax policy initiatives in the country.

Finally, we add to the literature on tax morale messaging. A key takeaway from this literature is that messages intended to prime non-pecuniary motivations for tax compliance have decidedly mixed effects (Luttmer and Singhal 2014; Jensen 2022). Tax morale messages can operate through various mechanisms, including reciprocity, i.e. the expectation of benefits from public services (Hoy et al. 2021), deterrence, i.e., the threat of possible sanctions from non-compliance (Mascagni and Nell, 2021), and peer effects, i.e., shame, pride, or other effects triggered by information comparing a taxpayer's performance with that of others in the community (Slemrod et al. 2021). Our messaging experiment differs from prior work in that our study firms are not registered for the VAT, and so are not the targets of the tax compliance initiative. Instead, they are the auditors. However, while study firms do not risk incurring tax-related penalties because of non-compliance, requesting invoices may lead them to incur other costs, either in the form of lost relationship benefits with suppliers, or through costly search for new suppliers. We find that messaging has a small effect on firms' willingness to incur these potential costs, but only for a short time.

The rest of the paper proceeds as follows. Section 2 outlines the motivation of the study, outlining the centrality of VAT in developing countries and the gaps that remain in studying its effectiveness. The section also lays out the research questions. Section 3 describes the experimental design, firm enrollment procedures, sampling criteria and dates and timings of interventions. Section 4 sets the empirical framework used to estimate the impact of the different inventions. Section 5 outlines the baseline findings and then proceeds to provide preliminary findings on the lottery

incentive experiment. Section 6 concludes.

2 The Value Added Tax in Zambia

As of 2016, VAT accounted for 25% of tax revenues in developing economies (Keen et al., 2016). VAT is intended to be self-enforcing because firms have a financial incentive to report purchases of intermediate goods on which they were charged VAT. Recent evidence suggests that this incentive does not always induce full compliance (Pomeranz, 2015). Even if VAT firms are compliant when selling to each other, the incentive breaks down at the "last mile" – the retail sale – where the buyer does not report the sale. Last-mile compliance problems are severe enough that some countries have reconsidered whether to have a VAT at all (Gérard and Naritomi, 2018; Waseem, 2023).

In Zambia, firms with revenue below a certain threshold do not register for VAT. These firms typically register to pay Turnover Tax (TOT), if they register at all. TOT firms and other small retailers are important buyers from VAT-registered suppliers. These suppliers may build mutually beneficial relationships with their buyers over time, sustained by relational contracts (Macchiavello, 2022). From these relationships, retail firms may receive favorable delivery terms, price discounts, goods on credit, advance notice of new products, or priority access to items in short supply. Their suppliers receive consistent and predictable demand, and, possibly, a buyer willing to forego VAT invoices on some purchases, which enables the supplier to evade taxes.

The program analyzed here was designed over multiple years of piloting and policy design consultation with ZRA economists and administrators. It is widely believed within ZRA that wholesale purchasing by small retail firms is a point of substantial VAT leakage. If so, study firms may be a key leverage point to increase compliance by VAT firms. By randomly incentivizing small retail firms to request formal VAT invoices from their suppliers, our lottery RCT is designed to raise the cost of supplier non-compliance.[2]

3 Experimental Design

We conducted two RCTs, in sequence, with a sample of retail firms in Lusaka, Zambia. The first experiment, the Lottery RCT, was a multi-arm intervention to encourage small firms to request VAT invoices from their suppliers. The second, the Messaging RCT, was a cross-randomized follow-up to the first RCT, to measure whether tax morale messaging could influence the persistence of any invoice-requesting habits formed during the lottery RCT.

In the following subsections, we describe the sample selection and enrollment process, incentives, assignment

[2]It is possible for us to observe an increase in issuance of VAT invoices to study firms, but with no increase in tax remittances to ZRA. This could occur if VAT-registered suppliers comply with invoice requests from study firms, but then shift their evasion to other transactions by not issuing invoices that they would have in the absence of the program (VAT-registered firms understand that they must issue some invoices in order to avoid scrutiny by auditors). In future analysis we will use ZRA administrative data to test whether we can detect an increase in tax remittances from the suppliers of study firms.

process, and timing.

3.1 Sample Selection and Enrollment

The target population for this RCT consists of all retail firms in Lusaka that are not registered to pay VAT, but are large enough to have a fixed physical location. We refer to this group as "small retailers." To recruit firms for the study, enumerators systematically walked through all markets and commercial areas of Lusaka, as identified from existing market lists and input from research team members and the ZRA. Table 2 provides the full list of market areas included in the enrollment process.

Because our goal was to recruit all non-VAT retail firms, enumerators did not approach manufacturing or service firms, large chains (e.g. Shoprite), shops that were clearly wholesalers, firms with a VAT registration certificate clearly displayed, or sellers with no permanent physical structure (e.g. roadside sellers). If an enumerator was unsure about the eligibility of a firm, they approached the firm and initiated a conversation. If the firm revealed itself to be VAT-registered during screening questions, the enumerator ended the interview, and the firm was not enrolled.

The sample recruitment and baseline survey process ran from late October to early December 2022. The team successfully interviewed 1,083 firms. The timeline of other completed and planned surveys is provided in Figure 1.

3.2 The Lottery RCT

3.2.1 Lottery RCT: Description of interventions

After the completion of the enrollment process and baseline survey, firms were randomized into three groups: Control, Low Treatment, and High Treatment. Firms in the Control group received no incentive for collecting VAT invoices. Firms in the Low Treatment group were eligible to win 35 ZWK if they won the lottery played during each follow-up visit. Firms in the High Treatment group were eligible to win 300 ZWK upon winning the lottery.

The lottery offered to treated firms consists of an opportunity to draw a card from a hat. The hat contains 9 black cards and 1 yellow card. The respondent was a winner if they drew the yellow card (10% win probability per play). Firms were given the opportunity to play the game during each of our follow-up visits, which were spaced approximately 1 month apart. The number of draws offered to a respondent during each follow-up visit, with replacement, was a function of the total value of official VAT invoices that they retain from purchases made since our previous visit. Invoices were eligible for the lottery if they contain the required details to be considered a formal invoice by ZRA staff, and if they represent purchases of items for (i) stock or inventory of the business, or (ii) capital improvements for the business. Enumerators verified the eligibility of every VAT invoice, including checking the dates and ensuring that the listed product was relevant to the business.

Treated firms were allotted chances to play the game as follows: 0 value VAT invoices = 0 draws; 1-1500 ZWK

value VAT invoices = 1 draw; 1501-3000 = 2 draws; 3001 or more = 3 draws. Each firm could only win one time per round. Lottery payouts were distributed via mobile money, within a few days of winning.

The stepwise mapping from the value of VAT invoices to the number of lottery draws was simple for participants to understand, but it created the risk of bunching at the kinks. We were especially concerned that firms would stop collecting invoices once they exceeded 3001 ZWK per round. To address this issue, we provided a second financial incentive to treated firms: a final "big lottery" at the end of study, with a payout of 1000 ZWK (about $55). The probability of winning the big lottery was proportional to the firm's share of total VAT invoices collected during the Lottery RCT. This ensured that the marginal expected value of additional VAT invoices was always positive.

3.2.2 Lottery RCT: Firm randomization

The 1,083 baseline firms were randomly assigned to the three Lottery RCT arms as follows: 407 firms to Control, 338 to Low Treatment, and 338 to High Treatment. We made the Control group slightly larger both because the cost to us of working with a Control firm was lower than an L or H firm, and because of possibly higher attrition from the Control group.

In assigning firms to experimental arms, we stratified on four baseline variables: market size (small, medium, or large, as categorized by knowledgeable team members), above/below median expenditure on stocks, above/below the median share of inventory purchases that generated a VAT invoice, and reliance on regular suppliers for stocking. Reliance on regular suppliers was captured by asking firms what percentage of their recent inventory was purchased from regular suppliers. We defined a "regular supplier" as one from which the retailer had purchased on at least two prior occasions, and from which the retailer planned to make purchases again in the future.

The intervention delivery took place in February-March 2023. A total of 898 out of the 1,083 baseline firms consented to participate in the Lottery RCT (83%). During the intervention delivery interview, respondents were given a plastic folder for safely retaining invoices and receipts, as well as an information sheet explaining the difference between valid and invalid invoices. This information sheet was provided to all participants.

