Tax audits and their economic distortions*

David Henning (UCLA) Joseph Okello (URA)

October 22, 2024

Find latest version here

Preliminary and incomplete. Please do not circulate or cite.

Abstract

Tax audits are essential for governments to raise revenue, but can create economic distortions. Using detailed administrative data from the Ugandan Revenue Authority and a regression discontinuity approach we show that receiving a comprehensive tax audit – the most intense tax enforcement measure – reduces the tax liability of firms in a developing country, Uganda. A key challenge is to disentangle whether such changes are due to changes in tax filing behavior (evasion) or changes in firm operations (real responses). We overcome this by combining the administrative data with a novel survey that we designed. Using the survey, we show that the reduction in revenue is driven by changes in firm operations, not only adjustments in what is reported to the tax authorities. The comprehensive audit causes 16% of audited firms to shut down. Further survey evidence supports the argument that the primary reason for shutting down is the audit. Conditional on remaining operational, firms reduce their size after the audit. Taken together, this indicates strong (negative) effects on marginally audited firms. Overall the results suggest that audits impose large costs on audited firms in our context and ultimately hurt both revenue collection efforts and the real economy.

*Contact information: djhenning@ucla.edu, jookello@ura.go.ug. David is heavily indebted to Natalie Bau, Adriana Lleras-Muney, Juliana Londoño-Vélez, and Manisha Shah for providing invaluable guidance and support throughout the project. We are grateful to staff at the Ugandan Revenue Authority for facilitating the project and providing data access. Special thanks to the Central Operations Office for their comments and insights throughout. We are indebted to Allan Gitta and Apollo Kamwebaze for helping us understand the audit selection process. We are also grateful to staff at the Ugandan Registry Services Bureau for providing registry information. The tracking and survey was done with the help of International Growth Research and Evaluation Center (IGREC). We gratefully acknowledge funding from the International Growth Center (IGC), Private Enterprise Development in Low-Income Countries (PEDL), Center for Effective Global Action (CEGA), and Structural Transformation for Economic Growth (STEG). All errors are our own.

1 Introduction

Formal firms are key for economic development and a pillar of tax collection efforts (Slemrod, 2019). While raising revenue is essential for economic growth (Besley and Persson, 2013), in a context where firms are small, statutory tax rates are high, and the informal sector is vast it is not evident that enforcement efforts targeted at formal firms – such as tax audits – would raise revenue. Specifically, one might be worried that tax audits induce firms to shut down, move to the informal sector or reduce their output to avoid future audits. All of these adjustments would cause firms to stay inefficiently small which creates misallocation and, ultimately, reduces aggregate output (Bachas et al., 2019; Hsieh and Klenow, 2009). Despite the potential negative impact of tax enforcement on firm output and economic development, it has received little attention in the literature (A. Jensen et al., 2024).

The combination of these factors are likely to be particularly acute in developing countries for several reasons. First, formal firms are on average smaller than their counterparts in rich countries (Bento and Restuccia, 2021). More limited accounting and auditing capacity due to lower size would make them more vulnerable to intense enforcement measures. Second, the informal sector tends to be significantly larger, offering firms an alternative operating environment where they avoid all taxes and regulatory costs (Ulyssea, 2018). Finally, statutory tax rates are higher (Enache, 2023). To the extent that these are enforced it creates a larger incentive for adjusting output to avoid scrutiny from the tax authority.

In this paper we investigate how receiving a comprehensive audit from the tax authority – the most intense tax enforcement measure – affects ex-ante formal firms in a developing country, Uganda. Comprehensive audits are extensive audit interventions. The median audit takes 3 months, and can investigate any part of a firms' accounts that relate to the past five years. We combine detailed administrative data with a survey and discontinuities to document three new findings. First, tax liabilities of the audited firms decline post-audit, driven by firms exiting the tax system and remaining firms reporting lower tax liabilities. A marginal value of public fund calculation in the style of Boning et al. (2023) would yield a negative value. Second, we track down firms (physically) two years after the audit to show that firms receiving comprehensive audits are more likely to shut down. Third, survey evidence from the firms that remained operational – in either the formal or informal sector – suggests that firms reduce their output after the audit.

Receiving a tax audit can affect firm operations in several ways. First, the audit induces an immediate cost to the firm due to the hassle cost of dealing with the audit and – if found liable – paying back taxes and associated penalties. If the cost to the firm is substantial, firms that are sufficiently liquidity constrained could be pushed towards insolvency. Second, the audit may cause the firm to change its perception about the enforcement environment. An audit may raise the belief that the authority can detect evasion. While this is usually considered "successful deterrence", an upward correction of such a belief increases the expected effective tax rate of the firm. Firms on the margin

might now expect to be unprofitable in the future. Finally, firms might also adjust their belief on the extent to which audit rates depend on firm output. An upwards adjustment in that belief could induce firms to lower their output.

Descriptive evidence from our survey supports the notion that taxation and tax enforcement imposes a large economic burden for firms. Overall costs of taxation ranges from 26-36% of *revenue*, with firms in the lowest quintile of sales reporting the highest shares. This is not because small firms pay more taxes, instead, the enforcement costs represents a larger share of revenue, suggesting that the relative cost of tax enforcement is particularly high for small firms. Furthermore, 63% of firms mention high taxes, tax administration costs or tax corrections as one of their three main obstacles to their performance. This is the most frequently mentioned reason, with the second most important reason – competition – mentioned by only 32% of firms.

Disentangling the impact of a comprehensive tax audit is notoriously difficult because firms selected for such interventions are different from other firms in both observable and unobservable ways. We make progress on this challenge by leveraging discontinuities in the audit selection process of the Ugandan Revenue Authority (the Ugandan IRS, hereafter referred to as the URA). Firms are selected into comprehensive audits based on risk scores that are calculated for the universe of tax filing firms annually. The firms with the highest risk scores are assigned to comprehensive audits, creating a discontinuity in the intensity of enforcement with very similar firms just above and below the threshold. Importantly, the firms below the comprehensive audit threshold receive light audits (not no enforcement). Our results should thus be interpreted as the effect of changing the audit assignment of one more firm from light to comprehensive.

Combining information on tax audits and audit selection criteria with the universe of Corporate Income Tax (CIT) and Value Added Tax (VAT) filings, we find consistent negative effects on tax liabilities. Firms receiving a comprehensive audit are 30 p.p. less likely to file a CIT return in the year after the audit, and the effect persists into – at least – the second post-audit year. This is not driven by "fictitious", firms that only exist to create and sell fake invoices.[1] This leads to an overall reduction in CIT revenue collected of 5,463 USD per firm audited. We find similar, though less precise, results for the VAT, probably because the VAT data only covers a short time period after the audit. While there is an increase in revenue through correcting firms past discrepancies in their tax returns – highlighting one reason why these audits might still be conducted – when considering post-audit reporting behavior, the overall effect of the marginal audit on revenue is unambiguously negative.

A compliers analysis reveals that the firms that exit the tax system tend to be high productivity *marginally formal* firms prior to the audit. The firms that exit are smaller service sector firms, with few employees and a low value of machines and plants. However, they submit substantial tax liabilities prior to the audit and have a high productivity, exactly the kind of firms that one would

[1] See Carrillo et al. (2023); Waseem (2023) for recent papers exploring this phenomenon in developing countries

want to keep in the economy and the tax system. The lack of extensive capital investments suggests that it is easier for these types of firms to relocate or shut down when under pressure from tax authorities. While this explains why they respond by leaving despite being productive, these are exactly the type of firms needed to improve future employment opportunities and growth (Quinn and Woodruff, 2019).

While the administrative data is rich in detail, it is challenging to measure the impact of receiving a comprehensive audit on changes in firm output using administrative tax data. Information in tax returns submitted post-audit might be influenced both by real production decisions of the firm and changes in reporting behavior. We overcome this challenge by designing and implementing a firm survey for the firms on the margin of receiving a comprehensive audit. The survey attempts to track down all the 839 firms that identify the effect of an audit in the administrative tax data. We implement the survey approximately two years after the audit, and are able to verify whether a firm closed or remained open for 91% of firms.

Leveraging information from our novel survey we find strong evidence that comprehensive audits cause firms to shut down. Our reduced form estimates show that firms on the margin of receiving a comprehensive tax audit are 11 p.p. more likely to shut down. Dividing the coefficient by the change in the probability of filing taxes in the two years after the audit suggests that slightly more than half of the firms that stop filing taxes (54%) shut down post-audit. The remaining 46% kept operating in the informal sector. 43% of closed firms mention that they closed because of challenges with high taxes and/or the URA. Crucially, less than half of respondents that shut down opened another firm, underlining that the exit is long-lasting.

We also find evidence that the firms that remain operational reduce their output after receiving a comprehensive audit. To improve our power for the smaller sample where we have survey data, we move away from the firms in the immediate surrounding of the discontinuity. Combining the survey and administrative data in a difference-in-difference strategy, we find that comprehensive audits cause firms that remain operational to reduce their sales by as much as 56%. This intensive distortion effect would not have been possible to discern from the administrative tax data alone, because the firms either do not file a tax return or report no sales in their tax return. We corroborate that the result is qualitatively similar when using the regression discontinuity design. Taken together, this strongly suggests that comprehensive audits both reduce the number of firms and reduce the output of the remaining firms.

This paper contributes to several strands of the literature. First, it contributes to the literature on the economic (efficiency) costs of taxation and tax enforcement. Previous literature has documented that the VAT distorts firm-to-firm trade (Gadenne et al., 2022) and that VAT rebates affect the export performance of firms (Chandra and Long, 2013). More broadly, firms actively respond to thresholds that change their tax liability (Best et al., 2015; Harju et al., 2019; Liu et al., 2021) or enforcement levels (Almunia et al., 2022) and respond to enforcement interventions by appearing

small (Carrillo et al., 2017) or stop filing taxes (Belnap et al., 2024). A common theme in this literature is that it relies on tax administrative data. While rich in detail and ideal for studying compliance responses it is unclear how much of a firms response is driven by changes in reporting (evasion) or changes in firm sales (real output). We make progress on this by linking administrative tax data with a survey on firms. The survey is a first in allowing us to measure the output of firms while accounting for changes in evasion behavior. The closest paper to ours is Harju et al. (2024) who document that bankruptcy rates increase after a tax audit in Finland. Unlike them, we also document that firms respond in less extreme ways, such as reducing output, which is crucial to understand the overall implication on aggregate output.

Second, we contribute to a long literature on the revenue and welfare impacts of tax audits. Previous literature tends to find that tax audits increase revenue collected by the government (Advani et al., 2021; Beer et al., 2020; Best et al., 2021; Boning et al., 2023; Christiansen, 2024; DeBacker et al., 2015; DeBacker et al., 2018; Gemmell and Ratto, 2012; Harju et al., 2024; Kleven et al., 2011; Kotsogiannis et al., 2024; Li et al., 2018; Løyland et al., 2019). Leveraging a marginal value of public funds approach suggests that audits have large positive welfare effects, at least in the US (Boning et al., 2023). We are one of the first to document that tax audits can have a negative effect on the potential revenue collected by the tax authority. A noteable exception is DeBacker et al. (2015) who document that firms in the US become more tax aggressive after an audit.[2] This might be due to two key features of our study: we focus on a developing country, Uganda, and on the most intense enforcement intervention, comprehensive tax audits. Other work on tax audits in developing countries finds zero (Best et al., 2021) or positive (Kotsogiannis et al., 2024) effects. In the case of the former it is because tax audits are simple cross checks, and in the case of the latter they focus on a balanced panel of firms shutting down extensive margin responses. As such, our work adds to previous literature in similar contexts by drawing attention to the importance of the extensive margin when enforcement measures are comprehensive. Our results also speak to a broader literature on tax enforcement in developing countries (Basri et al., 2021; Bergeron et al., 2024), which does not tend to find negative effects of enforcement on potential revenue collection. We suggest the reason for our different results is that we focus on the most intense enforcement measure.

Third, this speaks to a broad literature on one of the key question in development economics: why are there so few large firms? There is an extensive micro literature investigating different reasons why firms may be prevented from growing. Possible explanations include credit constraints (de Mel et al., 2008; McKenzie and Woodruff, 2008), barriers to hiring (Bassi and Nansamba, 2022; Carranza et al., 2022), limited business training (Blattman et al., 2016; de Mel et al., 2014; Field et al., 2010), delegation and organization of firms (Akcigit et al., 2021; Bassi et al., 2023; Bloom et al., 2012) poor contract enforcement (Boehm and Oberfield, 2020; Iyer and Schoar, 2015), high

[2]There is also a large theoretical and experimental literature documenting the possibility that firms and individuals reduce their tax liability in the immediate aftermath of the audit (Alm and Malézieux, 2021; Kasper and Alm, 2022; Kasper and Rablen, 2023)

cost of formalization (McKenzie and Seynabou Sakho, 2010), and constraints on the demand side (Bold et al., 2022; Hjort et al., 2020; R. Jensen and Miller, 2018; Vitali, 2023). However, one of the defining features of developing countries is that they have high statutory tax rates and high enforcement rates relative to firm size, potentially making it one of the key obstacles to firm growth. Tax enforcement has been recognized as a potential reason for misallocation in the structural and macro literature (Bachas et al., 2019; Corbellini, 2024; Di Nola et al., 2021; Leal Ordóñez, 2014). However, to the best of our knowledge, there has been limited causal evidence on the extent to which tax audits – and by extension enforcement – induces firms to reduce their output. We fill that gap.

Fourth, our paper speaks to the distortionary costs of regulation and government enforcement more broadly. There is extensive evidence that firms below regulatory thresholds are over-represented providing evidence of misallocation (Boeri and Jimeno, 2005; Evans, 1986; Garicano et al., 2016; Gourio and Roys, 2014; Schivardi and Torrini, 2008). There is also direct evidence that audits can create distortions in public procurement even in high-capacity low-corruption settings (Gerardino et al., 2024). Our paper adds to the literature by highlighting that tax audits can distort the output of firms, which has implications for aggregate output. In doing so we draw attention to the role of government policy in affecting the size distribution of firms through their choice of tax enforcement.

The rest of the paper is structured as follows. In section 2 we describe the Ugandan economic context, tax system and audit selection process. In section 4 we present the data used in our analysis and provide descriptives for our sample. Section 5 describes our estimation strategy, and ascertains its validity. In section 6 we present results for the administrative tax data, and in 7 we present results on firm output. Section 8 discusses implications for aggregate output and 9 explores mechanisms behind the results. Section 10 concludes.

2 Context: the Ugandan economy and tax system

2.1 Ugandan economy

Uganda is a low-income country with a low tax-to-GDP ratio. Uganda's per-capita income was $3,040 in PPP (World Bank, 2023b). Its tax-to-GDP ratio – 12.2% in 2020/21 (OECD et al., 2023) – is slightly below the 15.6% average in Sub-Saharan Africa (OECD et al., 2023) and is substantially lower than the 34.1% average in OECD countries (OECD et al., 2023). However, similar to other developing countries, the low tax-to-gdp ratio is not explained by low tax rates. The statutory Corporate Income Tax (CIT) rate is 30%, making it the 21st highest in the world (Enache, 2023).