3.3 The Messaging RCT

The Messaging RCT began in July 2023. The aims of this experiment were (i) to examine whether any invoice-requesting behavior induced by the Lottery RCT persisted after the incentives were removed, and (ii) to test whether tax morale messaging influenced the persistence of invoice requests. The setting differs from prior work on tax morale and intrinsic motivation because the subjects in our study are not the targets of tax enforcement (Luttmer and Singhal 2014). The costs that they may incur arise in the relationship and financial costs associated with requesting VAT invoices from suppliers or switching to suppliers that provide VAT invoices.

3.3.1 Messaging RCT: Description of interventions

In the Messaging RCT, treatment involved receiving a message that primed three aspects of tax morale. The full message was as follows:

> *We have appreciated your ongoing participation in this study, and we look forward to visiting 3-4 more times. Please continue to retain invoices and receipts. In future visits, we will conduct our normal short surveys (not as long as today) and will review the receipts and invoices that you retain. Remember that VAT-registered firms must provide you with an invoice if you request one, for any purchase.*
>
> *Over the last 4-5 months, ___% of firms like yours involved in our study have collected more VAT invoices than your firm, and ___% have collected the same amount or fewer VAT invoices.*
>
> *VAT firms almost always request invoices when buying and selling from each other, because they can use those invoices to claim tax refunds from ZRA. So ensuring that you get an invoice is a way to make sure that the tax you paid is sent to the government on your behalf.*
>
> *Essential public services, including market infrastructure, roads, electricity, schools, and hospitals are funded by taxes. When VAT-registered firms underpay taxes, everyone loses out. Asking for a VAT invoice is a simple yet effective way to help increase tax revenue and provide better services in Zambia.*

The first paragraph relays basic information about the study and the requirement that VAT-registered firms issue invoices, which constituted the placebo message for the Control firms. That paragraph was read aloud to all firms, and was included in written materials left with all respondents. The remaining three paragraphs were read only to the Treatment group from the Messaging RCT.

The second paragraph is intended to prime peer effects, by telling firms where they stand in the distribution of VAT invoice retention. It is unclear *ex ante* how learning one's relative position in the distribution might affect future behavior. One hypothesis is that high performance firms could be encouraged; yet they could also respond by slacking. Poor performance firms could be discouraged, or could respond by working harder to request and retain invoices. How

firms react may depend in part on how they perceive their suppliers, and how much effort they have exerted to get VAT invoices from their suppliers or to switch to more compliant suppliers.

The third paragraph of the Treatment message is intended to prime fairness. By highlighting that VAT firms typically request invoices when buying from their suppliers, because of the tax benefits, we aimed to imbue participants with a sense that it is only fair for VAT firms to provide invoices when requested.

The fourth paragraph primes reciprocity. We list some of the public goods that are funded with tax revenues, leading with those that are most directly relevant to participants in their capacity as small firms.

We opted for a multi-part tax morale message because the primary goal of this experiment is to test whether messaging is useful in converting an extrinsic motivation into an intrinsic one, not to measure the relative importance of different types of messages. Respondents will differ in how receptive they are to different types of tax morale messaging (if they are responsive at all). By priming a number of factors, we hoped to make the message relevant for the largest possible share of participants.

The Treatment and Control messages were read out loud by enumerators to respondents at the end of the fifth follow-up interview. Enumerators also provided respondents with a printed letter containing the relevant message in both English and Nyanja, on International Growth Center letterhead. The percentage fields written above as "___%" were automatically populated in the survey software; the enumerator then wrote in the appropriate percentages on the printed messages, before leaving with the participants.

3.3.2 Messaging RCT: Firm randomization

The 816 firms still participating in the Lottery RCT as of the fifth follow-up round were assigned to receive either the Treatment message (50%) or the Control message, stratifying on the treatment strata from the Lottery RCT, as well as on Lottery treatment status.

3.4 Data Collection and Timing

The data for this study were collected through 10 survey rounds conducted by the research team. The baseline survey ran from late October 2022 to early December 2022. The Lottery intervention delivery, and accompanying short survey, were implemented in February 2023. We then ran five post-treatment follow-up rounds for the Lottery RCT, spaced approximately one month apart, from March-July 2023 (we call those follow-up rounds 1-5). Enumerators delivered the Messaging treatment at the end of the round 5 follow-up interview. We then collected three more rounds of follow-up surveys, rounds 6-8, to measure the effects of the messaging treatment. Those rounds were also spaced roughly one month apart, from August-October 2023.

The long time-frame for the Lottery RCT follow-up surveys was important for three reasons. First, it gave firms

time to adapt. Second, it increased the net present value of requesting invoices and breaking the relational contract between retailer and supplier, if one existed (and was premised on tax avoidance). Third, it prevented firms from shifting expenses inter-temporally to increase their odds of winning the lottery. Study firms tend to keep inventory small and restock regularly.

The primary outcome variables, the total value of VAT invoices retained by the firm from inventory and capital improvement purchases, and the share of VAT-invoiced purchases as a percentage of all such expenditures, were measured during all follow-up surveys. All follow-up surveys also covered changes in shopping behavior (such as seeking out new suppliers or enquiring about VAT status of a supplier), the percentage of expenditures coming from regular suppliers, the number and total value of non-VAT receipts received, the reported value of all lost or damaged VAT invoices, and whether the firm was denied a VAT invoice by a supplier despite requesting one. We winsorize the 2% tails of all continuous variables.

Follow-up rounds 5 and 8, which we call the "midline" and the "endline", also included modules from the baseline survey that covered details about relational contracts and interactions with the most important suppliers. We use these data to examine whether relationships evolved over the course of the experiment.

4 Empirical Framework

4.1 Estimating the impact of lottery incentives

To estimate the impact of the Lottery RCT on the request and retention of VAT invoices, we estimate the following specification by linear regression, using data from follow-up survey rounds 1-5:

$$Y_{ist} = \alpha + \beta_l Low_{is} + \beta_h High_{is} + \mu_s + \gamma_t + \varepsilon_{ist} \tag{1}$$

where i indexes firms, s indexes randomization strata, and t indexes survey rounds. The variable Low_{is} is a dummy variable that takes on a value of 1 for firms assigned to the low treatment group, and 0 otherwise; $High_{is}$ is similarly defined, for the high treatment group; μ_s are strata fixed effects; γ_t are survey round fixed effects, and ε_{ist} is a statistical error term. The coefficients of interest are β_h and β_l. If financial incentives are effective in inducing firms to request and retain formal VAT invoices, we expect $\beta_h > \beta_l > 0$. In all specifications we cluster standard errors at the firm level, the unit of treatment assignment (Abadie et al., 2023).

The primary outcome variables in (1) are the total value of VAT invoices retained from business-related expenditures, and the share of VAT invoices in total stocking expenditure. In additional analyses we use variables related to shopping behavior that provide insight into mechanisms (e.g., changing suppliers, dropping non-VAT suppliers, losing relationship services, and similar). We also estimate variations on (1) that include interactions between treatment

variables and round dummy variables, to estimate impacts separately by survey round.

We interpret the estimated coefficients in (1) as average treatment effects (ATE). Control firms could not access the treatment, because we had complete control over the lottery and incentive payment system. Treated firms could not avoid learning about the possibility of playing the lottery and winning cash prizes, once they consented to participate. Even if some firms were unmoved by the incentives, they still received the treatment. Hence, compliance with treatment assignment was assured.

To test whether firms with stronger supplier relationships are less responsive to the lottery incentives, we estimate heterogeneous effects by baseline importance of supplier relationships. We use two measures of baseline relational contracts. The first is above/below the median share of recent purchases made from regular suppliers. The second is an index based on a series of questions from the baseline survey covering retailers' receipt of relationship-based services from suppliers, including price discounts, favorable delivery terms, items on credit, and others.

In order to estimate the ATE of the messaging treatment, we estimate the following regression using data from follow-up survey rounds 6-8:

$$Y_{ist} = \alpha + \beta_1 TreatMessage_{is} + \mu_s + \gamma_t + \varepsilon_{ist} \tag{2}$$

where i indexes firms, s indexes randomization strata, and t indexes survey rounds. The primary outcome variables are the same as those from the Lottery RCT. The variable $Treat_Message_{is}$ is a dummy variable that takes on a value of 1 for firms assigned to the messaging treatment group, and 0 otherwise. The coefficient of interest is β_1. The sign, magnitude, and statistical significance of this coefficient allow us to examine whether the messaging treatment increased firms' willingness to request and retain VAT invoices from their suppliers.