The informal sector in Uganda accounts for a large share of economic activity. In 2017, 84% and 87% of the male and female workforce outside agriculture was estimated to work in the informal sector (World Bank, 2017).[3] 72% of businesses are estimated to operate informally, and the informal

[3]If we include employment in agriculture this rises to an average of 95% across both genders (ILO, 2021).

sector accounts for 51% of GDP (Sanday, 2023). These are large estimates by any measure, but are similar to other countries in the region. Across Sub-Saharan Africa, 90% of employment is estimated to be in the informal sector (World Bank, 2019a), 86% of enterprises in Africa are informal (OECD and ILO, 2019), and informality accounts for around 40% of GDP (World Bank, 2019a).

These two facts may be related. High tax rates, if enforced, may encourage formal firms to leave the formal sector and prevent informal ones from becoming formal. Conversely, high levels of informality, incentivises the government to enforce taxes more heavily, retrieving as much tax revenue as possible from the small pool of firms that are formal.

2.2 Ugandan tax system

Uganda uses a modern electronic tax system. Since 2013 firms in Uganda have been required to file their taxes electronically, creating a large repository of digital information on firms' tax filings. The three main tax bases in Uganda are VAT, CIT, and Pay As You Earn (PAYE). Firms with sales above 150 million UGX ($\sim$ 40,000 USD) are required to register for the VAT and CIT. Below this threshold a simple sales tax replaces both tax heads.[4] In contrast to VAT and CIT, PAYE is supposed to be paid for any employee with a monthly salary above 235 thousands UGX ($\sim$ 60 USD), regardless of the firms' sales.

The CIT is filed annually, and is due 6 months after the close of the financial year. Firms are required to file a detailed balance sheet and profit/loss statement denoting each of their costs and revenue sources in their CIT return. The Ugandan fiscal year goes from the 1st of July to the 30th June, but firms can apply to change their accounting year to suit their operations. The CIT rate in Uganda is 30% and is calculated on the net income reported by companies, with few exceptions.[5] Once the return has been filed, firms can amend it – re-file the return with updated information – indefinitely. Not filing a tax return on time cost 200,000 UGX (55 USD) or 2% of the tax liability for the period, whichever is higher. It also gives the URA permission to determine how much the taxpayer owes in taxes.

VAT-registered firms have to submit monthly VAT declarations to the URA for the domestic part of their business due within 15 days of the end of the month. Similar to the CIT, firms may file amendments for past returns indefinitely. Negative liabilities can be carried over to subsequent months when less than 5 million UGX ($\sim$ 1,400 usd). Higher amounts can be claimed as a refund, but that triggers a refund audit. The VAT rate in Uganda is 18% and is calculated on the difference between total sales and inputs from VAT registered firms.

[4]This automatically excludes many of the smallest taxpayers in Uganda, however, based on conversations with officers at the URA it is clear that those taxpayers rarely experience any form of compliance intervention. Firms can choose to register for CIT and VAT even if they are below the threshold. The reason they might do so is because of wanting to deal with VAT firms, or bid for government procurement.

[5]Repatriated branch profits are taxed at 15%, income from non-residents providing shipping services is taxed at 2%, and income from residents providing telecom services is taxed at 5%. In the period 2013 – 2022, this is relevant for 0.1% of all the CIT returns filed by taxpayers in Uganda.

Similar to VAT, firms liable for PAYE have to submit monthly PAYE returns for their employees within 15 days of the end of the month. The PAYE tax in Uganda is based on the total income of an employee, including any non-salary payments, and the rates vary from 10% to 40% depending on the employee's income.[6] The PAYE tax is withheld by the firm and remitted directly to the URA, similar to payroll taxes in other countries.

2.3 Taxation and enforcement as an obstacle to firm growth

To explore to what extent high tax rates and tax enforcement impose an obstacle to firm growth in Uganda we draw on our survey of formal firms (for more details on sample selection and implementation, see section 4.2), and a survey of informal firms conducted by the World Bank (for details on sample selection and methodology, please see World Bank (2019b)).

Information from our survey suggests that high statutory tax rates translate to high tax-induxed operation costs in Uganda. This is not a given. In countries with limited enforcement capacity and/or extensive tax incentives, the effective tax rate faced by firms may be significantly lower than the statutory tax rates (Bachas et al., 2023).[7] Using the results from our survey (details in section 4.2) we document relevant facts about the cost of filing taxes in Uganda, as reported by the firms.

The results from the survey are shown in Figure 1, which reports the cost of taxes as a share of sales, grouping firms into five quintiles based on their sales. There are three key takeaways. First, the costs of taxation as a share of sales are substantial, ranging from 26-36%. Second, the overall costs of taxation (as a share of sales) are substantially higher for the smallest firms, and relatively stable over the remaining quintiles, despite tax payments being lower for small firms. This is because, third, the enforcement costs to the firm – interactions with the URA, and tax corrections – represent a large share of sales for small firms. Overall, results from our survey indicate that taxation imposes a significant economic burden on firms and that the burden of tax enforcement – measured as a share of sales – disproportionately affects the smallest firms.

Not only are the taxation-induced operation costs high, firms also say that taxation and its associated costs is one of the main obstacles to their growth. In the survey, we asked firms for the three main obstacles to their performance. Panel (a) of Figure 2 shows that 63% of firms mention high taxes, tax administration costs or tax corrections as one of their three main obstacles. This is the most frequently mentioned reason, with the second most important reason – competition – mentioned by only 32% of firms. Furthermore, firms do not seem to have strong intrinsic motivations to pay taxes. In panel (b) of figure 2 we show the reply from firms when asked about the three main reasons for paying taxes. Only 38% mention it contributes to economic development, and less than 20% mention it helps their business. Together this demonstrates that formal firms in Uganda consider taxes to be a major obstacle to their performance, and do not see many benefits from paying them.

[6] In appendix figure B1 we show how the tax rate varies with income. It quickly rises to 25%.
[7] The effective corporate income tax rate in Uganda is 15.1%, half of the statutory rate (Bachas et al., 2023).

In addition to presenting a major obstacle for formal firms, high perceived costs of being formal discourage informal firms from formalizing. To shed light on this aspect we draw on a survey of the informal sector in the greater Kampala area conducted by the World Bank in 2016 (for details on sample selection and methodology, please see World Bank (2019b)). In the questionnaire, non-registered firms are asked why they are not registered. In panel (a) of figure 3 we show the distribution of their answers. 36% of firms say they are not formalized because it is too expensive or it "could be bad for business". Together, these are the two most frequently mentioned reasons. Bear in mind, the question is about registering with the local administrative authorities *not even* about registering with the tax authority. The same set of informal firms are asked what would encourage them to formalize, and the number one reason mentioned by 42% of firms are low fees, as shown in panel (b). This underlines that formalization is seen as prohibitively costly or as damaging to business operations by informal firms.

3 Intervention: audit process and selection

The URA conducts three types of audits, ranked from most to least intense: comprehensive, issue and desk. In this section we both introduce each audit intervention, and describe the selection process. The focus of this study is the relative effect of a comprehensive audit compared to a desk audit.

Comprehensive and desk audits vary starkly in their intensity. In a comprehensive audit, auditors can investigate any tax base up to 5 years in the past and the median comprehensive audit takes 93 days.[8] They always involve direct interaction with taxpayers and often include a physical visit to the firm. The average size of the auditing team for a comprehensive audit is 3 officers. In stark contrast, during a desk audit the taxpayer receives a letter from the URA detailing what discrepancy in their tax return has been detected. The taxpayer is then asked to either amend the return to correct the discrepancy, or provide documentation explaining the discrepancy. A draft letter of the desk audit intervention can be found in appendix figure B2. Issue audits are a mix of the two audit interventions. Depending on the case, the audit may look more like a comprehensive audit or a desk audit, which is why we avoid comparing comprehensive audits to issue audits.

A data driven risk-calculation and the capacity of a taxpayer office together determine which taxpayers receive each audit type. Each year the Central Operations Office (COO) at the URA picks risk categories to focus on. In our year of analysis they picked 35 risks. They then take the tax returns from two years prior and determine whether the risk applies to each taxpayer. If it does, they calculate a *risk score* for that taxpayer-risk category combination and estimate the *expected potential revenue* from resolving the risk for that taxpayer. If a risk category does not apply, the implied risk score and expected potential revenue is 0. Finally, for every taxpayer they aggregate the risk score and expected revenue of each underlying risk category to compute a taxpayer specific *total*

[8] While this is the official guidance, during qualitative interviews with firms, we heard stories of auditors raising tax issues that were more than 15 years old.

risk score and *total expected potential revenue*. While the URA avoids auditing the same taxpayer for two years after an audit, tax return submitted in the immediate aftermath of an audit will be used in future risk calculations.

Once it has computed the risk scores, the COO determines how many comprehensive audits the central audit centre – the office responsible for conducting comprehensive audits – can do. Equipped with these numbers, it creates an ordinal ranking of firms. The ordinal rank prioritizes the total risk score over the total expected potential revenue. That is, firms with the highest risk score are always ranked higher. It is only for firms with the same risk score that the expected potential revenue determines their rank. The firms with the highest ordinal rank are assigned a comprehensive audit until the capacity of the central audit centre is reached.

Having concluded the comprehensive audit selection, it calculates the number of issue and desk audits each *local* taxpayer station can do. Taxpayers are generally registered at the local taxpayer station that is closest to their premises.[9] For each taxpayer station, it ranks the remaining taxpayers – those not selected for comprehensive audits – by their expected potential revenue. The highest ranked firms within a local station are assigned to issue audits and the next set of firms are assigned to desk audits, until the capacity of the station is reached for each audit type. The remaining firms do not receive any compliance intervention that year. The details of the selection process vary annually. Our analysis exclusively focuses on financial year 2021, which is the earliest year for which we were able to pin it down.

Due to prioritizing different risk scores for different audits, the composition of audit assignment around the comprehensive audit threshold is not obvious. Figure 4 provides a scatter plot of firms along the ordinal ranking of the firm (horizontal axis) and their normalized total expected potential revenue (vertical axis). The total expected revenue is normalized to the station-specific cut-off that determines whether a firm receive issue or desk audits. There are two takeaways from the figure. First, as mentioned before, the comprehensive audit selection is entirely driven by the ordinal ranking of the firm. The highest ranked firms always receive comprehensive audits regardless of their normalized total expected potential revenue. Second, for firms close but below the comprehensive audit threshold, whether they receive issue or desk audits is driven by the normalized expected potential revenue of the firm. Because issue audits are hard to interpret and there are three times more desk audits around the threshold than issue audits, we compare comprehensive audits against desk audits. As a robustness check we compare comprehensive against issue and desk audits. The results are consistent.

Upon selection, all audit cases are sent to the relevant taxpayer station, where the supervisor distributes the cases among their officers. The officer who receives the case verifies that the data-

[9]The exceptions are: Large taxpayers, Medium taxpayers, government entities, and oil and gas companies. These have their own specialized taxpayer office. Taxpayers from the medium taxpayer office form part of our analysis, but the rest do not. The specialized taxpayer stations – apart from the Medium taxpayer office – do not follow the general compliance plan that the rest of the URA utilizes.

driven risks are still present in their tax return, before moving forward with the audit.

For comprehensive audits, once the risk is verified, a notification is sent to the taxpayer informing them that they have been selected for an audit and asking them to prepare their records. The auditor either visits the taxpayer's premises and looks through the records, or the taxpayer comes to the URA offices with records in hand. If the auditor finds that the taxpayer is compliant, the audit is concluded. If the auditor determines that the taxpayer evaded taxes they can issue an assessment. An assessment is effectively a re-filing of a taxpayer's tax returns with revised amounts.[10] The taxpayer then has 45 days to object to the assessment. If they object, the case gets sent to a new department at the URA that will go through the audit and check that everything was done correctly.[11] If the taxpayer accepts the assessment, the audit is concluded and the higher tax liability is added to the taxpayer's ledger. If a firm was found to have underdeclared their tax liability the auditor is legally empowered to add interests to the outstanding liability and levy fines on the taxpayer. In practice fines are not frequently used to reduce the economic burden on the taxpayer. Appendix figure B3 graphically illustrates the audit process.

The process for desk audits is much shorter. Upon verification of the discrepancy by the auditor, a letter is sent to the taxpayer informing them about the exact discrepancy that was discovered. The letter requests the taxpayer to either amend their tax returns eliminating the discrepancy, or provide documentation for why the discrepancy is correct. Taxpayer are given 7 days to respond to the request, after which the officer can issue an assessment. Similar to issue and comprehensive audits the taxpayer can object to the assessment, otherwise the audit is concluded. This compliance strategy is similar to the letter intervention analyzed in Almunia et al. (2023). In fact, the desk audits were modelled on the letter intervention analyzed in our paper.

4 Data & descriptives

We combine detailed administrative data on the universe of firms tax filings with audit records, risk scores, and a survey, to evaluate the impact of receiving a comprehensive tax audit. In this section we describe each dataset, discuss how we construct our main sample of analysis, and present descriptive statistics.

4.1 Administrative data

The primary administrative dataset leveraged in this paper is a panel that combines the universe of CIT and VAT tax returns with audit records and risk scores, all provided by the URA. We also introduce the PAYE data from the URA and Census data from the Ugandan Bureau of Statistics (UBOS), because we use it to contextualize and describe our sample.

[10] For example, if a taxpayer was found to have underreported their sales by 10 million UGX in the VAT return for the month of June 2018, the assessment would override the original return and specify the higher amount.

[11] If the taxpayer is unhappy with the outcome of the objection they can file a legal complaint. Theoretically, this can be escalated all the way up to the Ugandan supreme court. For a recent example see (The Independent, 2023).

4.1.1 Sources

Administrative tax (CIT) The CIT dataset contains the universe of tax filings from all CIT registered firms for financial years 2013 to 2022. The CIT tax returns includes a full balance sheet and profit-loss statement of the firm. It contains information on total sales and a detailed breakdown of each type of cost. In our analysis focus on CIT returns for the years 2019 – 2022. Baseline characteristics are drawn from the last unaffected year before the audit, which is financial year 2019.[12]

Administrative tax (VAT) The VAT data contains the universe of monthly returns filed between January 2013 and November 2022. The VAT data compliments the CIT data in two key ways. First, it is filed monthly, providing us with information at higher frequency. Second, it is an important tax for revenue collection in Uganda, comprising 30% of all revenue collected (OECD et al., 2023). We focus on VAT returns filed between June 2019 and November 2022. Baseline characteristics are based on the returns filed between June 2019 and July 2020. This is the financial year the URA considers in their risk calculations, and corresponds with the baseline data from the CIT.

Administrative tax (PAYE) The PAYE data contains the universe of monthly returns filed between January 2013 and June 2022. The PAYE complements the other administrative datasets by providing us with the number of employees at the firm. We do not have any information on PAYE returns post-audit, so we use this dataset exclusively to describe the firms in our sample.

Audits and risk scores Data on comprehensive and issue audits come from the URA's online database, which contains the universe of all audits that were *assigned* from January 2013 to January 2024. This is important because even if an audit was assigned and sent to a taxpayer station it was not necessarily executed. The data contains information on the type of audit (issue/comprehensive), when the audit was assigned, started, and completed, and what tax bases the audit was conducted on. If an audit does not have a start date and there is no record of an audit outcome, we assume it did not happen.[13] The data also records which supervisor was in charge of the audit and the audits' outcome. For our analysis we focus on audits that were assigned in financial year 2021/22.

Desk audit information is not included in the URA's online database, because they are considered "compliance advisory" interventions rather than audits by the URA. We received information on desk audit allocations for financial year 2021/22 from the URA's annual compliance report. While the report does not include the start date of the desk audit, it does include information on when the desk audit finished. Given the timeline of a desk audit, we assume that the desk audit started

[12] We exclude financial year 2020 because depending on when a firm files a tax return and when it was audited, the 2020 return may be affected.

[13] Based on conversations with auditors at the URA that seems a reasonable assumption. The number of executed audits form part of the auditors annual evaluations, hence they have every incentive to add it to the online system.

7 days before its completion date. Furthermore, we assume the desk audit did not happen if there is no completion date.[14]

Risk score information was shared with us by the URA officer in charge of audit selection during financial year 2021/22. In that financial year the risk scores and expected potential revenue calculations were based on a taxpayers' discrepancy across 35 different risk categories. We do not only observe a firms total risk score and expected potential revenue, but also their risk score and expected potential revenue for each individual risk category. If a risk category did not apply to a given firm the implicit risk score and expected revenue was set to 0 for that category. The aggregate risk score and expected revenue are an unweighted aggregate of the scores and expected revenues of each of the underlying risks.

It is possible for a firm to receive several audits in a given year. In particular, after receiving a desk audit, the auditor may trigger an issue or comprehensive audit because they believe there are more compliance risks to investigate. For the purpose of our analysis, we consider the first audit intervention that a firm receives to be the assigned one. Anything happening afterwards is endogenous to the outcome of the first audit intervention.

Census The Census of Business Establishments (CoBE) allows us to compare our sample of firms to the universe of all firms in Uganda. The CoBE was conducted by the Ugandan Bureau of Statistics (UBOS) in financial year 2020/21, and contains information on the geographic location, sector, number of employees and turnover of all firms in Uganda. We cannot match firms one-to-one so we construct size-categories and compare the number of firms in the size categories.

Figure A1 provides a visual timeline of each of the datasets used in the analysis.

4.1.2 Sample selection criteria

To construct our main sample we merge the administrative tax data with the risk scores from the audit selection process and audit information. We impose a number of restrictions to facilitate our analysis. Table A1 in the appendix highlights how the sample changes when imposing each restriction.

We start with the universe of firms that were flagged for at least one risk in financial year 2021.[15] We proceed to remove firms registered at specialized taxpayer stations. This excludes firms in the Large Taxpayer Office, Oil and Gas office and the Public Sector Office. These offices do not follow the compliance improvement plan of the rest of the URA, instead conducting their own compliance strategies. Furthermore, the administrative capacity at these stations is different from non-specialized offices, they typically have a higher officer-to-taxpayer ratio. We also remove firms

[14] Desk audits are short interventions, so they are unlikely to lead to long-standing disputes which could cause an audit not to be completed.

[15] Firms are defined as taxpayers who are registered as non-individuals and who are a private entity. This excludes entities registered as clubs, estates or trusts, government bodies, international organizations, and NGOs.

in the mining sector, because 40% of firms in the mining sector are registered with the Oil and Gas office. The remainder would be a heavily selected sample.[16]

We remove firms that are audited the year before our year of analysis. The URA typically avoids auditing the same taxpayer two years in a row. If it does it is usually to follow-up on a prior issue. Including these set of audits could confound our estimates for two reasons. First, the firm is selected into the audit for reasons not related to the risk scores, hence the audit is not induced by the instrument. Second, the audit in the previous year could already have changed the firm's behaviour, leading us to identify a disparate treatment effect.

We restrict to "active" firms, which we define as any firm that filed either a CIT or VAT return at some point prior to the audit. Finally, we only consider stations where at least one of the firms received a comprehensive audit.

After imposing each step, we remain with 20,838 firms, of which 231 were selected for comprehensive audits, 1,041 for issue audits, 12,783 for desk audits, and 6,783 received no intervention. The final sample used to identify our treatment effect is restricted to the 13,014 firms assigned to comprehensive and desk audits.

4.1.3 Descriptives

In Table 1 we show that our sample of firms is skewed toward the upper tail of the firm size distribution, both among formal firms and the universe of firms in Uganda. We leverage information from tax filing records to assess how our sample compares to the universe of tax filing firms. We leverage the Census of business establishments from UBOS to determine how our sample compares to the universe of all firms in Uganda.

Table 1 presents summary statistics for our samples, and p-values for whether the means across samples are significantly different. In Column (1) we present the averages for all firms that filed CIT or VAT taxes at some point prior to the audit year. This also includes firms that did not receive a risk score. Column (2) restricts to the full comprehensive sample, whereas Column (4) restricts to the part of the comprehensive sample that are used to identify our treatment effects, namely the set of firms that tend to fall within the optimal bandwidth.

Firms in the comprehensive sample – shown in Column (2) – are on average larger than the universe of firms filing CIT or VAT. They report higher sales and purchases, higher VAT, but lower VAT liabilities. Similarly, they report higher sales and costs in the CIT data, higher profits and in this case also higher tax liabilities. One caveat is that they tend to report roughly half the number of employees. Note that firms in the comprehensive sample are far more likely to file tax returns across each of the three tax bases presented here. This is consistent with the URA leveraging information from a firms' tax return when calculating the risk scores. Finally, we consider the

[16]The highest percentage of firms registered with one of the specialized offices in any of the other sectors is manufacturing at 8%.

sector decomposition in the two samples. Interestingly, a much higher share of the firms in the universe of tax filers are firms in the retail sector, consistent with the URA avoiding targeting some of the smallest formal firms in their database. Column (3) presents the p-values for whether the averages of the full comprehensive sample are significantly different from the universe of firms. 26 of the 32 variables are significantly different at the 5% significance level.[17]

These differences become more pronounced when we consider the subset of firms in the comprehensive sample that tend to fall within the optimal bandwidth. In this case, firms are larger across all measures we have access to, even number of employees. Despite being larger, they report lower VAT liabilities, profits and CIT liabilities than the universe of firms. This is in line with the objectives of the URA. Comprehensive audits are the heaviest tax compliance intervention the URA has at its disposal, these types of audits are targeted towards large taxpayers that are not contributing sufficiently (or at all) towards revenue collection. Overall, Table 1 demonstrates that the subset of firms we study are some of the largest formal firms in Uganda.

To compare our sample of firms to the universe of all firms in Uganda, we turn to the 2020 Census of Business Establishments. The number of employees in the census is measured in bins. To facilitate the comparison we construct equivalent bins for firms that file PAYE returns, and divide the number of firms in each category that file PAYE taxes by the number of firms in each category that are present in the census. This allows us to indirectly assess the share of tax filing firms in each size category.[18] We limit the comparison to focus on the universe of firms filing PAYE and the comprehensive sample. In appendix Figure A2 we show the share of firms filing taxes in each of the size categories. As expected, both the set of firms filing taxes and the set of firms in the comprehensive sample are heavily skewed towards the upper tail of the size distribution. The number of firms that file PAYE taxes add up to 95% percent of all firms that employ more than 100 employees (19% for firms in the comprehensive sample). In stark contrast, the share of firms filing taxes among firms with one employee is merely 2% percent of all firms in the census (0.5% for firms in the comprehensive sample). Recall that the firms that tend to fall within the optimal bandwidth are, if anything, even larger than the firms in the comprehensive sample, underlining that we are investigating the impact of heavy enforcement interventions on some of large firms in Uganda. However, the very largest firms are not included, since they are registered in specialized taxpayer offices.

4.2 Survey data

Our survey was designed to determine whether firm existed, and if they did, what their financial situation looked like. To achieve this, we divided the survey into two parts. The first, intended to determine whether firms exist, was an extensive tracking exercise where we tracked down every firm that was on the margin of receiving a comprehensive audit. The second component was the

[17] The p-value of a joint f-test on all variables is 0.000 as shown in the second to last line.
[18] There is no unique identifier that matches across the tax data and the census.

survey itself. The survey included detailed questions about the firms' financial situation, such as: revenues, costs, profits, number of employees, and prices and quantities of goods sold. Because we were interested in the effect of tax enforcement measures, we also included an extensive module on the costs of filing taxes. Here we took a broad view on what such costs might include. We asked firms about: 1) The administrative filing costs, 2) The direct cost of receiving compliance interventions, and 3) The cost of interacting with the tax authority (through audits or otherwise). Together this was designed to capture the full monetary burden imposed on firms by taxation.

The survey compliments the administrative tax data in two ways. First, it allows us to discern whether audits affect firm sales. Firm sales are notoriously difficult to measure using administrative tax data because, by definition, non-compliant firms misrepresent their sales and profits in their tax returns. As such, it is hard to disentangle whether an audit changes reporting behavior (evasion) or firm sales (output), or both. We both investigate whether audits affect firm sales through both firm shut downs, and reduction in output conditional on remaining operational. Second, the survey provides us with an in-depth view into the cost of audits and tax compliance in Uganda. To the best of our knowledge, there is little extensive information on the total costs of filing taxes, including administrative costs of interacting with the tax authority and the cost of tax correction in developing countries.[19] We conducted the survey in the period November 2023 to July 2024.

Selection: We select firms for the survey to align with our identification strategy. Initially we focused on all firms in the Kampala Metropolitan Area that were close to the universal regression discontinuity cut-off for comprehensive audits *and* the various station-specific cut-offs for issue audits. Later on, given budgetary constraints, we redirected our effort to focus on firms that were within the optimal bandwidth for one of our key outcomes; the probability of filing a tax return in the year after the audit. For this part of the sample we also included firms outside Kampala Metropolitan Area. The narrow sample thus contains all the firms that identify our main treatment effect. To gain power, we use the broader sample to understand whether firms reduced their scale in section 7 and for descriptives and context in section 2. A total of 839 and 3,323 were reached out to in the narrow sample and total sample, respectively. Of the firms reached out to XXX and XXX were surveyed in the narrow and broad sample.

Tracking: To maximize our chances of finding firms two years after they were audited, we merge contact information from four sources: 1) Online information ("Googling"), 2) The Ugandan Registry Services Bureau (URSB), 3) The 2020 Census of Business Establishments, and 4) Contact information registered with the URA.[20] The URSB is the official registration agency for all firms in Uganda and firms are supposed to renew their registration every year. Around December 2022 the

[19] The most recent version of the World Bank (2023a) includes information on hours spent on tax compliance, but it is hard to assess what that implies in terms of costs to the firm. Furthermore, there is no information on tax corrections which we find to be important.

[20] There are no unique identifiers across these datasets. We conduct a fuzzy merge based on names, and restrict to cases where the names are extremely similar.

URSB switched to an online database, facilitating access to the registry information.[21] The contact information from the URSB and the URA contains information on both firms addresses and phone numbers. The 2020 Census of Business Establishments only contains information on the geographic location of the firms.

Once we combine all source of information we proceeded in three steps. First, we called all the phone numbers we had on file that were not registered with the URA. We avoid the phone number registered with the URA because we know from a pilot that we conducted in the year prior that some firms know which numbers are registered with the URA. To avoid biasing respondents we only used these phone numbers as a last resort. Second, if we did not reach the firm through a phone number, we sent enumerators to each location on file that within the Kampala Metropolitan Area.[22] If we did not find the firm at any of the locations – and the neighbors did not know the firm or what happened to it – we called firms on the phone numbers they registered with the URA.

Table 2 presents summary statistics for the tracking exercise, and survey completion rates. We focus on the narrow sample. Of the 839 firms that we attempted to track down, we found evidence on whether a firm existed for 90.9% of them. That is to say, we either reached the firm or we learnt from people familiar with the firm – such as neighbors or previous employees – that the firm closed. The remaining 9.1% we label "vanished", and consider them closed in the analysis.[23] To our surprise, 17 firms on our list turned out to be non-governmental organisations (NGOs). Because they registered as a business with the URA it is unclear to what extent they behave as a NGO or enterprise. We include them in our analysis.

We completed interviews with 45% (N = 681) of all the firms that we found and were operational at the time of the survey. This is slightly below the response rates in the World Bank Enterprise Surveys, which focus on similar sized firms (<**empty citation**>).

To check whether firms believed we were from the URA – despite our best efforts – we asked enumerators to indicate wether the respondent did in the survey. Based on qualitative work and our pilot it was apparent if the respondent thought we were affiliated with the URA. Of the 308 firms we interviewed from the narrow sample 25 (8%) thought we were from the URA. For the broader sample the equivalent numbers are 43 (6%).

[21] One important limitation is that the online database only holds information for firms that updated their registry after the switch to the online system. Due to limited resources it was not possible to look through the manual information on firms for years prior to the switch to the online system.

[22] This includes firms that were located in Mukono and Entebbe, but excludes firms that were registered in Jinja. For firms within our narrow sample we also sent an enumerator to Jinja. We did not send enumerator to any of the other cities around Uganda.

[23] We conduct several checks to verify this is a reasonable assumptions. More details in section 7.

5 Empirical strategy

We exploit discontinuities created by the audit selection process to estimate the local average treatment effect of a comprehensive tax audit. Our estimation strategy is designed to mimic the audit selection process. The validity of our design relies on three key assumptions: 1) No manipulation of audit assignment, 2) Observed differences around the discontinuity are only due to the difference in audit assignment, 3) Changes in audit assignment has large effects on the likelihood of receiving an audit. We provide evidence that each assumption holds.

Our first-stage equation – capturing the increase in the probability of receiving a comprehensive audit above the threshold – is written formally in equation 1.

$$D_f = \alpha \, \mathbb{1}\{z_f \geq l\} + \underbrace{\lambda_1 \, \mathbb{1}\{z_f \geq l\}(z_f - l + 0.5) + \psi_1 z_f}_{\text{linear slopes in ordinal rank}} + \underbrace{\Gamma_1 \mathbf{X}_f}_{\text{Controls}} + \varepsilon_{1f} \tag{1}$$

Where D_f is an indicator for whether firm f received a comprehensive audit, z_f is the ordinal rank of firm f used for selecting comprehensive audits, and $\mathbf{X}_f$ is a vector of controls. Our estimation model allows for separate slopes on each side of the discontinuity. l is the threshold for the assignment of comprehensive audits; the ordinal rank of the last firm that was *not* assigned to a comprehensive audit. We recentre the threshold by $l+0.5$, such that the discontinuity is measured midway between the last firm not assigned to a comprehensive audit and the first firm assigned to a comprehensive audit.