To examine whether habits formed during rounds 1-5 persist after the financial incentives are removed, we estimate the following model:

$$Y_{ist} = \alpha + \beta_1 TreatMessage_{is} + \beta_2 VAT_{is} + \beta_3 \{VAT X TreatMessage_{is}\} + \mu_s + \gamma_t + \varepsilon_{ist} \tag{3}$$

where VAT_{is} is the total value of VAT invoices retained by firm i in follow-up rounds 1-5 (when the financial incentives are active), and all other variables are as before. To isolate the exogenous component of VAT_{is} in (3), we instrument using treatment assignment in the Lottery RCT. The estimated value of the β_2 coefficient tells us whether there is persistence in VAT retention after the removal of financial incentives, possibly due to habit formation. A positive and significant value of β_3 would tell us that tax morale messaging has reinforcing effects on any habits formed by the exogenous inducement to participate in tax enforcement.

Finally, to estimate the impact of treatment on VAT remittances of the supplier VAT firms, we estimate the

following model:

$$VATRemittances_{it} = \alpha + \sum_{k=-8}^{4} (\beta_k \cdot TreatmentStrength_i \cdot Dummy_{t+k}) + \gamma_i + \varepsilon_{it} \quad (4)$$

where $VATRemittances_{it}$ represents the VAT remittances (in millions of Kwacha) for firm i in quarter t, as recorded in the Zambian Revenue Authority's administrative data. The term $Dummy_{t+k}$ is a set of event study dummies indicating whether quarter t is k quarters before or after the treatment. The variable $TreatmentStrength_i$ captures the relative size of the intervention and we use two measures, value of invoices issued relative to firm average sales and the count of invoices issued. And its interaction with $Dummy_{t+k}$ allows for differential treatment effects depending on treatment strength. The coefficients β_k measure the dynamic effect of the treatment on VAT remittances, scaled by $TreatmentStrength_i$. Firm fixed effects are included as γ_i to account for unobserved, time-invariant differences across firms, and ε_{it} represents unexplained variation in VAT remittances. The inclusion of interaction terms highlights how treatment effects vary with the intensity of treatment. For example, a positive and significant β_k for post-treatment quarters ($k \geq 0$) would indicate that higher treatment strength is associated with increased VAT remittances in those periods.

5 Results

5.1 Summary statistics from the baseline survey

Table 3 provides summary statistics from baseline firms. The average firm opened in 2016, and employs 1.7 full-time workers, 1.3 part-time workers, and 0.6 unpaid workers (typically family members of the owner). The mean number of daily customers is 43 on a weekday and 49 on a weekend day. Just over half of firms (53%) report keeping formal books or ledgers to track business costs and revenues; 31% of firms report retaining receipts for their business-related purchases, for at least a short period of time. The mean firm spent 11,851 ZWK on inventory purchases in the month before the baseline survey, of which 2,177 ZWK triggered the receipt of a VAT invoice. We did not verify the validity of these invoices at baseline (but will do during post-treatment follow-up visits).

Table 4 provides a characterization of the items sold by study firms. The most commonly sold item, "Drinks," includes everything from bottled water and soda to tea, coffee, and other non-alcoholic beverages. Other popular items include meat and fish, snacks, grocery items for home cooking (oils, sauce, spices), and dry foods such as beans, rice, and similar. The items sold by the fewest firms are cosmetics, alcohol, home goods (such as cleaning products), and fruits and vegetables. Fruits and vegetables are most commonly sold in open-air markets, on the street, or at small neighborhood kiosks, none of which were included in our sampling frame.

Figure 2 shows a histogram of the number of regular suppliers that each firm reported having over the previous

year. A regular supplier is defined as a business from which the firm made at least two prior purchases, and from which the firm intends to make purchases again in the future. Firms vary widely in their use of regular suppliers. A substantial share report having 0, 1, or 2 regular suppliers; the overwhelming majority have 6 or fewer. This variation hints at the different approaches to sourcing inventory deployed by study firms, which is potentially indicative of heterogeneity in the strength of their relational commitments to not request invoices and thereby avoid triggering tax remittances.

Variation in supplier relationships is even more evident in Figure 3, which shows a histogram of the percentage of recent inventory purchased from regular suppliers, measured at baseline. The alternative to sourcing from regular suppliers is to acquire inventory on the open market. This measure is essential to our study, as it is a key way to understand the importance of supplier relationships at baseline. The distribution shown in Figure 3 is bimodal. A plurality of firms report sourcing 100% of their inventory from regular suppliers. The next largest group, representing over 30% of firms, reports sourcing 0% of their inventory from regular suppliers. The remaining firms are concentrated between 51-99%, with a small share reporting a value between 1-50%. This heterogeneity seems to indicate that there are two primary and highly differentiated ways to be a small retailer in Lusaka; one involves building and maintaining relationships with specific suppliers, while the other involves explicitly avoiding relationships and sourcing inventory on the market (presumably by searching for products based on quality, price, and availability). One caveat to this finding is that we re-asked this question in our intervention delivery survey (which concluded on March 13), and the correlation between the two responses is approximately 0.2. Why this between-round correlation is not higher is something we are still puzzling over and something that we plan to address in future surveys.

A natural question arising from Figure 3 is whether the firms that source primarily from regular suppliers are different on other dimensions—age, sector, size, and others—from the firms that source primarily from the open market. We address this question by regressing various firm characteristics on a dummy variable that takes a value of 1 if the firm is above the median in the share of recent stock acquired from regular suppliers (the median is approximately 0.75). Results are shown in Table 5. There is little evidence of systematic differences between the firms that do and do not use regular suppliers. The only statistically significant coefficient is for the binary dependent variable that takes a value of 1 if the firm keeps receipts from its purchases.

In Table 6 we explore these differences further, by testing the statistical significance of the difference in the probability that each type of firm (above/below median) sells items in a particular category. We find three differences, for fruits and vegetables, phone accessories, and stationary products. Overall, we find little evidence of substantial differences between firms based on whether they are above or below the median in their recent use of regular suppliers.

5.2 Average Treatment Effects from the Lottery RCT

Table 7 shows estimated lottery treatment effects in the five survey rounds. In column 1 the ATE for the High Treatment group is 624 ZWK, representing a 54% increase over the Control mean of 1162 ZWK ($p < 0.01$). The Low treatment effect is smaller in magnitude, at 392 ZWK, but statistically significant ($p < 0.06$). In the top panel of Figure 5 we plot treatment effects by round. ATEs are largely stable over the survey rounds, except for a drop in the Low group effect in the final round.[3]

In column 2 of Table 7 we report ATEs on the VAT share, defined as the total value of VAT invoices retained divided by total expenditure on stocking. The treatment effect for High treatment group is 0.109, representing a 58% increase over the Control mean of 0.187. The ATE for the Low treatment group is much smaller, at 0.032. For the VAT share, estimated treatment effects for the High and Low groups are statistically different. The time-path of effects on the VAT share is similar to that for the VAT value (Figure 5, lower panel).

In Table 8 we report estimates of heterogeneous treatment effects by the strength of supplier relationships at baseline. In columns 1-2, the strength of relationships is defined as the share of recent stock that was purchased from regular suppliers (as opposed to on the open market). In columns 3-4, we measure relationships using an index that aggregates extensive margin responses to questions about services received from suppliers (goods on credit, price discounts, favorable delivery terms, and others).

There are two main takeaways from Table 8. The first is that across all outcomes and relationship definitions, ATEs are larger in magnitude for firms with weaker relationships at baseline. This suggests that firm-to-firm relationships act as a deterrent to tax-related incentives. Second, among control firms, outcomes are systematically larger for firms with strong supplier relationships, regardless of which relationship measure is used. This suggests that firms with stronger supplier ties face lower compliance costs, possibly due to larger order sizes, greater bargaining power, or transactions with more formalized suppliers. While we do not show it here, Appendix A provides evidence that this effect is driven by relational contracting rather than firm size alone.