Our vector of controls, $\mathbf{X}_f$, includes station fixed effects and separate indicators for whether a firm is in the VAT or CIT sample. Both set of controls are included to mimic the selection procedure at the URA. We include station fixed effects because the counterfactual – desk audits – is conducted and selected by officers at the local stations, not the central audit centre. While the central operations office assigns an initial set of firms for desk audits, officers at the stations have some leeway to adjust the selection. Station fixed effects control for any differential selection into desk audits.[24] We include two indicators for whether the firm is in the VAT or CIT sample because the URA targets comprehensive audits at firms that are in more than one tax base.

Thresholds: While we do know how firms are selected for audits we do not know exactly where the cut-off was set. We therefore employ a data-driven method to identify the universal threshold for comprehensive audit assignment. We take the universe of firms (N = 20,838) that remain after imposing each cleaning step and rank the firms from highest to lowest based on their risk score (first) and potential expected revenue (second). That is, we rank firm by their risk score, and for firms that share the same risk score we rank them by their potential expected revenue. Once ranked, we conduct a structural break test to find the single largest change in the likelihood of being *assigned*

[24] There is evidence that the assignment of desk audits from the central audit centre was not always binding. Some stations decided to give all firms that had a risk score a desk audit.

a comprehensive audit (which is distinct from receiving an audit). Because we know the initial number of firms assigned to comprehensive audits, an alternative approach would be to assume that the 289 firms with the highest rank were assigned a comprehensive audit. Both approaches yield similar results, and we conduct robustness checks using the alternative threshold.

Running variable: Our running variable is the ordinal ranking of the firm among the 13,014 that were assigned to comprehensive and desk audits. Once we derive the threshold, we rank firms again according to their risk score (first) and potential expected revenue (second). We then calculate the ordinal rank of the firm relative to the firm just below the threshold. We leverage the ordinal rank because it drives selection into comprehensive audits as illustrated in Figure 4.[25] By calculating it on the sample of firms assigned to desk and comprehensive audits we avoid having gaps in the data from not including firms assigned to issue audits or that did not receive any intervention. As a robustness, we include everyone assigned to any enforcement measure.

Counterfactual: As discussed extensively in section 3 we compare comprehensive audits against desk audits. This is both because issue audits are hard to interpret and there are three times as many desk audits around the comprehensive audit threshold as illustrated in figure 4.

To capture the reduced form effect of being above the comprehensive audit threshold on an outcome of interest, we run the following reduced form regression discontinuity equation:

$$Y_f = \beta_{RF}\, \mathbb{1}\{z_f \geq l\} + \underbrace{\lambda\, \mathbb{1}\{z_f \geq l\}(z_f - l + 0.5) + \psi z_f}_{\text{linear slopes in ordinal rank}} + \underbrace{\Gamma \mathbf{X}_f}_{\text{Controls}} + \epsilon_f \tag{2}$$

The only difference with equation 1 is that we replace D_f with an outcome of interest Y_f.

Finally, we are also interested in understanding the implied effect of *receiving* a comprehensive audit on the outcome of interest, not just the effect of being above the threshold. To derive the effect of receiving a comprehensive audit we run a two-stage least-squares (2SLS) regression, where the first stage is given by equation 1 and the second stage is written below:

$$Y_f = \beta_{2SLS} \hat{D}_f + \underbrace{\lambda_2\, \mathbb{1}\{z_f \geq l\}(z_f - l + 0.5) + \psi_2 z_f}_{\text{linear slopes in ordinal rank}} + \underbrace{\Gamma_2 \mathbf{X}_f}_{\text{Controls}} + \varepsilon_{2f} \tag{3}$$

Where $\hat{D}_f$ is the predicted increase in the probability of receiving a comprehensive audit from the first stage. Throughout our results we present both reduced form RD estimates and 2SLS estimates. β_{RF} indicates the former whereas β_{2SLS} indicates the latter.

Bandwidth: The bandwidths are selected using a data-driven approach implemented through the "rdrobust" package by Calonico et al. (2014). We use the universal optimal bandwidth, but also

[25] Other studies with similar set-ups have used a score standardized around the cut-off (Busso et al., 2023). We prefer the ordinal ranking because it is the driving force behind selection into comprehensive audits.

show that the main results are robust to changes in the bandwidth.

5.1 Increase in the likelihood of receiving a comprehensive audit

The discontinuity created by the audit selection process captures large differences in the likelihood of receiving a comprehensive audit, the most intense tax enforcement intervention at the URA. As outlined in more detail in section 3 these audits last for 3 months on average, and can involve examination of all records in relation to all tax bases up to 5 years in the past. As such, they represent in-depth examination of a firms' accounts.

Figure 5 shows that crossing the threshold increases the likelihood of receiving a comprehensive audit by 69 p.p., which is highly statistically significant. The F-stat for a two-sided test of the coefficient being greater than 0 is 107. We plot the residualized probability of receiving a comprehensive audit along the ordinal ranking of firms and the share of firms receiving a desk audit in each bin. We construct residualized outcomes to match the variation shown in the figure to the regression results.[26]

The figure highlights three important aspects of the audit selection process. First, the probability of receiving a comprehensive audit and the share of firms receiving a desk audit do not add up to 1 because not all the firms assigned a desk audit received one, underlining the importance of controlling for station fixed effects. Second, no firms above the threshold were assigned a desk audit because *all* of them were assigned comprehensive audits.[27] If the comprehensive audit was not executed, the firm was not downgraded to a desk audit. Instead the audit would be carried over into the next financial year or dropped. Finally, the probability of receiving a comprehensive audit decreases in the ordinal ranking above the threshold. This is driven by auditors choosing "easier" audit cases first. Firms with a higher risk score have more risk factors to investigate, implying more challenging audits. The performance of auditors is measured in the number of audits they execute and the amount of revenue they collect, incentivizing them to avoid difficult and long-lasting audit.

5.2 Instrument validity

We have shown that the discontinuity causes large changes in the likelihood of receicing a comprehensive audit. The two remaining assumptions for the validity of our RD design are: 1) No manipulation of audit assignment, and 2) Observed differences around the discontinuity are only due to the difference in audit assignment. We perform a series of balance and validation tests to verify that both assumptions hold.

[26] We run a regression of the outcome variable against a constant, α, station fixed effects, and two indicators for whether the firm was in the CIT or VAT sample, captured by $\mathbf{X}_f$. Then we add back the average of the outcome for scaling purposes. Formally, the residualized outcome is given by: $\tilde{Y}_f = Y_f - (\hat{\alpha} + \hat{\mathbf{X}}_f) + \bar{Y}_f$.

[27] In appendix figure A3 we show this more clearly by plotting the residualized probability of being assigned to a comprehensive audit. Literally every firm above the threshold was assigned a comprehensive audit.

Manipulation by firms: The first concern is that firms might manipulate their risk score to not be assigned a comprehensive audit. To successfully manipulate selection into comprehensive audits, a firm would have to do the following. First, find out which risk parameters are used to conduct the risk analysis in a given year. This is hard to do. The group in charge of conducting the risk analysis is small, there were 35 risk parameters involved in the calculation, and both the officers in charge of risk calculations and the risk parameters change every year. Furthermore, the risk score calculations are divided up across officers within the risk score team so that no single officer knows how every risk parameter is computed. Second, find out how the risk profile of the firm compares to that of other firms. What matters for selection is not just a firms' risk score, but its risk score relative to the risk score of other firms. Not only would the firm have to manipulate its own risk score, but ensure that it is low relative to that of other firms. Finally, the firm would have to adjust its tax returns from two years prior. Risk calculations are done retrospectively. Even if a firm gained access to someone that knew how each parameter was calculated in the current year, it would have to file an amendment for its old tax return *before* the risk calculations were completed. Since risk selection and calculations go hand-in-hand, it will be extremely difficult for a firm to adjust its past tax return in time.

Although corruption is always a possibility, given the institutional context outlined above it would be challenging to affect the *assignment* process through corruption. To the extent corruption affects who is audited or the outcome of an audit this would be reflected in the data. Specifically, it would reduce the likelihood of receiving an audit, or the revenue collected from the audit. Our estimates can thus be interpreted as a lower bound.

A density test would not be appropriate in this context because the running variable is the ordinal rank of firms. To alleviate concerns that the ordinal ranking is hiding attempts at manipulation we conduct two exercises. First, we show in appendix figure A4 the mass of firms at each discrete risk score. There is no evidence of differential mass on each side of the threshold. Second, we conduct the test for density manipulation suggested and developed by Cattaneo et al. (2018); Cattaneo et al. (2020) for an alternative running variable designed to capture the relative importance of risk score and expected revenue in selection. The alternative running variable is defined as the risk score value plus the expected revenue of the firm relative to the maximum revenue of all firms with the same risk score.[28] The density is highly non-normal towards the lower tail of the distribution, as seen in figure A4. To ensure that the density test can capture differences in mass around the discontinuity we restrict the test to run within the optimal bandwidth. There is no evidence of density manipulation around the threshold. The resulting p-value is 0.206. A histogram with the density plot and function estimated using the "rddensity" command is shown in appendix figure A5.

[28]Formally, let zn_f be the new running variable for firm f, $risk_{j(f)}$ be the risk score of the group of firms j that f belongs to, and r_f be the expected revenue of firm f. Then, formally $zn_f = risk_{j(f)} + \frac{r_f}{\max(r_{j(f)})}$.

Balance: One might be worried that the selection procedure by the URA causes firms above and below the threshold to be systematically different. To assess whether this is a concern we conduct a series of balance checks. First, we run equation 1 for a large set of baseline covariates, where the baseline covariates replaces the indicator for whether a firm was audited in the regression. The results can be seen in appendix table A2. We reject the null hypothesis of no difference for 1 of the 30 baseline covariates. Second, we conduct a test for the joint significance of all baseline covariates combined, and do not reject the joint null hypothesis of balance in all baseline covariates across various specifications. We conduct this test varying both the bandwidth and whether we impute 0's for missing values. As seen in the last rows of appendix table A2, the p-value of an f-test of joint significance is 0.129, 0.240 when using the minimum and maximum optimal bandwidth, respectively. Finally, we conduct a dimension reduction exercise where we predict our main outcomes using all the 30 baseline covariates shown in table A2, and then run the RD on the predicted outcome. The results of this exercise, shown in appendix table A3, demonstrate that imbalances in the baseline covariates cannot explain the large impacts we find.

Manipulation by URA: One might be concerned that officers at the URA change the threshold. It could be the case that the central audit centre changes the threshold to include or exclude firms with certain characteristics. We show that this is unlikely to be the case. First, if firms right below the threshold are systematically different from the ones right above, this should show up as violation in the balance tests. As shown in table A2, we did not find evidence of such violations. Second, the difference in the threshold from the data-driven method and what would be implied from the initial assignment of comprehensive audits is minimal. Using the threshold from the initial assignment of comprehensive audits would shift the threshold 4 steps lower on the ordinal ranking. As a result, our estimates are robust to using the assigned threshold that would be implied from the initial assignments of audits. We show this in appendix table A10.

6 Results using administrative tax data

The purpose of an audit for the URA is to raise revenue through correcting past evasion and in the post-audit periods through deterrence. In this section we show that while audits do detect evasion, post-audit compliance drastically reduces both because firms exit the tax system and because the remaining firms report lower tax liabilities.

6.1 Tax corrections

We first investigate whether comprehensive audits differentially increase revenue potential through detecting and correcting past evasion. From the audit data we know what corrections were filed by the URA for previous tax returns as a result of the audit. Merging the audit corrections with the original tax returns we can discern to what extent the corrections increased firms tax liability. Unlike comprehensive audits, desk audits do typically not lead to corrections, but instead ask the

taxpayer to self-correct their past tax returns. To measure the differential increase in tax liability from a tax correction we therefore compare the increase from corrections (among firms receiving comprehensive audits) against the increase from self-corrections (among firms receiving desk audits).

Tax liabilities and corrections are notoriously noisy, making it often hard to detect average effects even in experiments with a large sample size (Pomeranz, 2015). We present three transformations of the amount variable when discussing impacts on tax liabilities and corrections. First, we winsorize the top and bottom 1% of the amount variable (Advani et al., 2021). Second, we scale the amount variable by firms' sales at baseline and winsorize the scaled variable at the top and bottom 1%, following Kennedy et al. (2024). Third, because we do have a significant number of zeros, we also use the inverse hyperbolic sine transformation (IHS) of the outcome variable. All three variations are commonly done in the literature.

Figure 6 shows that comprehensive audits induce differential changes in firms tax liabilities through tax corrections. In panel (a) we show that firms above the discontinuity are 33 p.p. more likely to be caught evading on any of their VAT returns (246% increase over the control mean). The implied effect of receiving an audit using the 2SLS is an increase of 48 p.p. In panel (b) we show that the corrections lead to an increase in the tax liability of 34 thousand USD, significant at the 1% level. Scaling the increase in the tax liability by average baseline sales (shown in panel (c)) suggests that the correction amount is approximately equal to 9.4% of baseline annual sales.

Our estimate do not imply that the URA actually collected 34 thousand USD from the marginal comprehensive audit. We do not have information on the actual payment made by firms, hence the number should be interpreted as the upper bound on the potential revenue collected by the URA through correction of past evasion. We discuss robustness in section 6.3.

6.2 Post-audit compliance behavior

Having established that comprehensive tax audits increase firms tax liability through correcting past evasion, we next investigate the tax filing behavior of firms in the financial years after the audit. We primarily focus on CIT because we have information on filing behavior from two years post-audit for the CIT and only 5 months post-audit for the VAT.

6.2.1 Post-audit tax filing

To measure extensive margin filing behavior of firms we focus on the CIT tax base, which we can follow for two years post-audit.

In the year immediately after the audit, financial year 2021/22, firms are less likely to file a CIT tax return. Figure 7(a) shows that firms above the discontinuity are 20 p.p. (35% compared to the control mean) less likely to file a CIT return in the year after the audit. The implied effect of receiving an audit, given by the 2SLS, is a decrease of 30 p.p. Both the reduced form and two stage-least-squares effects are statistically significant at the 1% significance level. The result is driven by

firms that were *marginally* selected for a comprehensive audit. Note however that each point in the figure contains roughly 35 firms. It is not the case that the results are driven by a singular firm.

The effect persists and becomes slightly larger in the second year post-audit. Panel (b) of figure 7 shows that firms receiving comprehensive audits are 34 p.p. (51%) less likely to file a CIT return in the second year post-audit, compared to 30 p.p. one year post-audit. This suggests that the results are not driven by aggressive tax planning in the immediate aftermath of the audit. If firms believe they are not likely to be audited again for some time following an audit, the best time to evade taxes may be in the immediate aftermath of one. In the literature this has been colloquially referred to as the "bomb-crater" effect (Alm and Malézieux, 2021; DeBacker et al., 2015; Kasper and Alm, 2022; Kasper and Rablen, 2023). Consistent with our results there are no institutional incentives in the Ugandan tax system for firms to behave in such a manner. Not filing a tax return on time leads to a penalty, and the information submitted in the immediate aftermath of the audit will be used to calculate the risk score in subsequent years. We interpret our results as being driven by firms leaving the formal sector in the aftermath of an audit.

We do a number of checks to verify that the results can be interpreted as firms exiting the formal sector. In appendix figure A6 we show that the result also hold when we use an indicator for whether a firm either filed a VAT or CIT return. The reduced form coefficient reduces to 16 p.p., but remains statistically significant at the 1% level. This suggests that firms do not just stop filing the CIT, but also other taxes such as the VAT. The results are also not driven by fictitious firms. It could be the case that the "firms" leaving are fictitious firms that only exist to create and sell fake invoices (see Carrillo et al. (2023); Waseem (2023) for two recent papers exploring this phenomenon in developing countries). To verify this is not the case in Column (4) of table A8 we include a control for firms that have been identified as fictitious by the URA.[29] The result does not change. Our survey results, presented in section 7 further support the interpretation that these are not "fictitious" firms.

We discuss further robustness checks in section 6.3. Taken together we interpret our results as a strong evidence that audits are inducing real firms to exit the tax system.

6.2.2 Post-audit tax liability

The extensive margin effects are strikingly large, but may not have implications for the tax liabilities submitted by firms in the aftermath of the audit. If the firms that stop filing taxes are small and did not pay substantial taxes at baseline, removing them from the tax system would have little revenue implications for the tax authority. Furthermore, if the firms that kept filing taxes submitted substantially higher liabilities post-audit, tax liability may actually increase despite some firms exiting the tax system. We again focus primarily on CIT taxes, but show that the results hold for the VAT as well. Because we are interested in the overall change in tax liability we impute 0's for

[29] We received information on all firms identified as being "suspicious" and all the firms "confirmed" to be fictitious by the URA up until November 2023. We take a conservative approach and include an indicator for all firms labeled as "suspicious".

firms that did not submit a tax return.

When we focus on CIT returns that pertain to the year after the audit (but filed two years after) – effectively allowing business operations to be affected by the audit – we find strong negative effects on CIT liability. Note this does not include the increase in tax liability triggered by the tax corrections, such increases are paid separately and would not be reflected on a tax return. As shown in Column (1) of table 3 we find that the probability of reporting a positive tax liability for firms that received a comprehensive audit, as opposed to a desk audit, reduces by 26 p.p. (52%), which is statistically significant at the 1% level. Similarly, we present three variations of the tax liability amount variable, all of which point to a negative effect. The CIT liability submitted to the URA decreases by 5,463 USD (shown in Column (2)), this represents a decrease of 2 p.p. relative to baseline sales (shown in Column (3)). Finally, the inverse hyperbolic sine transformation also points towards a large negative effect. RD figures with the equivalent outcomes can be found in appendix figure A8. Overall we take this as strong evidence that comprehensive audits have a negative impact on post-audit CIT liabilities.[30]

Interestingly, when we focus on tax returns filed in the immediate aftermath of the audit, but that pertain to the year of the audit, we find no effect on CIT liabilities as shown in appendix table A4. Throughout the coefficients are small and statistically insignificant. This is probably because most comprehensive audits occur towards the end of the financial year, hence it is unlikely to affect the operations of the firm.[31] Any differences would therefore have been driven by post-operation adjustments in what is reported to the URA. This provides the first indication that the effects we observe are driven by adjustments in firm operations.

In appendix table A7 we break down the overall decrease in tax liability into extensive and intensive filing margins. We leverage reduced form coefficients for this exercise. Because firms can submit tax returns with zero tax liability, there are three margins of interest: 1) whether firms file a tax return, 2) conditional on filing, whether firms report a positive CIT liability, and 3) conditional on filing a return with a positive CIT liability, do they reduce the amount filed. As shown in Column (2) firms are 23 p.p. less likely to file a tax return in the second year post-audit.[32] Conditional on filing, firms are 18 p.p. less likely to file a tax return with a positive liability, albeit this is only marginally significant. Conditional on filing a return with a positive liability, firms report liability that are approximately 19% smaller, but this is statistically insignificant. We conclude that our results are primarily driven by the first margin, with the second margin also playing a minor role. The first margin is shut down in previous work on the revenue implication of tax audits in similar contexts (Kotsogiannis et al., 2024), and helps explain why our results are different.

As a final check, we verify that our results are not only driven by the CIT tax base. Leveraging

[30] In appendix table A6 we attempt to investigate what lines of the tax return is driving the decrease. The results are difficult to interpret, there does not seem to be a consistent story in what is driving the decline.

[31] Firms are notified of the audit shortly before it starts, ruling out extended anticipatory effects.

[32] Same coefficient as shown in panel (b) Figure 7

information on 5 months post-audit from the VAT data, we investigate whether the VAT liability reduces. We use the same outcomes as for the CIT analysis, but take the average across the five months.[33] The results are shown in table 4. RD figures with the equivalent outcomes can be found in appendix figure A9. While all variations of the outcome variable point towards a negative effect, the coefficient is only significant in Column (2) where the outcome is the amount of VAT liability submitted. The coefficient suggests that the VAT liability submitted in the first five-months post-audit reduces by 6 thousand USD on average. This is a far larger magnitude than what we found in the CIT – recall that the VAT data is monthly – but given that only one of the specifications is significant, we interpret this as suggestive evidence of a negative effect.

Altogether we conclude that comprehensive audits have a negative effect on tax liability for the marginally audited firms through two channels: 1) causing firms to leave the tax system, and 2) reducing the probability of reporting positive liabilities. However, it is the former that is driving the aggregate decrease in revenue. We discuss robustness in the section below, section 6.3. We next attempt to characterize the type of firms that stop filing taxes post-audit.

6.3 Robustness

We conduct a series of robustness checks to verify that our results are not sensitive to key adjustments.

Alternative bandwidths. We first present estimates using alternative bandwidths for our main outcomes of interest, shown in appendix figure A7. Our estimates are robust to even large changes in the bandwidths. Furthermore, as expected, the estimates become larger the closer we are to the threshold. This is consistent with the results shown earlier. The firms closer to the threshold are driving the results.

Alternative specifications. Next we show that our results are robust to using alternative controls. We fix the bandwidth to be the bandwidth used in our main specification and change the controls include. In table A8 we focus on whether a firm received a comprehensive audit and whether they filed tax returns. In table A9 we focus on firms tax liabilities. In column (1) we include no controls, in column (2) we only include station fixed effects, column (3) is our baseline specification, and column (4) adds controls for whether firms are fictitious. In column (5) we use the controls from our baseline specification and add sector fixed effects, and in column (6) we include controls for each of the underlying risk categories. This adds 70 separate controls which control for the risk score and expected revenue of each of the underlying 35 risk parameters, an extremely demanding specification. Finally, in Column (7) we run a lasso specification to select the controls. In table A8 all estimates point in the same qualitative direction, remain significant at the 5% or 1% significance level and do not vary by more than 20%. The outcome variables in A9 are more noisy, but are

[33] We will extend the time-frame of the analysis once more recent data becomes available.

still remarkably robust. In all cases the coefficients point in the same qualitative direction. When the outcome is the amount of CIT liability, we lose significance if we don't include any controls. When the outcome is the average VAT due over the five months post-audit we lose significance when we include all risk controls. Adding all risk controls is an extremely demanding specification and not including station fixed effects is incorrect given that stations could conduct desk audits in various ways. As such, we take our results as indicating that they are robust across reasonable specifications.

Alternative threshold. As discussed in section 5, there are two ways of identifying the threshold. Our baseline approach is conducting a structural break test to find the largest change in the probability of being assigned a comprehensive audit. Since we know the initial number of firms assigned to comprehensive audits, we can also identify the threshold by assuming the the highest ranked 289 firms were assigned comprehensive audits. In table A10 we show that our main results are generally robust to using this alternative threshold.

Alternative sample. In our main specification we restrict to firms that are assigned to comprehensive or desk audits. We then calculate the ordinal ranking of firms based on the restricted sample. One may be worried that by moving some firms closer to the threshold on the ordinal ranking our results are relying on assigning incorrect weights to some firms. Or, one may be worried that by comparing firms with different normalized expected revenue (see figure 4) we are introducing imbalances to our estimation (though we do not find any evidence of that when we conduct our balance tests). To address both concerns we run a specification where we include all firms assigned to any intervention. That is, we run an RD on all the firms that are shown in figure 4.

Having verified the robustness of our results we next investigate which firms stop filing taxes in the aftermath of the audit.

6.4 Characterizing compliers: who are the firms that stop filing taxes?

To explore which firms stop filing taxes post-audit, we estimate the average characteristics of the compliers (Abadie, 2002; Imbens and Rubin, 1997) following the methodology by Pinotti (2017). We conduct the analysis with two outcomes: the probability of receiving a comprehensive audit, and the probability of not filing taxes in the immediate aftermath of the audit (same outcome as shown in panel (a) of figure 7). Our analysis relies on the monotonicity assumption. That is, being above the threshold does not decrease the probability of receiving an audit, and it does not increase the probability of filing taxes. Given our analysis up until this stage, both assumptions seem plausible on average. However, we recognize that it is possible some firms are more likely to file taxes after receiving an audit because they perceive to be under greater scrutiny from the tax authority (i.e. there may be some defiers).

We run a two-stage least squares regression whose first and second stages are given by:

$$h_f = \alpha \, \mathbb{1}\{z_f \geq l\} + \mathbb{1}\{z_f \geq l\}(z_f - l + 0.5) + z_f + \Gamma \mathbf{X}_f + \varepsilon_f \tag{4}$$

$$h_f \times k_f = \theta h_f + \mathbb{1}\{z_f \geq l\}(z_f - l + 0.5) + z_f + \Gamma \mathbf{X}_f + \varepsilon_f \tag{5}$$

Where h_f changes depending on which compliers we are looking at and k_f is the baseline characteristic of interest. Specifically, h_f is an indicator for either 1) whether the firm received a comprehensive audit, or 2) the probability of not filing taxes in the first year post-audit. The characteristics of the compliers are given by θ. Similar to equation 1, $\Gamma \mathbf{X}_f$ is a vector of controls including station fixed effects and separate indicators for whether the firm is in the VAT and CIT sample, l is the threshold for comprehensive audits, and z_f is the ordinal ranking of firms.

Table 5 reports the average characteristics of the two set of compliers, the sample mean, and p-values for tests of significance between each of the complier means and sample averages. Focusing on Column (2) and (3) firms induced into audits are somewhat larger than the rest of the sample. Firms induced into audits report significantly higher sales and purchases in the VAT (and marginally significant higher cost in CIT), are selling to more firms, and are more likely to be in the manufacturing sector. They also report higher number of employees, are more likely to file each of the tax returns, and report somewhat higher tax liabilities (but none of these are statistically significant). This is consistent with the URA using comprehensive audits to target large firms that pay several tax heads. The latter part underlines the importance of including controls for whether the firms are in the CIT or VAT sample at baseline.

Interestingly, the pattern is completely reversed for the firms that get induced to stop filing taxes after the audit. Column (4) and (5) show that firms induced to stop filing taxes are less connected to other firms (lower number of suppliers and buyers), employ fewer workers, and are more likely to be in the service sector. This is important since firms in the service sector are less likely to require machines or other physical capital to operate, making it less taxing to close down a firm or move location. The detailed nature of the CIT returns allow us to test for this explicitly by looking at the value of physical machines reported at baseline. Consistent with our interpretation, the value of machines and plants reported at baseline is significantly lower for firms induced to stop filing taxes. While they also report lower sales and purchases, their tax liabilities are comparable to the rest of the sample at baseline. Because their tax liabilities are comparable to the rest of the firms, they are an equally important contributor to revenue collection *despite* being smaller on average.

Overall we conclude that comprehensive audits induce the *marginal formal* firm – small formal firms that do not require a lot of employees or capital – to exit. Despite being at lower end of the firm size distribution of formal firms, these are of significant size compare to the overall firm size distribution. Naturally, most of these are found in the service sector. However, despite being small, this set of firms report substantial tax liabilities to the URA and are an important contributor to revenue collection reflected in their high tax liabilities.

6.5 Discussion

Results from the administrative tax data clearly indicate that comprehensive audits can "backfire": they cause firms to exit the formal sector and reduce revenue potential. However, our results also help rationalize why the URA still audits these firms. We find that comprehensive audits lead to an increase in tax liability of around 10 thousand USD from correcting evasion among marginally audited firms. If that translates to increase in collected revenue – something we cannot verify – these represent large short-term increase in revenue collection.

However, post-audit reporting behavior demonstrates that this comes at the cost of overall reduction in potential revenue collection in the aftermath of the audit. While the URA is raising 10,000 USD from correcting evasion among marginally audited firms, it is losing 6,500 thousand USD from firms reducing their CIT liabilities in the year after the audit, and 10,000 USD from a reduction in VAT liabilities. This suggests that the URA may be "killing the goose that lays the golden eggs" in an attempt to raise revenue in the short run. If they allow firms to evade they are still collecting some revenue from them. Encouraging compliance potentially leads to short-term revenue increases, but at the cost of subsequent decreases in revenue collection. The net effect, even ignoring the administrative costs of conducting the *marginal* audit is unambiguously negative.

Conducting a marginal value of public funds approach as done in for example Boning et al. (2023) would suggest that the welfare effect of the marginal tax audit is negative. The welfare effect of a tax audit in this framework is the ratio between the willingness to pay to avoid an audit by the taxpayer and the net revenue gain from the audit. Given that net revenue gain for the government from the marginal audit is negative and the willingness to pay to avoid an audit is positive, this immediately demonstrates that the welfare effect of the marginal audit is negative. This only speaks to the marginal audit and does not consider general deterrence – the idea that auditing one taxpayer may affect other taxpayers' behavior.

7 Tracking marginal firms through a survey

We have shown that firms exit the tax system post-audit and that tax liability decreases substantially in the post-audit period. While the effect on tax liability is unambiguously negative, this does not necessarily imply changes in firms' output. Changes in output are hard to measure with administrative tax data because information in tax returns will be influenced both by real production decisions and changes in reporting behavior. This is especially the case for the group of firms we study: designated to be likely tax evaders by the URA.

To overcome this challenge we design and implement a novel survey that tracks down firms on the margin of receiving a comprehensive tax audit. Due to our close collaboration with the URA we are able to link the survey with the administrative tax data. This allows us to investigate both whether firms close down and whether existing firms reduce their sales (reported in the survey) in

the aftermath of a comprehensive tax audit.