We also examine whether the financial incentives in the lottery RCT have persistent effects in rounds 6-8, after the removal of the incentives. In columns 3 and 4 of Table 9, we see that the effects of the High Treatment do persist, and remain substantial in magnitude. The Low treatment ATE is not distinguishable from zero once the possibility of winning cash is removed. In Figure 7 we see that outcomes for the Low group converge to the Control group soon after the end of the lotteries.

[3] That drop is driven by a large increase in VAT invoice retention by the Control group, not a reduction in VAT invoice retention by the Low Treatment firms. See Figure 7.

5.3 Average Treatment Effects from the Messaging RCT

In columns 1-2 of Table 9 we report the ATEs of the messaging treatment over follow-up rounds 6-8. Recall that study firms were re-randomized after the end of the Lottery RCT, stratifying on lottery treatment status, so these findings are not contaminated by lingering effects of the lotteries. The ATE on the VAT value is 233, which represents a statistically significant increase of 27% over the Control mean. We find no effect of messaging on the VAT share. In Figure 6 it is clear that the positive effect of the Messaging treatment is due entirely to the first month after delivery of the message.

Table 10 reports estimates of specification (3). The goal of these regressions is to test for habit formation, and to examine whether messaging influences habit formation. In columns 1-2 we see the OLS and IV estimates of total VAT value from rounds 1-5 on VAT values in rounds 6-8. The OLS estimates provide evidence of both habit formation (the coefficient of 38.1 on the round 1-5 VAT value), and suggest an additive role for messaging (coefficient = 15.13). However, when we use Lottery treatment assignment to isolate the exogenous variation in VAT values over rounds 1-5, we see that messaging has no effect, and habituation is even stronger than it appeared.

5.4 Supplier Reaction

Figure 7 reports estimates of specification (4). It presents event study estimates of VAT remittances in million Kwacha, where treatment effects are measured by two metrics: the value of invoices issued relative to firm average sales (top panel) and the total number of invoices issued (bottom panel). The coefficients represent the dynamic treatment effects over the quarters before and after the intervention, with $t = 0$ denoting the treatment quarter (2023 Q1).

First, both panels show a clear increase in VAT remittances after the intervention ($t \geq 0$), with coefficients becoming positive and significant two quarters after the treatment. The magnitude of the treatment effects varies depending on how the treatment is measured. For the first treatment, where the value of invoices is normalized by firm average sales, a one standard deviation increase in treatment strength is associated with a 13.8% increase in VAT remittances. For the second treatment, which measures the total count of invoices issued, a one standard deviation increase in treatment strength corresponds to a 24% increase in VAT remittances.

These results provide compelling evidence that supplier VAT firms responded to the intervention, reporting more VAT remittances as shown in the administrative data. This reaction is particularly surprising given the small size of the retailer firms in our sample, which might typically lack the leverage to compel suppliers to formalize transactions. The findings suggest that even small, informal retailers can influence the compliance behavior of larger VAT-registered suppliers when properly incentivized. This indicates that engaging retailer firms as intermediaries holds promise as a policy tool to improve VAT compliance, potentially addressing a key point of VAT leakage in supply chains.

6 Conclusion

This paper investigates how tax policy interventions influence the behavior of small retailer firms and their VAT-registered suppliers in a low-income country context. Using two sequential randomized controlled trials (RCTs) conducted in Lusaka, Zambia, we provide evidence that targeted incentives for retailer firms can significantly impact the retail firm's invoice requesting and retaining behaviors, and the VAT compliance and remittance behavior in supply chains.

The Lottery RCT showed that financial incentives, both high and low, effectively increased VAT invoice retention by retailer firms. High treatment firms retained VAT invoices worth 54% more than the control group, while the low treatment group achieved a 34% increase. Importantly, firms with weaker baseline relationships with suppliers exhibited larger treatment effects, suggesting that relational contracts often deter the adoption of tax-compliant practices. This finding underscores the dual role of supplier relationships: while they provide operational benefits like credit and discounts, they can also perpetuate informal and tax-avoiding behaviors.

The Messaging RCT extended these insights by examining the persistence of invoice-requesting behavior and the role of tax morale. While the high financial incentives from the Lottery RCT continued to influence behavior after their removal, tax morale messaging had only modest and short-lived effects. These results highlight the challenges of converting extrinsic incentives into intrinsic motivations for tax compliance. Nevertheless, the enduring impact of financial incentives suggests that sustained, targeted interventions can foster habit formation, even in contexts characterized by high informality.

Finally, we find that supplier VAT-registered firms not only responded to retailer interventions by issuing more formal VAT invoices but also increased their tax reporting to the government. This is a surprising yet promising outcome, given the small size of the retailer firms in our sample. The finding demonstrates the potential for indirect policy mechanisms to influence behavior upstream in the supply chain.

A notable contribution of this study is its focus on the potential of small retailer firms as agents of change in improving VAT compliance. By incentivizing small retailers with fixed locations, tax authorities can keep administrative cost manageable, yet indirectly influence larger VAT-registered suppliers, addressing a critical source of VAT leakage in developing economies. The results point to the promise of scaling such interventions as part of broader tax reform efforts, particularly in settings where administrative capacity is constrained. We hope this study advances our understanding of firm-to-firm relationships in the context of tax compliance—shedding light on how these relationships shape non-VAT and VAT firm's incentives to request and issue VAT invoices and highlights the untapped potential of such dynamics to foster greater formalization and tax compliance within supply chains.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge, "When should you adjust standard errors for clustering?," *The Quarterly Journal of Economics*, 2023, *138* (1), 1–35.

Alm, James, Yongzheng Liu, and Kewei Zhang, "Financial constraints and firm tax evasion," *International Tax and Public Finance*, 2019, *26*, 71–102.

Baker, George, Robert Gibbons, and Kevin J Murphy, "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics*, 2002, *117* (1), 39–84.

Carrillo, Paul E, Edgar Castro, and Carlos Scartascini, "Do rewards work?: Evidence from the randomization of public works," 2017.

Fafchamps, Marcel, "Development and social capital," *The Journal of Development Studies*, 2006, *42* (7), 1180–1198.

Gérard, François and Joana Naritomi, "Value Added Tax in developing countries: Lessons from recent research," *IGC Growth Brief Series*, 2018, *15*.

Ghani, Tarek and Tristan Reed, "Relationships, Risk and Rents: Evidence from a Market for Ice," 2017.

Hoy, Christopher, Luke McKenzie, and Mathias Sinning, "Improving tax compliance without increasing revenue: Evidence from population-wide randomized controlled trials in Papua New Guinea," *Economic Development and Cultural Change*, 2024, *72* (2), 000–000.

Jensen, Anders, "Evidence-Based Insights to Build Tax Capacity in Developing Countries: A Critical Review of the Literature," 2019.

_ , "Employment structure and the rise of the modern tax system," *American Economic Review*, 2022, *112* (1), 213–234.

Keen, Michael, Russell Krelove, and John Norregaard, "The financial activities tax," *Can. Tax J.*, 2016, *64*, 389.

Khan, Adnan Q, Asim Ijaz Khwaja, and Benjamin A Olken, "Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings," *American Economic Review*, 2019, *109* (1), 237–270.

Kleven, Henrik Jacobsen, Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez, "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica*, 2011, *79* (3), 651–692.

Luttmer, Erzo FP and Monica Singhal, "Tax morale," *Journal of economic perspectives*, 2014, *28* (4), 149–168.

Macchiavello, Rocco, "Relational contracts and development," *Annual Review of Economics*, 2022, *14*, 337–362.

_ and Ameet Morjaria, "The value of relationships: evidence from a supply shock to Kenyan rose exports," *American Economic Review*, 2015, *105* (9), 2911–2945.

_ and _ , "Competition and relational contracts in the Rwanda coffee chain," *The Quarterly Journal of Economics*, 2021, *136* (2), 1089–1143.

Mascagni, Giulia and Christopher Nell, "Tax compliance in Rwanda: Evidence from a message field experiment," *Economic Development and Cultural Change*, 2022, *70* (2), 587–623.

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

Pomeranz, Dina, "No taxation without information: Deterrence and self-enforcement in the value added tax," *American Economic Review*, 2015, *105* (8), 2539–2569.