In this section we work with two sample. The "narrow sample" – used to measure whether firms shut down after an audit – is the sample of firms that identifies the effect on whether a firm files a tax return. This is the set of firms used in the regression in Figure 7(a). We invested extensive resources to determine whether these firms exist or closed down, making us confident that the classification is accurate. Importantly, we do not find differential survey response rates – defined as whether a firm agrees to be interviewed – above and below the discontinuity. We show in appendix figure A11 that the coefficient on probability of completing a survey is small (-0.018) and statistically insignificant.

Despite no differential response rates, we have limited statistical power to assess whether firms reduce their sales post-audit because we only interview 45% of the firms in the narrow sample. To gain power we move away from firms in the vicinity of the discontinuity and include all firms from the 13,014 firms in our final sample that were interviewed or that we know exited. In other words, we include all firms where we collected information on their sales or where we can confidently determine that they had no sales. In moving away from the discontinuity we lose the possibility to use the regression discontinuity as an identification strategy. We therefore leverage a difference-in-difference (DD) specification, combining the administrative and survey data. Specifically, we use the the administrative data to test for pre-trends and the survey data as the outcome in the post-audit period.

7.1 Results on firm exit

Our first key outcome is firm exit. We take a conservative approach defining a firm as exiting if it does not file CIT taxes for two years post-audit *and* either someone informs us that they closed or they "vanish".[34] This is important because firms might be "dormant", shutting down production and awaiting a profitable large contract – for example from the government – to restart. In such cases, firms keep filing taxes despite not operating. Combining the administrative tax data and the survey allows us to identify real exits. We define a firm as being informal if it does not file taxes post-audit, but we find evidence that it still exists. We choose this definition – as opposed to whether a firm is tax registered – because firms do not have an incentive to change their tax registration status when they leave the formal sector. As such, non-filing is the definition that most closely resembles what is used in the informality literature (Ulyssea, 2020).

Figure 8 suggests that the majority of firms that stop filing taxes shut down. In panel (a) we show that firms above the threshold are 4 p.p. more likely to have exited when we track them down approximately two years after the audit. The implied effect of receiving an audit is 7 p.p. (175%). Both the reduced form and two stage least squares estimates are significant at the 5% significance

[34]To verify that labelling vanished firms as closed is a reasonable assumption we compute the tax filing rates for each of our categories. In appendix table A11 we show that filing rates between closed and vanished firms are remarkably similar. 39% of closed firms file a CIT in the latest financial year and 39% of vanished firms do. The p-value for a t-test on whether the two averages are different the same is 0.992.

level. In panel (b) we show the effect on informality. The coefficient on the probability of being informal is also large, suggesting an increase in informality of 11 p.p. but this is only marginally significant. The coefficient on the probability of not filing a CIT return for two years post-audit is -0.20. Dividing the coefficient on exit and informal by the coefficient of not filing suggests that slightly more than half of firms (54%) exit, whereas the remaining 46% of firms that stop filing taxes keep operating informally.

Robustness. We conduct the same robustness checks for the survey data as we did when using administrative tax data. In appendix table A12 we show that the point estimates are similar when using different controls and in appendix figure A10 we show that the results are stable over narrow bandwidths.[35]

7.2 Causes of exit and subsequent occupations

Firm closures were identified in various ways. Sometimes the neighbors of the firms former premises told us the firm shut down, in which cases retrieving information on why the firm shut down was difficult. However, for the cases where we reached the former owner or a former employee we asked the respondent for the cause of the shut-down. We display the reasons mentioned in panel (a) of figure 9. While the number of observations is not high – and hence this should be interpreted with caution – note that the number of answers represents 63% of the firms we confirmed exited. The evidence suggests that the main reason for exit was taxes or troubles with the URA. 43% of firms mentioned high taxes or issues with the URA as the reason for why the firm exited. The next highest category was low demand at 16%.

Finally, we also asked the respondent what they were doing after the firm closed (or if we did not reach the firm, if they knew what the owners of the firm were doing now). We find that 39% of respondents reported being an employee in another firm. Crucially, less than half of respondent (40%) said they started a new firm, suggesting that the majority of firms that exit do not open again. We view this as suggestive evidence the audits are not only leading firms to close, but also discouraging the owner from starting a new firm in its aftermath.

7.3 Results on reduction in output

To assess whether output changes post-audit we combine the administrative and survey data for firms whose sales can be inferred from the survey. That is, firms who reported their sales in the survey or firms that closed or "vanished" and did not report any sales in their tax returns. This includes "dormant" companies as well.[36] We use administrative tax data for the pre-audit period

[35] The only time the coefficient changes by more than 20% is in Column (6), where we add sector fixed effects and a full set of risk controls (70 total). This is an extremely demanding specification leading us to conclude that the results are robust.

[36] This removes 133 (14%) of the sample. These were firms where the respondent was afraid or not allowed to disclose financial information (100) or firms that were mistakenly labeled as closed (33).

and run two regressions: one where we only use administrative tax data and one where we substitute for the administrative tax data with the survey data in the last financial year. We run the following DD specification:

$$y_{ft} = \beta_{DD} \, \mathbb{1}\{z_f \geq l\} \times \mathbb{1}\{t = 2022\} + \alpha_f + \delta_t + \varepsilon_{ft} \tag{6}$$

Where y_{ft} is the outcome of interest of firm f measured at time t. z_f is the ranking of firm f used for selecting comprehensive audits, and l is the threshold for the assignment of comprehensive audits. α_f and δ_t are firm and year fixed effects, respectively.

If the parallel trends assumption holds – that is, in the absence of the comprehensive audit, the outcomes of firms would have evolved on parallel trends – β_{DD} captures the average treatment effect on the treated. To test whether the parallel trends assumption holds, we run a two-way fixed effect version of equation 6, which is formally written below

$$y_{ft} = \beta_k \sum_{\substack{2013 \\ k \neq 2019}}^{2022} \mathbb{1}\{z_f^c \geq l^c\} \times \mathbb{1}\{t = k\} + \alpha_f + \delta_t + \varepsilon_{ft} \tag{7}$$

β_k captures the dynamic treatment effect of being above the threshold on the outcome of interest. All event-study results are shown in figure 10. We do not find evidence of pre-trends for any of the specifications.

A separate concern is what constitutes the post-period. As discussed extensively in section 6 firms file CIT returns retrospectively. The first year in which all firms are unambiguously filing taxes for a period after the audit is therefore 2022. We therefore exclude 2020 and 2021 from the analysis altogether. This is because these years may be affected by the audit, but to a lesser extent, mechanically driving the treatment effect towards zero.

Table 6 presents results on sales using the DD specification. In Column (1) – (5) we show results using only the administrative data, whereas in Column (6) – (10) we show results where we replace the post-audit administrative information with the survey data. Throughout we use the same set of firms. That is, firms that reported their sales to us in the survey or where we can discern that sales were zero in the last financial year.

Focusing on the results using only the administrative tax data, we find a 20.8 p.p. drop in the likelihood of reporting any sales to the tax authority. Note that this is broadly in line with the reduced form coefficient on the probability of filing a tax return in Figure 7(b) (-0.23). For the firms that do submit sales information to the tax authority we observe a drop in the probability of reporting positive sales of 19.3 p.p. (Column (2)), in line with the reduction in the probability of having a positive tax liability, discussed in section 6 (-0.18). Using the logarithmic transformation of sales, effectively conditioning on firms reporting sales in their tax return in 2022, we find no effect on the amount of sales reported (Column (3)). The coefficient is small (0.017) and statistically

insignificant. However, if we instead impute zeros for firms not filing taxes, we find large and significant reductions in total sales (Column (5)). The coefficient in Column (5) suggests a decrease in the average sales of 258 thousand USD. Note that this is largely driven by the firms not filing taxes. If we exclude firms that did not file taxes the coefficient reduces to 118 thousands USD and becomes statistically insignificant (Column (4)).

The story changes significantly when we leverage information from the survedy data for the same set of firms. Because we restricted to firms where we can infer sales throughout the analysis, there is no significant differences in the likelihood of reporting sales (Column (6)). More interestingly, the reduction in the probability of reporting positive sales attenuates from 19.3 p.p. (Column (2)) to 6.2 p.p. (Column (7)) and is only marginally significant. In other words, while audited firms are less likely to file taxes or report positive sales in their tax return (conditional on filing taxes), they are only marginally less likely to report positive sales in the survey, highlighting the importance of combining administrative and survey data.

We next turn to the amount of sales reported. Using the logarithmic transformation of sales (Column (8)) we find that firms above the discontinuity reduce their sales by 0.842 log points (56%). This is entirely based on survey data and hence controls for any changes in evasion. The results are significant at the 1% level. This suggests that firms reduce their output after an audit, but instead of reporting a reduction in output in their tax returns, firms stop filing taxes or reporting sales altogether. Finally, we include firms that exited and estimate the overall effect on sales. Interestingly, our effect on overall sales using the survey data is similar to what we find when using the administrative data. Using the survey data we find a reduction in overall sales of 267 thousand USD on average (Column (10)), while we find a reduction in sales of 258 thousand USD using the administrative data (Column (5)). It suggests that the extensive margin identified in the administrative data closely resembles the intensive margin in the survey data in terms of overall magnitude on sales. This is not a given, and we view this as a coincidence.

In appendix table, A13 we show results for sales leveraging the RD specification. Like in table 6, we include exiters and dormant firms imputing zero's, and leverage both the winsorized and log transformation of the amounts. Standard errors are large making it hard to discern what is going on. The RD coefficient on the logarithmic transformation of sales (Column (2)) suggest a reduction of 49%, which is similar to what we found in the DD specification (57%). We also included other measures of firm size, such as costs and total number of employees in our analysis, shown in Column (4) – (6). Our estimates suggests a reduction in costs of 49% and a reduction in the number of employees of 7.4 workers. However, none of these are statistically significant. Overall, the RD analysis corroborates what we found in the DD analysis: firms that remain operational are reducing their size.

These results are striking. Not only do audits reduce the tax liability of firms in the post-audit period, but they also affect the output reported by firms. There are two key questions remaining:

are these effects relevant at the aggregate level? What part of the audit process is driving these results, i.e. what is the mechanism behind the effect we find?

We start by estimating the effect of audits on aggregate output.

8 Aggregation

The results from our survey demonstrates that comprehensive audits do not only reduce the tax liability of firms, but also cause them to reduce output. A central remaining question is whether this has any implications on aggregate output.

Instead of estimating a model we leverage theoretical results from XXX, that show that to a first oder, the effect of an intervention on aggregate output is given by the following equation:

$$\sum_i [sales_i - cost_i] \times \Delta \log inputs_i$$

9 Mechanism

We have documented that comprehensive tax audits have a negative effect on tax liabilities, firm output and that the reduction in firm output is relevant at the aggregate level. In this section we explore the mechanism behind our results.

9.1 A simple firm problem and potential mechanisms

There are several potential mechanisms that could be driving firm behavior, each of which would have different policy implications. To guide our discussion we introduce a simple firm optimization problem. The model adds audits to a simplified version of the model used in Best et al. (2015).

Assume there is a revenue tax τ, that firms produce output y at a strictly convex and differentiable cost $c(y)$. They can choose to engage in evasion $e = y - z$, where z is the amount of revenue declared to the tax authority. Evasion is not for free, but instead firms face a convex, increasing and differentiable cost $g(e)$ of evading. If they get audited they have to pay back the tax liability of everything they evaded at interest rate r. The firm chooses y and e to maximize after-tax profits taking into account the perceived probability of being audited $\hat{p}(y)$, which depends on firm output. That is,

$$\max_{y,e} \pi_f = (1-\tau)y - c(y) + \tau e - g(e) - \hat{p}(y)\left[(1+r)\tau e + c_a\right] \tag{8}$$

Yielding the following optimal evasion and output decision,

$$g'(e) = \tau\left[1 - \hat{p}(y)(1+r)\right] \tag{9}$$

$$c'(y) = 1 - \underbrace{\left[\tau + \hat{p}'(y)\left[(1+r)\tau e + c_a\right]\right]}_{\text{tax and audit induced distortion}} \tag{10}$$

The intuition behind equation 9 is that firms will set the marginal cost of evasion $g'(e)$ equal to the marginal gain from τ minus the expected additional cost of paying back what it evaded $\hat{p}(y)(1+r)\tau$. Similarly, the firm will set the marginal cost of production $c'(y)$ equal to the marginal benefit of production. Note that if the perceived audit probability does nto depend on firm size, $\hat{p}'(y) = 0$, then the only feature preventing from setting $c'(y) = 1$ is the tax rate. The simple choice model captures three potential channels through which audits may affect firm behavior.

First, if a firm gets audited, the amount of taxes the firms has to repay may be too high. That is, $(1+r)\tau e$ might be so high that $\pi_f < 0$ forcing the firm to exit. This would especially acute in the presence of credit constraints, which are highly like to be a feature in the Ugandan context.

Second, if a firm gets an audit, conducting an audit might put an onerous financial burden on the firm. That is, c_a is sufficiently high such that $\pi_f < 0$. Note that this channel does not include the taxes a firm has to repay, but focuses on the "hassle" cost of an audit. i.e. the fixed cost to the firm of an audit *regardless* of whether it evaded or not.

The third mechanism points to the central role of firms believes about the probability of being audited and how it varies with firm size. Receiving an audit may change $\hat{p}$, according to equation 9 that would affect the extent to which the firm evades. If $\hat{p}$ is revised upwards, by the first order condition, the firm evades less leading to a higher effective tax rate. Under the updated perception on audit rates firms may not think they are profitable leading them to exit. However, receiving an audit may also change $\hat{p}'(y)$, that is, firms might adjust their believes about to what extent audit rates vary with firm output. If $\hat{p}'(y)$ is revised upwards following an audit, by equation 10 firms will reduce their output to reduce the possibility of future audits.

Finally, tax morale – broadly defined – might also play a key role but is mostly not captured by the firm problem above. Luttmer and Singhal (2014) identify five channels through which tax morale may affect compliance: 1) *intrinsic motivation*, capturing the idea that someone is intrinsically motivated to pay taxes, 2) *reciprocity* individuals pay taxes because they believe they get something in return, 3) *peer effect and social influence*, the behavior of other people you know may affect your actions, 4) *cultural factors* that may affect willingness to pay taxes, and 5) *information imperfections and deviation from utility maximization*. The last one, 5), is partially captured by us including a perceived audit probability which may be incorrect.

9.2 Empirical evidence

Having discussed the various channels we combine descriptive and causal evidence to understand which ones may be driving the results. Recall there are two key results we are interested in explaining: 1) why firms exited, and 2) why firms reduced their output.

In our survey we asked the firms that closed their business why they closed. The most frequently mentioned reason by far was high taxes. Concern about high taxes could either come from high tax corrections – channel one – or increases in the perceived effective tax rate – channel three.

How we can test for the presence of each mechanism and, ideally, differentiate which one is driving the results? Let us first abstract from tax morale and focus on the channels represented by the firm choice model.