Rudder, Jessica and Brian Dillon, "Search Costs and Relational Contracting: The Impact of a Digital Phonebook on Small Business Supply Chains," *Working paper*, 2023.

Slemrod, Joel, Obeid Ur Rehman, and Mazhar Waseem, "How do taxpayers respond to public disclosure and social recognition programs? Evidence from Pakistan," *Review of Economics and Statistics*, 2022, *104* (1), 116–132.

Waseem, Mazhar, "Overclaimed refunds, undeclared sales, and invoice mills: Nature and extent of noncompliance in a value-added tax," *Journal of Public Economics*, 2023, *218*, 104783.

Zambia Statistics Agency, "Average Household Income by Rural/Urban Zambia - 2022 Living Conditions Monitoring Survey," 2022. [Accessed March 18, 2025].

7 Figures

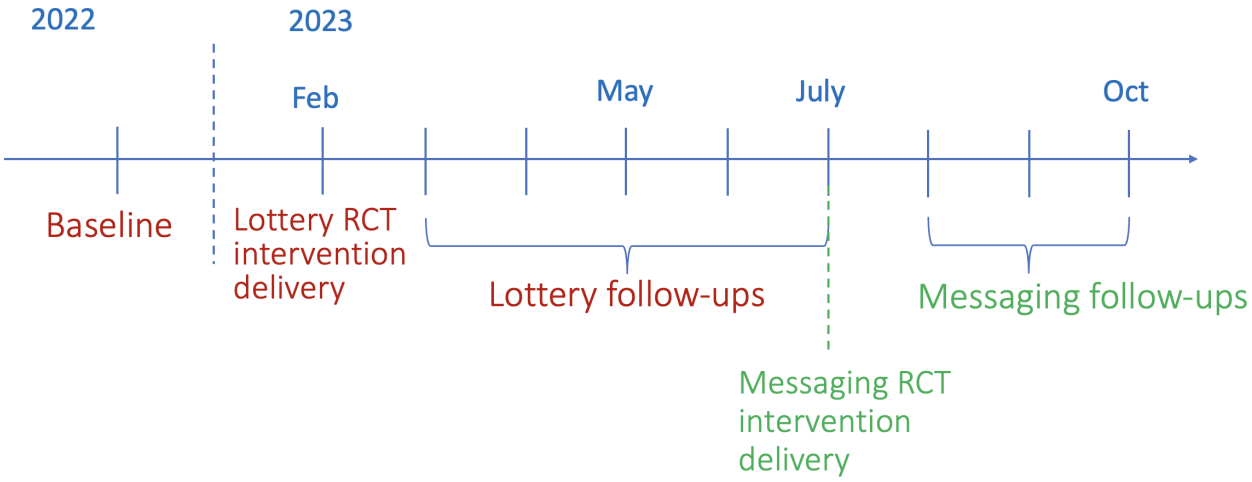

Figure 1: Survey and Intervention Timeline

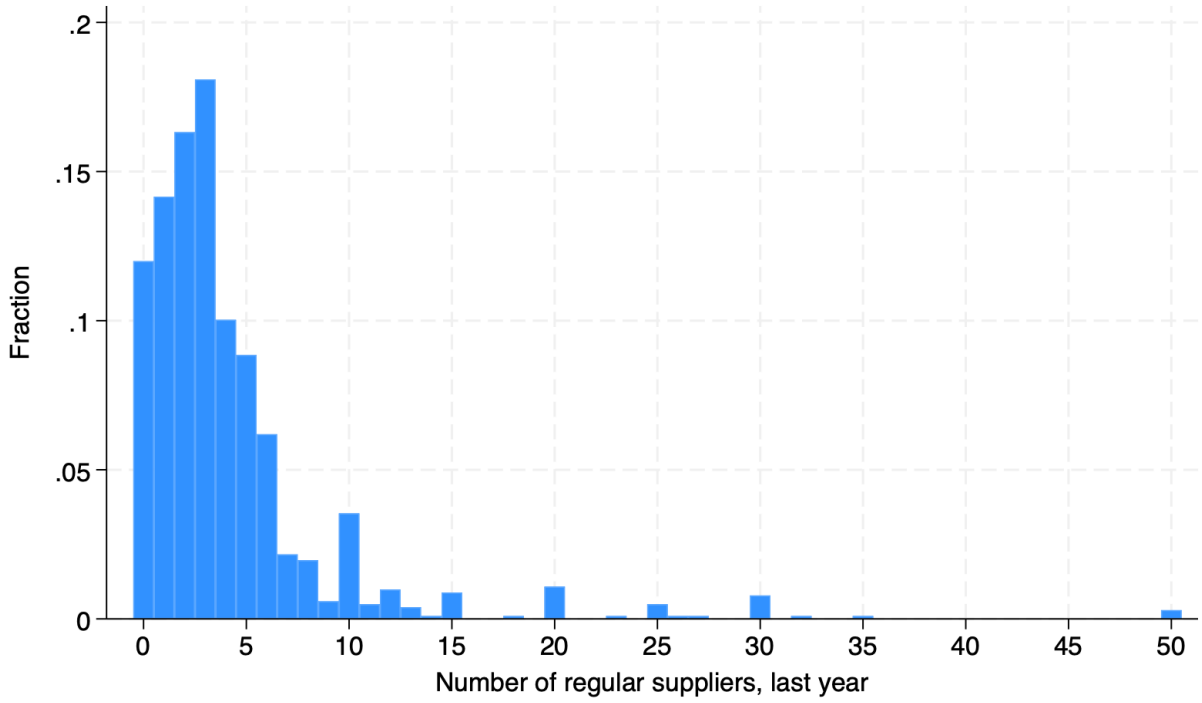

Figure 2: Number of regular suppliers, last year

Notes: Authors' calculations from baseline survey with 1,083 retail firms in Lusaka. A regular supplier is defined as a wholesaler that the retailer firm has purchased from at least twice in the recent past, and that they plan to purchase from again at some future point.

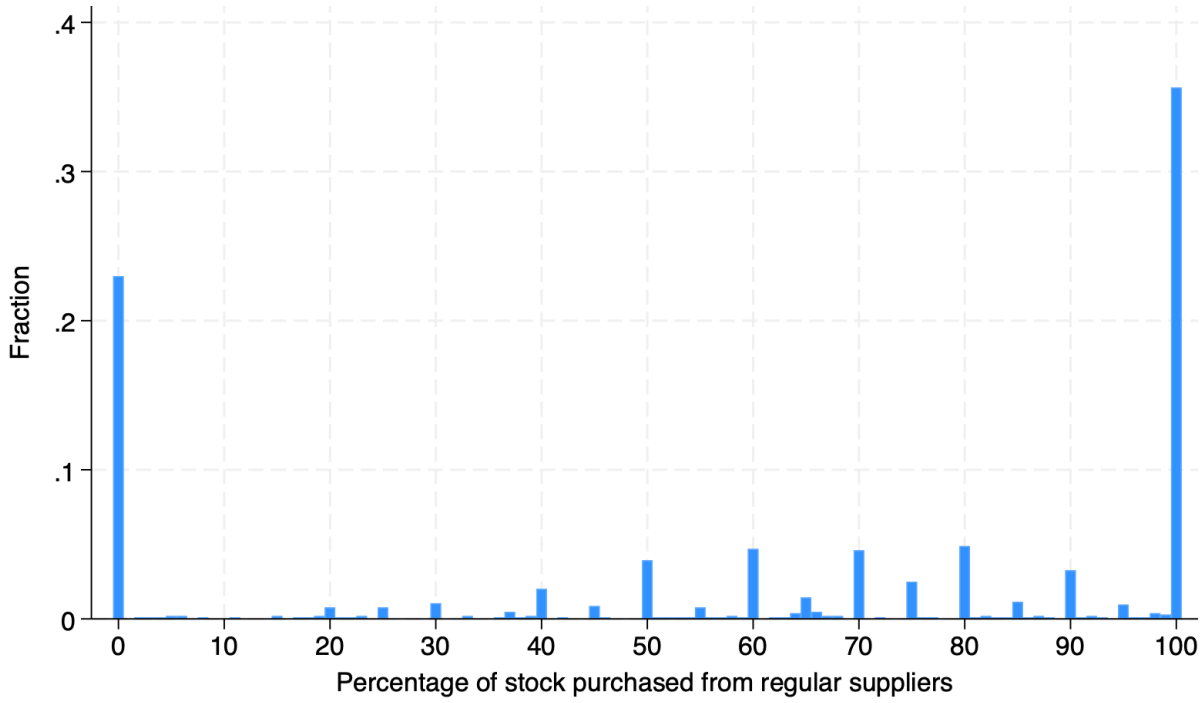

Figure 3: Percentage of stock purchased from regular suppliers

Notes: Authors' calculations from baseline survey with 1,083 retail firms in Lusaka. A regular supplier is defined as a wholesaler that the retailer firm has purchased from at least twice in the recent past, and that they plan to purchase from again at some future point.