If the main mechanisms is the first channel, we would expect the firms exiting to be the ones that received the highest tax corrections as a share of baseline sales. This is testable using the administrative tax data.

If it is the second channel, we would expect firms that report high hassle cost to be driving the result. Sadly, we did not interview firms that already closed hence we canno

10 Conclusion

In this paper we investigate the effect of increasing enforcement on firms – through comprehensive audits – in a context with limited administrative capacity, Uganda. We show that increasing enforcement for the *marginal* firm backfires. Overall tax liability is reduced due to changes in firm operations during the post-audit period.

These results are an important reminder that tax enforcement is not costless. Indeed, the focus in the tax audit literature has turned towards thinking about which taxpayers it is most beneficial to audit (Boning et al., 2023). The underlying assumption is that audits do not impose large economic distortions on any of the taxpayers in question. Our result show that this assumption is violated in the context of a developing country, Uganda. Given the magnitude of our results, we argue that understanding the economic cost imposed on taxpayers is of primary importance when thinking about which taxpayers to audit (or whether to audit at all).

We show that the reduction in revenue reflects real economic distortions, not aggressive tax planning. Two-thirds of the firms that stop filing taxes shut down post-audit. The ones that keep filing taxes reduce their scale substantially.

A compliers analysis reveals that the firms that stop filing taxes are smaller, require less capital and are primarily situated in the service sector. This highlights a daunting challenge for tax authorities in developing countries: how do you increase enforcement on a group of firms that can easily move locations to avoid the increased pressure from the tax authority?

References

Abadie, Alberto (2002). "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models". In: *Journal of the American Statistical Association* 97.457, pp. 284–292. ISSN: 0162-1459. JSTOR: 3085782. (Visited on 07/16/2024).

Advani, Arun, William Elming, and Jonathan Shaw (2021). "The Dynamic Effects of Tax Audits". In: *The Review of Economics and Statistics*, pp. 1–45. ISSN: 0034-6535. DOI: 10.1162/rest_a_01101. (Visited on 03/28/2022).

Akcigit, Ufuk, Harun Alp, and Michael Peters (2021). "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries". In: *American Economic Review* 111.1, pp. 231–275. ISSN: 0002-8282. DOI: 10.1257/aer.20180555. (Visited on 10/18/2024).

Alm, James and Antoine Malézieux (2021). "40 Years of Tax Evasion Games: A Meta-Analysis". In: *Experimental Economics* 24.3, pp. 699–750. ISSN: 1573-6938. DOI: 10.1007/s10683-020-09679-3. (Visited on 07/14/2024).

Almunia, Miguel, Justine Knebelmann, Lin Tian, David Henning, and Dorothy Nakyambadda (2022). *Information, Fiscal Capacity, and Tax Enforcement: An Experimental Evaluation*. Tech. rep.

– (2023). *Leveraging Trading Networks to Improve Tax Compliance: Experimental Evidence from Uganda*.

Bachas, Pierre, Anne Brockmeyer, Roel Dom, and Camille Semelet (2023). *Effective Tax Rates and Firm Size*. Tech. rep. The World Bank. DOI: 10.1596/1813-9450-10312. (Visited on 08/18/2024).

Bachas, Pierre, Roberto N. Fattal Jaef, and Anders Jensen (2019). "Size-Dependent Tax Enforcement and Compliance: Global Evidence and Aggregate Implications". In: *Journal of Development Economics* 140, pp. 203–222. ISSN: 0304-3878. DOI: 10.1016/j.jdeveco.2019.05.001. (Visited on 09/06/2024).

Basri, M. Chatib, Mayara Felix, Rema Hanna, and Benjamin A. Olken (2021). "Tax Administration versus Tax Rates: Evidence from Corporate Taxation in Indonesia". In: *American Economic Review* 111.12, pp. 3827–3871. ISSN: 0002-8282. DOI: 10.1257/aer.20201237. (Visited on 03/28/2022).

Bassi, Vittorio, Jung Hyuk Lee, Alessandra Peter, Tommaso Porzio, Ritwika Sen, and Esau Tugume (2023). *Self-Employment Within the Firm*. Working Paper. DOI: 10.3386/w31740. National Bureau of Economic Research: 31740. (Visited on 10/18/2024).

Bassi, Vittorio and Aisha Nansamba (2022). "Screening and Signalling Non-Cognitive Skills: Experimental Evidence from Uganda". In: *The Economic Journal* 132.642, pp. 471–511. ISSN: 0013-0133. DOI: 10.1093/ej/ueab071. (Visited on 09/29/2024).

Beer, Sebastian, Matthias Kasper, Erich Kirchler, and Brian Erard (2020). "Do Audits Deter or Provoke Future Tax Noncompliance? Evidence on Self-Employed Taxpayers". In: *CESifo Economic Studies* 66.3, pp. 248–264. ISSN: 1610-241X. DOI: 10.1093/cesifo/ifz018. (Visited on 03/28/2022).

Belnap, Andrew, Jeffrey L. Hoopes, Edward L. Maydew, and Alex Turk (2024). "Real Effects of Tax Audits". In: *Review of Accounting Studies* 29.1, pp. 665–700. ISSN: 1573-7136. DOI: 10.1007/s11142-022-09717-w. (Visited on 03/08/2024).

Bento, Pedro and Diego Restuccia (2021). "On Average Establishment Size across Sectors and Countries". In: *Journal of Monetary Economics* 117, pp. 220–242. ISSN: 0304-3932. DOI: 10.1016/j.jmoneco.2020.01.001. (Visited on 09/08/2024).

Bergeron, Augustin, Gabriel Tourek, and Jonathan L. Weigel (2024). "The State Capacity Ceiling on Tax Rates: Evidence From Randomized Tax Abatements in the DRC". In: *Econometrica* 92.4, pp. 1163–1193. ISSN: 1468-0262. DOI: 10.3982/ECTA19959. (Visited on 08/29/2024).

Besley, Timothy and Torsten Persson (2013). "Chapter 2 - Taxation and Development". In: *Handbook of Public Economics*. Ed. by Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez. Vol. 5. Handbook of Public Economics, Vol. 5. Elsevier, pp. 51–110. DOI: 10.1016/B978-0-444-53759-1.00002-9. (Visited on 09/11/2022).

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem (2015). "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan". In: *Journal of Political Economy* 123.6, pp. 1311–1355. ISSN: 0022-3808. DOI: 10.1086/683849. (Visited on 01/04/2024).

Best, Michael Carlos, Jawad Shah, and Mazhar Waseem (2021). *Detection Without Deterrence: Long-Run Effects of Tax Audit on Firm Behavior*. Tech. rep., p. 91.

Blattman, Christopher, Eric P. Green, Julian Jamison, M. Christian Lehmann, and Jeannie Annan (2016). "The Returns to Microenterprise Support among the Ultrapoor: A Field Experiment in Postwar Uganda". In: *American Economic Journal: Applied Economics* 8.2, pp. 35–64. ISSN: 1945-7782. DOI: 10.1257/app.20150023. (Visited on 09/29/2024).

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012). "The Organization of Firms Across Countries*". In: *The Quarterly Journal of Economics* 127.4, pp. 1663–1705. ISSN: 0033-5533. DOI: 10.1093/qje/qje029. (Visited on 10/18/2024).

Boehm, Johannes and Ezra Oberfield (2020). "Misallocation in the Market for Inputs: Enforcement and the Organization of Production*". In: *The Quarterly Journal of Economics* 135.4, pp. 2007–2058. ISSN: 0033-5533. DOI: 10.1093/qje/qjaa020. (Visited on 09/29/2024).

Boeri, Tito and Juan F. Jimeno (2005). "The Effects of Employment Protection: Learning from Variable Enforcement". In: *European Economic Review* 49.8, pp. 2057–2077. ISSN: 0014-2921. DOI: 10.1016/j.euroecorev.2004.09.013. (Visited on 10/18/2024).

Bold, Tessa, Selene Ghisolfi, Frances Nsonzi, and Jakob Svensson (2022). "Market Access and Quality Upgrading: Evidence from Four Field Experiments". In: *American Economic Review* 112.8, pp. 2518–2552. ISSN: 0002-8282. DOI: 10.1257/aer.20210122. (Visited on 09/29/2024).

Boning, William C., Nathaniel Hendren, Ben Sprung-Keyser, and Ellen Stuart (2023). *A Welfare Analysis of Tax Audits Across the Income Distribution*. Working Paper. DOI: 10.3386/w31376. National Bureau of Economic Research: 31376. (Visited on 07/11/2023).

Busso, Matias, Sebastián Montaño, and Juan Muñoz-Morales (2023). *Signaling Specific Skills and the Labor Market of College Graduates*. Discussion Paper Series 16449. IZA Discussion Papers.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". In: *Econometrica* 82.6, pp. 2295–2326. ISSN: 1468-0262. DOI: 10.3982/ECTA11757. (Visited on 07/18/2023).

Carranza, Eliana, Robert Garlick, Kate Orkin, and Neil Rankin (2022). "Job Search and Hiring with Limited Information about Workseekers' Skills". In: *American Economic Review* 112.11, pp. 3547–3583. ISSN: 0002-8282. DOI: 10.1257/aer.20200961. (Visited on 09/29/2024).

Carrillo, Paul, Dave Donaldson, Dina Pomeranz, and Monica Singhal (2023). "Ghosting the Tax Authority: Fake Firms and Tax Fraud in Ecuador". In: *American Economic Review: Insights* 5.4, pp. 427–444. ISSN: 2640-205X. DOI: 10.1257/aeri.20220321. (Visited on 08/20/2024).

Carrillo, Paul, Dina Pomeranz, and Monica Singhal (2017). "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement". In: *American Economic Journal: Applied Economics* 9.2, pp. 144–164. ISSN: 1945-7782. DOI: 10.1257/app.20140495. (Visited on 03/28/2022).

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma (2018). "Manipulation Testing Based on Density Discontinuity". In: *The Stata Journal* 18.1, pp. 234–261. ISSN: 1536-867X. DOI: 10.1177/1536867X1801800115. (Visited on 07/18/2023).

– (2020). "Simple Local Polynomial Density Estimators". In: *Journal of the American Statistical Association* 115.531, pp. 1449–1455. ISSN: 0162-1459. DOI: 10.1080/01621459.2019.1635480. (Visited on 10/11/2022).

Chandra, Piyush and Cheryl Long (2013). "VAT Rebates and Export Performance in China: Firm-level Evidence". In: *Journal of Public Economics* 102, pp. 13–22. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2013.03.005. (Visited on 10/17/2024).

Christiansen, Tobias Gabel (2024). "Dynamic Effects of Tax Audits and the Role of Intentions". In: *Journal of Public Economics* 234, p. 105121. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2024.105121. (Visited on 09/13/2024).

Corbellini, Nicola (2024). "The Effects of Tax Enforcement on the Firm Size Distribution and Aggregate Productivity".

de Mel, Suresh, David McKenzie, and Christopher Woodruff (2008). "Returns to Capital in Microenterprises: Evidence from a Field Experiment*". In: *The Quarterly Journal of Economics* 123.4, pp. 1329–1372. ISSN: 0033-5533. DOI: 10.1162/qjec.2008.123.4.1329. (Visited on 09/29/2024).

– (2014). "Business Training and Female Enterprise Start-up, Growth, and Dynamics: Experimental Evidence from Sri Lanka". In: *Journal of Development Economics* 106, pp. 199–210. ISSN: 0304-3878. DOI: 10.1016/j.jdeveco.2013.09.005. (Visited on 09/29/2024).

DeBacker, Jason, Bradley T. Heim, Anh Tran, and Alexander Yuskavage (2015). "Legal Enforcement and Corporate Behavior: An Analysis of Tax Aggressiveness after an Audit". In: *The Journal of Law and Economics* 58.2, pp. 291–324. ISSN: 0022-2186. DOI: 10.1086/684037. (Visited on 03/28/2022).

DeBacker, Jason, Bradley T. Heim, Anh Tran, and Alexander Yuskavage (2018). "Once Bitten, Twice Shy? The Lasting Impact of Enforcement on Tax Compliance". In: *The Journal of Law and Economics* 61.1, pp. 1–35. ISSN: 0022-2186. DOI: 10.1086/697683. (Visited on 03/28/2022).

Di Nola, Alessandro, Georgi Kocharkov, Almuth Scholl, and Anna-Mariia Tkhir (2021). "The Aggregate Consequences of Tax Evasion". In: *Review of Economic Dynamics* 40, pp. 198–227. ISSN: 1094-2025. DOI: 10.1016/j.red.2020.09.009. (Visited on 09/29/2024).

Enache, Cristina (2023). *Corporate Tax Rates around the World, 2023*. https://taxfoundation.org/data/all/global/co tax-rates-by-country-2023/. (Visited on 07/14/2024).

Evans, David S. (1986). "The Differential Effect of Regulation across Plant Size: Comment on Pashigian". In: *The Journal of Law & Economics* 29.1, pp. 187–200. ISSN: 0022-2186. JSTOR: 725408. (Visited on 10/18/2024).

Field, Erica, Seema Jayachandran, and Rohini Pande (2010). "Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India". In: *American Economic Review* 100.2, pp. 125–129. ISSN: 0002-8282. DOI: 10.1257/aer.100.2.125. (Visited on 09/29/2024).

Gadenne, Lucie, Tushar K. Nandi, and Roland Rathelot (2022). *Taxation and Supplier Networks: Evidence from India*. Tech. rep.

Garicano, Luis, Claire Lelarge, and John Van Reenen (2016). "Firm Size Distortions and the Productivity Distribution: Evidence from France". In: *American Economic Review* 106.11, pp. 3439–3479. ISSN: 0002-8282. DOI: 10.1257/aer.20130232. (Visited on 10/18/2024).

Gemmell, Norman and Marisa Ratto (2012). "Behavioral Responses to Taxpayer Audits: Evidence from Random Taxpayer Inquiries". In: *National Tax Journal* 65.1, pp. 33–57. ISSN: 0028-0283. DOI: 10.17310/ntj.2012.1.02. (Visited on 03/28/2022).

Gerardino, Maria Paula, Stephan Litschig, and Dina Pomeranz (2024). "Distortion by Audit: Evidence from Public Procurement". In: *American Economic Journal: Applied Economics* 16.4, pp. 71–108. ISSN: 1945-7782. DOI: 10.1257/app.20220512. (Visited on 10/18/2024).

Gourio, François and Nicolas Roys (2014). "Size-Dependent Regulations, Firm Size Distribution, and Reallocation". In: *Quantitative Economics* 5.2, pp. 377–416. ISSN: 1759-7331. DOI: 10.3982/QE338. (Visited on 10/18/2024).

Harju, Jarkko, Kaisa Kotakorpi, Tuomas Matikka, and Annika Nivala (2024). "How Do Firms Respond to Risk-based Tax Audits?" In.