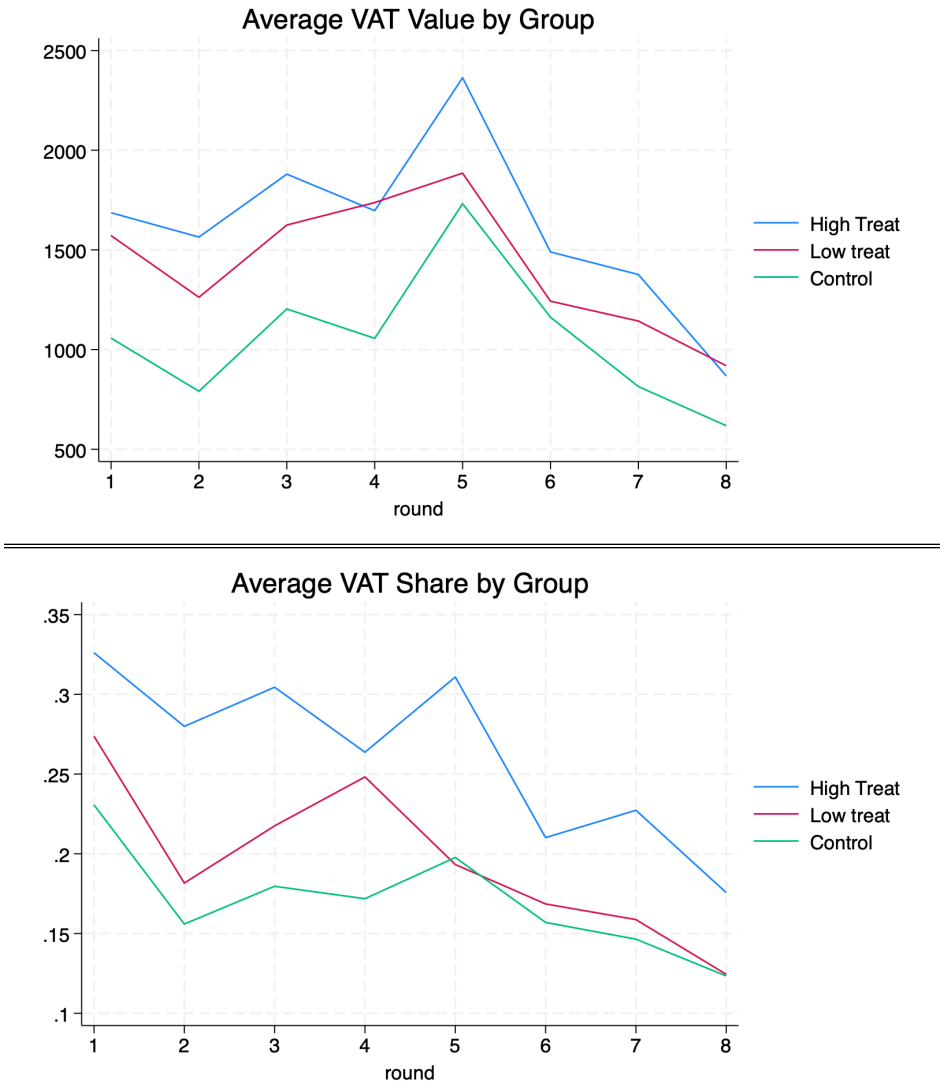

Figure 4: Average VAT Value and Share by Round

Notes: Authors' calculations from original survey data. VAT value is the total value of purchases reflected on valid VAT invoices retained by study firms. VAT share is the share of VAT value in total stocking expenditure. Each round is approximately one month long, running from March-October 2023. High Treat, Low Treat, and Control refer to treatment assignment in the lottery RCT, which was active from rounds 1-5.

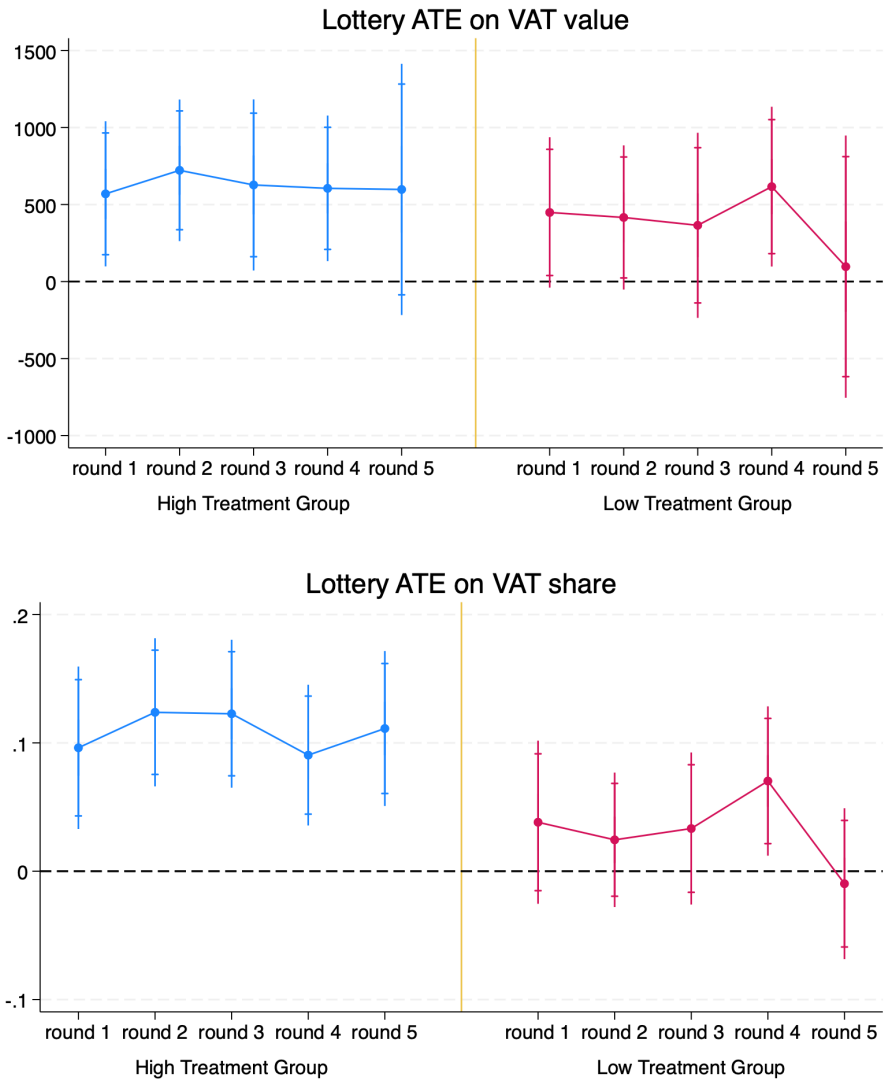

Figure 5: Lottery RCT: Average Treatment Effect by Round

Notes: Average treatment effect plotted separately for the five rounds of follow-up survey. Vertical bars show 95% (full bar) and 90% (horizontal marker) confidence intervals. Upper and lower figure respectively show the effect for two primary outcome variables, the total value of VAT invoices retained by the firm from inventory and capital improvement purchases, and the share of VAT-invoiced purchases as a percentage of all such expenditures. Standard errors are clustered at firm level. High Treatment and Low Treatment refer to assignment in the Lottery RCT.

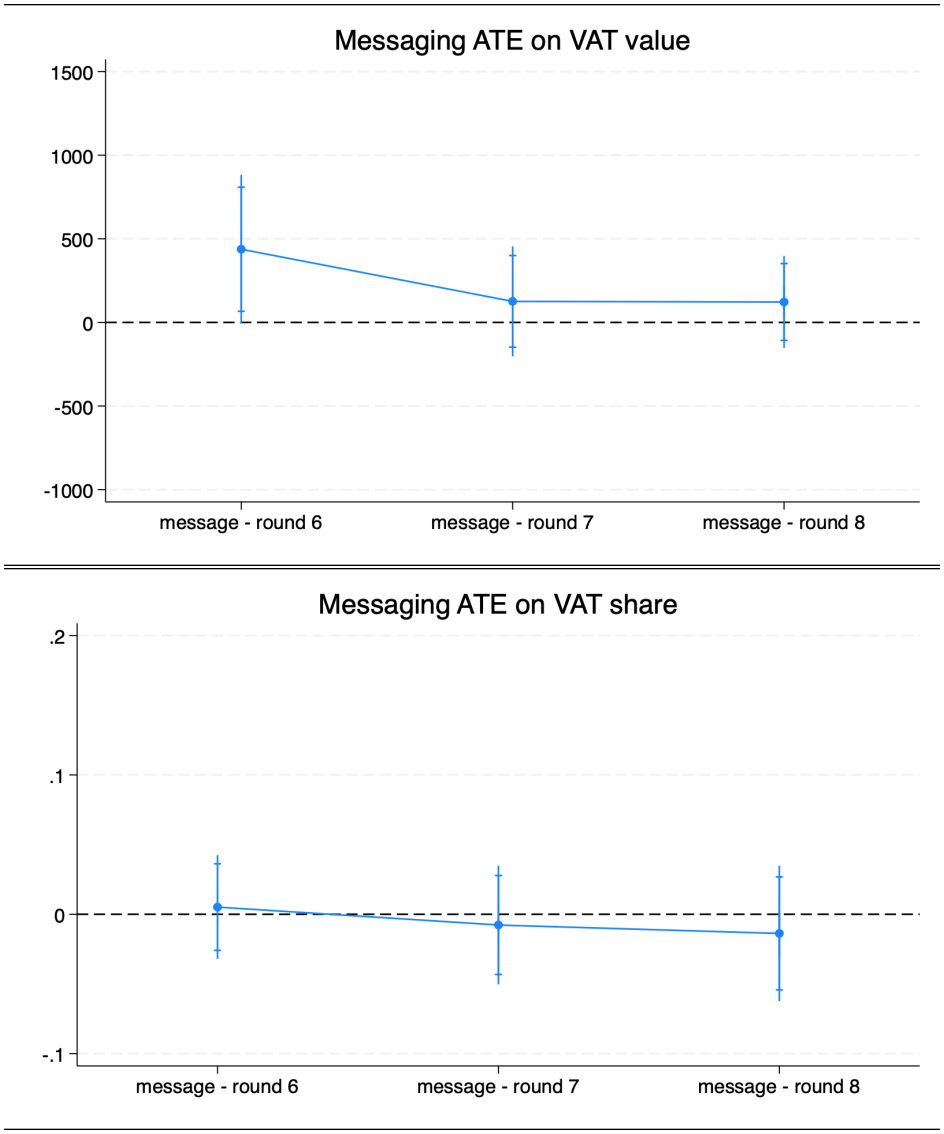

Figure 6: Messaging RCT: Average Treatment Effect by Round

Notes: Figure shows the average treatment effect of the message treatment, plotted separately for the three rounds of survey, after five rounds of follow-up survey. Point estimates, statistical significance level (*** p<0.01, ** p<0.05, * p<0.1), 95% and 90% confidence intervals are reported in the figure. Upper and lower figure respectively show the effect for two primary outcome variables, the total value of VAT invoices retained by the firm from inventory and capital improvement purchases, and the share of VAT-invoiced purchases as a percentage of all such expenditures. Standard errors are clustered at firm level.

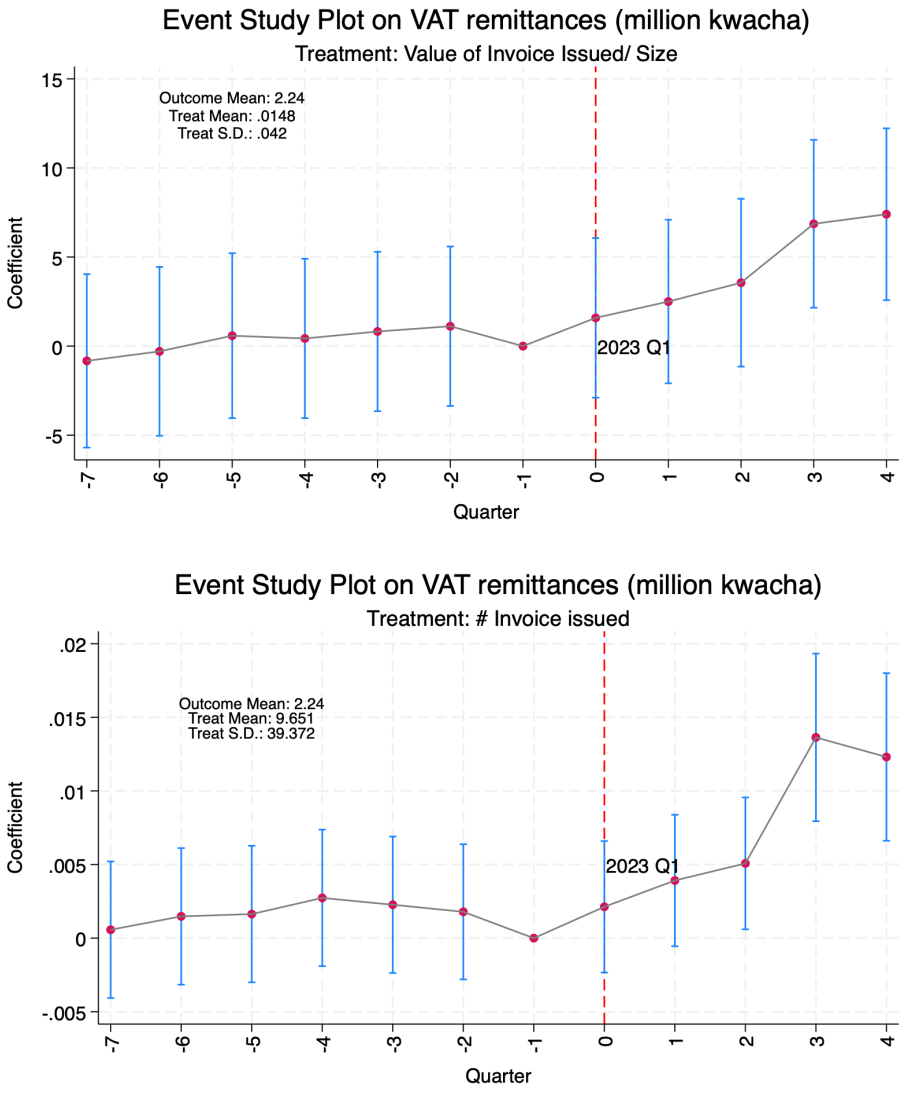

Figure 7: Event Study of Supplier Response

Notes: Two treatment variables are: the value of invoices issued normalized by average firm sales, and the number of invoices issued to our sample firm.

8 Tables

Table 1: Lottery RCT, number of draws for treated firms

Total kwacha value of purchases with valid VAT invoices	Number of draws
0	0
1-1500	1
1500-3000	2
3000+	3

Table 2: Lusaka Markets Covered at Baseline

bauleni	kalingalinga	matero
chainda	kamanga	misisi
chaisa	kamwala	mukuyu
chalala hillview	kamwala south	munyaule
chawama	kangwa	mutendere
chelstone	katambalala	mutendere east
chilenje	katungu/SDA	mutonyo
chingwere	kaunda square	new soweto
chipata	kuku	ngombe
emmasdale	kulima tower	northmead/olympia
garden	lilayi	nyumbayanga
helen kaunda	longacres	obama last stop
jack	lumumba	sekelela/apollo
john howard	lungu	soweto
john lenge	mandevu	town center
kabwata		

Table 3: Firm Characteristics and Share of Inventory from Regular Suppliers

	Mean	s.d.	N
Year firm opened	2016.3	6.2	1043
Number of full-time workers	1.7	1.4	1056
Number of part-time workers	1.3	1.2	1055
Number of unpaid workers	0.6	0.9	1053
Number of customers, typical weekday	43.1	75.8	902
Number of customers, typical weekend	49.3	84.9	881
Keeps formal books (=1)	0.53	0.50	1045
Keeps business receipts (=1)	0.31	0.46	1045
Total stocking expenditure (last month)	9547	11419	904
Value of VAT invoices for stocking (last month)	1207	2869	766

Notes: Authors' calculations from baseline survey with 1,083 retail firms in Lusaka.

Table 4: Items Sold by Study Firms (non-exclusive)

	Share	N
Drinks	0.39	1058
Cosmetics	0.36	1058
Cooking oil, sauces, and spices	0.35	1058
Snacks	0.35	1058
Dry food, including mealie meal, rice, soya pieces, and beans	0.34	1058
Sweets	0.27	1058
Hardware	0.25	1058
Household cleaning products	0.22	1058
Other (specify)	0.17	1058
Phone accessories	0.17	1058
Stationary products	0.14	1058
Meat and fish	0.09	1058
Home goods, including plates, cups, utensils, and cooking vessels	0.04	1058
Prepared food	0.04	1058
Fruits and vegetables	0.03	1058
Alcohol	0.03	1058

Notes: Authors' calculations from baseline survey with 1,083 retail firms in Lusaka.

Table 5: Firm Characteristics and Share of Inventory from Regular Suppliers

Variable	Coeff	s.e.	p-value	Control mean
Year firm opened	-0.53	0.39	0.17	2016.53
Number of full-time workers	-0.01	0.09	0.90	1.73
Number of part-time workers	-0.03	0.07	0.67	1.29
Number of unpaid workers	0.03	0.06	0.58	0.59
Total employees, incl. unpaid	0.00	0.14	0.99	3.59
Number of customers, typical weekday	-5.68	5.14	0.27	46.43
Number of customers, typical weekend	-7.78	5.85	0.18	54.01
Keeps formal books (=1)	0.04	0.03	0.16	0.51
Keeps business receipts (=1)	0.09	0.03	0.00	0.27
Total stocking expenditure (last month)	532.36	775.54	0.49	9256.62
Value of VAT invoices for stocking (last month)	276.37	205.12	0.18	1055.72
Num. item categories sold	-0.30	0.17	0.08	3.44

Notes: Authors' calculations from baseline survey with 1,083 retail firms in Lusaka. This table reports the coefficients, standard errors, p-values, and control group mean for regressions of each variable on a dummy for the Above median share of stock purchased from regular suppliers.

Table 6: Differences in items sold: above and below median use of regular suppliers

	Below median	Above median	p-value
Dry food, including mealie meal, rice, soya pieces, and beans	0.35	0.34	0.66
Drinks	0.41	0.38	0.30
Alcohol	0.04	0.03	0.35
Snacks	0.37	0.33	0.25
Sweets	0.29	0.27	0.45
Cooking oil, sauces, and spices	0.37	0.34	0.30
Fruits and vegetables	0.03	0.04	0.80
Meat and fish	0.08	0.10	0.34
Prepared food	0.05	0.03	0.09
Cosmetics	0.39	0.35	0.18
Household cleaning products	0.25	0.19	0.02
Phone accessories	0.18	0.16	0.54
Stationary products	0.18	0.12	0.00
Hardware	0.24	0.25	0.74
Home goods, including plates, cups, utensils, and cooking vessels	0.03	0.04	0.38
Other (specify)	0.18	0.17	0.84

Notes: Authors' calculations from baseline survey with 1,083 retail firms in Lusaka. This table reports the difference in the probability that each type of firm (above/below median share of stock purchased from regular suppliers) sells items in a particular category.

Table 7: Average treatment effect of lottery incentive on VAT invoice request and retention

	(1)	(2)
	VAT value	VAT value
High Treat	624.2***	0.109***
	(184.3)	(0.0181)
Low Treat	392.4*	0.0318*
	(203.0)	(0.0182)
Constant	453.3*	0.167***
	(233.6)	(0.0441)
N	3962	3402
R-squared	0.0573	0.0671
Control Mean	1161.7	.187
Round FE	YES	YES
Strata FE	YES	YES

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Average treatment effect for all five rounds of follow-up survey. Column 1 and 2 respectively show the effect for two primary outcome variables, the total value of VAT invoices retained by the firm from inventory and capital improvement purchases, and the share of VAT-invoiced purchases as a percentage of all such expenditures. Standard errors are clustered at firm level.

Table 8: Heterogenous effect by relationship

	(1)	(2)	(3)	(4)	(5)	(6)
	Stocking		Relationship Index		Duration	
	VAT value	VAT share	VAT value	VAT share	VAT value	VAT share
Treat X Strong	382.4	0.0631***	460.8**	0.0673***	304.9	0.0606*
	(265.5)	(0.0240)	(207.3)	(0.0193)	(471.1)	(0.0367)
Treat X Weak	608.3***	0.0767***	738.7*	0.0779**	429.1**	0.0826***
	(177.2)	(0.0201)	(419.9)	(0.0391)	(195.2)	(0.0222)
Strong	1457.5***	0.0836**	73.85	0.00148	573.2	0.0394
	(403.5)	(0.0326)	(370.7)	(0.0362)	(422.5)	(0.0342)
Constant	416.9*	0.139***	496.7	0.131**	238.3	0.102
	(247.7)	(0.0384)	(419.7)	(0.0604)	(434.8)	(0.0625)
N	3873	3326	3349	2877	2780	2376
R-squared	0.0664	0.0648	0.0677	0.0673	0.0681	0.0775
Control Mean	1161.7	.187	1161.7	.187	1161.7	.187
Round FE	YES	YES	YES	YES	YES	YES
Strata FE	YES	YES	YES	YES	YES	YES

Notes: Authors use two measures of relationship. The first definition is the percentage of the stocking sourced from regular suppliers. The second definition is the index that captures how much benefit (delivery, discount, gift, trade credit) firms get from the supplier firms. Columns 1 and 2 show the results on the regular supplier definition. Columns 3 and 4 show the results on the relational benefit definition. And columns 5 and 6 show the results on duration of the relationship with key suppliers. Standard errors are clustered at firm level.

Table 9: Messaging and Lottery ATE on VAT Value and Share at Roun 6-8

	(1)	(2)	(3)	(4)
	VAT value	VAT share	VAT value	VAT share
Message treat	232.9**	-0.00511		
	(115.4)	(0.0131)		
High treat			356.8*	0.0642***
			(180.8)	(0.0213)
Low treat			204.8	0.00501
			(190.0)	(0.0200)
Constant	1046.0***	0.135***	946.2***	0.110***
	(67.50)	(0.0103)	(122.5)	(0.0124)
N	2101	1918	2101	1918
R-squared	0.0535	0.0450	0.0549	0.0551
Control Mean	875.2	.143	875.2	.143
Round FE	YES	YES	YES	YES
Strata FE	YES	YES	YES	YES

Notes: Average treatment effect of messaging and treatment assignment for messaging treatment rounds 6-8. Column 1 and 2 respectively show the effect for messaging treatment variables, Column 3 and 4 respectively show the effect for treatment assignment. The total value of VAT invoices retained by the firm from inventory and capital improvement purchases, and the share of VAT-invoiced purchases as a percentage of all such expenditures. Standard errors are clustered at firm level.

Table 10: Habit Formation

	(1)	(2)	(3)	(4)
	VAT value		VAT share	
	OLS	IV	OLS	IV
value (round 1-5)	38.10***	83.52**		
	(4.709)	(36.60)		
value (round 1-5) X message	15.13***	2.205		
	(5.409)	(13.92)		
share (round 1-5)			0.417***	0.642***
			(0.0324)	(0.142)
share (round 1-5) X message			-0.0281	-0.0341
			(0.0382)	(0.0548)
Constant	644.2***	291.2	0.0760***	0.0260
	(57.39)	(296.0)	(0.00861)	(0.0322)
N	2101	2101	1918	1918

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: First stage F-statistics from the F test of excluded instruments for the the explanatory variables are: value (5.78), value-message (59.38), share (14.73), share-message (203.72). The mean and s.d. of VAT value (round 1-5) are 1523 and 3498, The mean and s.d. of VAT share (round 1-5) are 0.24 and 0,.33