Harju, Jarkko, Tuomas Matikka, and Timo Rauhanen (2019). "Compliance Costs vs. Tax Incentives: Why Do Entrepreneurs Respond to Size-Based Regulations?" In: *Journal of Public Economics* 173, pp. 139–164. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2019.02.003. (Visited on 10/17/2024).

Hjort, Jonas, Golvine de Rochambeau, Vinayak Iyer, and Fei Ao (2020). *Informational Barriers to Market Access: Experimental Evidence from Liberian Firms*. Working Paper. DOI: 10.3386/w27662. National Bureau of Economic Research: 27662. (Visited on 09/29/2024).

Hsieh, Chang-Tai and Peter J. Klenow (2009). "Misallocation and Manufacturing TFP in China and India*". In: *The Quarterly Journal of Economics* 124.4, pp. 1403–1448. ISSN: 0033-5533. DOI: 10.1162/qjec.2009.124.4.1403. (Visited on 10/15/2024).

ILO (2021). *Informal Employment Rate by Sex (%)*. https://rshiny.ilo.org/dataexplorer31/?lang=en&id=EMP_TE (Visited on 08/13/2024).

Imbens, Guido W. and Donald B. Rubin (1997). "Estimating Outcome Distributions for Compliers in Instrumental Variables Models". In: *The Review of Economic Studies* 64.4, pp. 555–574. ISSN: 0034-6527. DOI: 10.2307/2971731. (Visited on 07/15/2024).

Iyer, Rajkamal and Antoinette Schoar (2015). "Ex Post (In) Efficient Negotiation and Breakdown of Trade". In: *American Economic Review* 105.5, pp. 291–294. ISSN: 0002-8282. DOI: 10.1257/aer.p20151077. (Visited on 09/29/2024).

Jensen, Anders, Anne Brockmeyer, and Lucie Gadenne (2024). *Taxation and Development*. Tech. rep. 12.

Jensen, Robert and Nolan H. Miller (2018). "Market Integration, Demand, and the Growth of Firms: Evidence from a Natural Experiment in India". In: *American Economic Review* 108.12, pp. 3583–3625. ISSN: 0002-8282. DOI: 10.1257/aer.20161965. (Visited on 09/28/2024).

Kasper, Matthias and James Alm (2022). "Audits, Audit Effectiveness, and Post-Audit Tax Compliance". In: *Journal of Economic Behavior & Organization* 195, pp. 87–102. ISSN: 0167-2681. DOI: 10.1016/j.jebo.2022.01.003. (Visited on 07/09/2024).

Kasper, Matthias and Matthew D. Rablen (2023). "Tax Compliance after an Audit: Higher or Lower?" In: *Journal of Economic Behavior & Organization* 207, pp. 157–171. ISSN: 0167-2681. DOI: 10.1016/j.jebo.2023.01.013. (Visited on 07/09/2024).

Kennedy, Patrick J, Paul Landefeld, Christine L Dobridge, and Jacob Mortenson (2024). *The Efficiency-Equity Tradeoff of the Corporate Income Tax: Evidence from the Tax Cuts and Jobs Act*. Tech. rep.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez (2011). "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark". In: *Econometrica* 79.3, pp. 651–692. ISSN: 1468-0262. DOI: 10.3982/ECTA9113. (Visited on 03/28/2022).

Kotsogiannis, Christos, Luca Salvadori, John Karangwa, and Theonille Mukamana (2024). "Do Tax Audits Have a Dynamic Impact? Evidence from Corporate Income Tax Administrative Data". In: *Journal of Development Economics* 170, p. 103292. ISSN: 0304-3878. DOI: 10.1016/j.jdeveco.2024.103292. (Visited on 07/07/2024).

Leal Ordóñez, Julio César (2014). "Tax Collection, the Informal Sector, and Productivity". In: *Review of Economic Dynamics* 17.2, pp. 262–286. ISSN: 1094-2025. DOI: 10.1016/j.red.2013.07.004. (Visited on 09/29/2024).

Li, Wanfu, Jeffrey A. Pittman, and Zi-Tian Wang (2018). "The Determinants and Consequences of Tax Audits: Some Evidence from China". In: *Journal of the American Taxation Association* 41.1, pp. 91–122. ISSN: 0198-9073. DOI: 10.2308/atax-52136. (Visited on 03/28/2022).

Liu, Li, Ben Lockwood, Miguel Almunia, and Eddy H. F. Tam (2021). "VAT Notches, Voluntary Registration, and Bunching: Theory and U.K. Evidence". In: *The Review of Economics and Statistics* 103.1, pp. 151–164. ISSN: 0034-6535. DOI: 10.1162/rest_a_00884. (Visited on 10/17/2024).

Løyland, Knut, Oddbjørn Raaum, Gaute Torsvik, and Arnstein Øvrum (2019). *Compliance Effects of Risk-Based Tax Audits*. Tech. rep. 7616. CESifo. (Visited on 03/28/2022).

Luttmer, Erzo F. P. and Monica Singhal (2014). "Tax Morale". In: *Journal of Economic Perspectives* 28.4, pp. 149–168. ISSN: 0895-3309. DOI: 10.1257/jep.28.4.149. (Visited on 10/21/2024).

McKenzie, David and Yaye Seynabou Sakho (2010). "Does It Pay Firms to Register for Taxes? The Impact of Formality on Firm Profitability". In: *Journal of Development Economics* 91.1, pp. 15–24. ISSN: 0304-3878. DOI: 10.1016/j.jdeveco.2009.02.003. (Visited on 09/29/2024).

McKenzie, David and Christopher Woodruff (2008). "Experimental Evidence on Returns to Capital and Access to Finance in Mexico". In: *The World Bank Economic Review* 22.3, pp. 457–482. ISSN: 0258-6770. DOI: 10.1093/wber/lhn017. (Visited on 09/29/2024).

OECD, African Union Commission, and African Tax Administration Forum (2023). *Revenue Statistics in Africa 2023*, p. 388. DOI: https://doi.org/10.1787/15bc5bc6-en-fr.

OECD and ILO (2019). *Tackling Vulnerability in the Informal Economy*.

Pinotti, Paolo (2017). "Clicking on Heaven's Door: The Effect of Immigrant Legalization on Crime". In: *American Economic Review* 107.1, pp. 138–168. ISSN: 0002-8282. DOI: 10.1257/aer.20150355. (Visited on 07/08/2024).

Pomeranz, Dina (2015). "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax". In: *American Economic Review* 105.8, pp. 2539–2569. ISSN: 0002-8282. DOI: 10.1257/aer.20130393. (Visited on 03/28/2022).

Quinn, Simon and Christopher Woodruff (2019). "Experiments and Entrepreneurship in Developing Countries". In: *Annual Review of Economics* 11.Volume 11, 2019, pp. 225–248. ISSN: 1941-1383, 1941-1391. DOI: 10.1146/annurev-economics-080218-030246. (Visited on 09/28/2024).

Sanday, Amos (2023). *Unlocking the Potential of the Informal Sector in Uganda through Business Formalisation*. (Visited on 08/13/2024).

Schivardi, Fabiano and Roberto Torrini (2008). "Identifying the Effects of Firing Restrictions through Size-Contingent Differences in Regulation". In: *Labour Economics* 15.3, pp. 482–511. ISSN: 0927-5371. DOI: 10.1016/j.labeco.2007.03.003. (Visited on 10/18/2024).

Slemrod, Joel (2019). "Tax Compliance and Enforcement". In: *Journal of Economic Literature* 57.4, pp. 904–954. ISSN: 0022-0515. DOI: 10.1257/jel.20181437. (Visited on 04/01/2022).

The Independent (2023). *Supreme Court Dismisses Heritage Appeal in $435m Oil Tax Row*. (Visited on 06/21/2024).

Ulyssea, Gabriel (2018). "Firms, Informality, and Development: Theory and Evidence from Brazil". In: *American Economic Review* 108.8, pp. 2015–2047. ISSN: 0002-8282. DOI: 10.1257/aer.20141745. (Visited on 10/06/2022).

– (2020). "Informality: Causes and Consequences for Development". In: *Annual Review of Economics* 12.1, pp. 525–546. DOI: 10.1146/annurev-economics-082119-121914. (Visited on 04/16/2023).

Vitali, Anna (2023). *Consumer Search and Firm Location: Theory and Evidence from the Garment Sector in Uganda*. Tech. rep.

Waseem, Mazhar (2023). "Overclaimed Refunds, Undeclared Sales, and Invoice Mills: Nature and Extent of Noncompliance in a Value-Added Tax". In: *Journal of Public Economics* 218, p. 104783. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2022.104783. (Visited on 08/20/2024).

World Bank (2017). *Informal Employment by Gender (% of Total Non-Agricultural Employment)*. https://databank.worldbank.org/source/gender-statistics/Series/SL.ISV.IFRM.FE.ZS#advancedDownloadOpti (Visited on 08/13/2024).

– (2019a). *Global Economic Prospects, January 2019: Darkening Skies*. Washington, DC: World Bank. ISBN: 978-1-4648-1343-6 978-1-4648-1386-3. DOI: 10.1596/978-1-4648-1343-6. (Visited on 08/13/2024).

– (2019b). *National Manpower Survey 2016, Kampala Informal Sector Survey*. DOI: 10.48529/8H55-GF47.

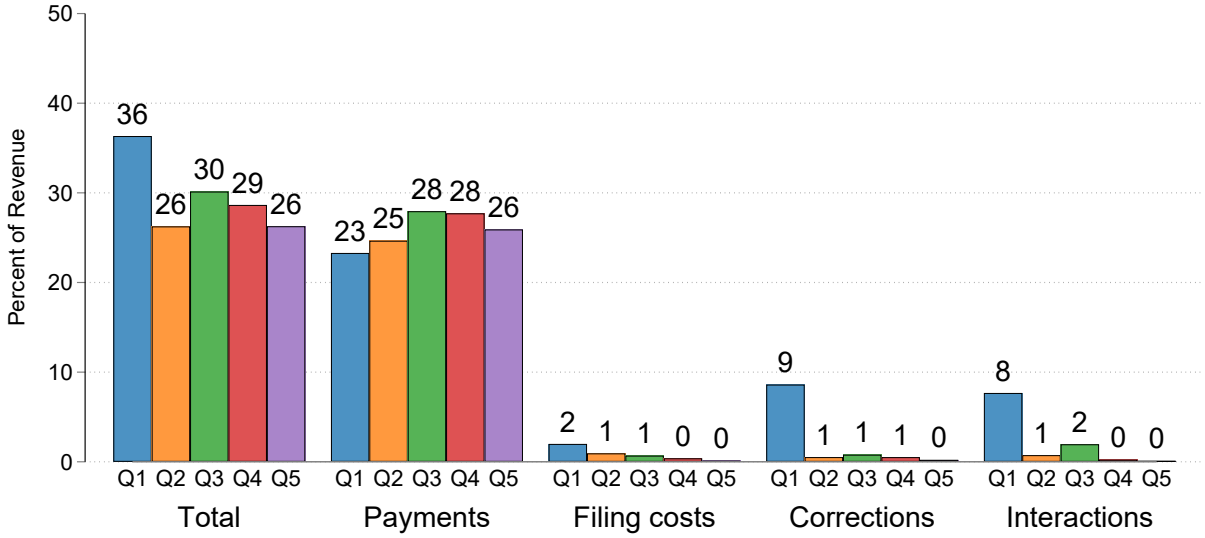
– (2023a). *Enterprise Surveys*.

– (2023b). *GNI per Capita in PPP (Current International Dollars)*. https://data.worldbank.org. (Visited on 07/08/2024).

Tables & Figures

Figures

Figure 1
Costs of filing taxes as a share of revenue

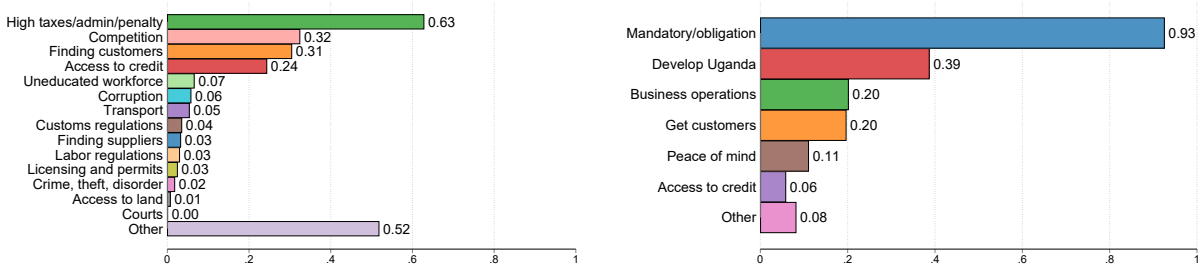


Notes: This figure displays both the total and different average costs associated with filing taxes as a share of revenue across the quintile of the firm-size distribution. Size is measured by revenue reported for the last accounting year. "Payments" are total tax payments across all tax heads. "Filing costs" are the total administrative costs (internal and outsourced) of filing taxes. "Corrections" is the average value of the assessments filed by the URA over the last two accounting years. "Interactions" are all the costs of interacting with the URA, including audits. "Total costs" is the sum across all of them. Each term is presented as a share of the firms revenue. We restrict to firms that reported revenues, tax payments, and either reported the correction value or did not report receiving a correction. Sample consists of 528 firms. Source: Authors' survey and computation.

Figure 2
Obstacles to firm performance and reasons for paying taxes

(a) Top three obstacles to firm growth

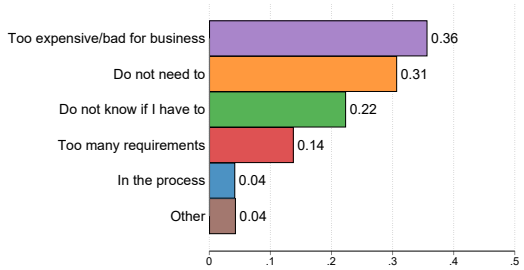
(b) Top three reasons for paying taxes



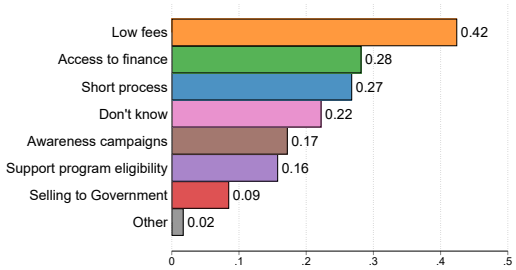
Notes: This figure displays the top three obstacles for firm performance and the top three reasons for paying taxes. We aggregate across each of the three reasons mentioned. In *Panel (a)* we display the top three obstacles to firm performance. We aggregate across the three options related to taxes and the URA, namely: "High taxes", "High administrative costs of taxes", and "Tax assessments". N = 754. In *Panel (b)* we display the top three reasons for paying taxes mentioned by firms. N = 767. Source: Authors' survey and computation.

Figure 3
Reasons for not formalizing among informal firms

(a) Why is the firm not registered

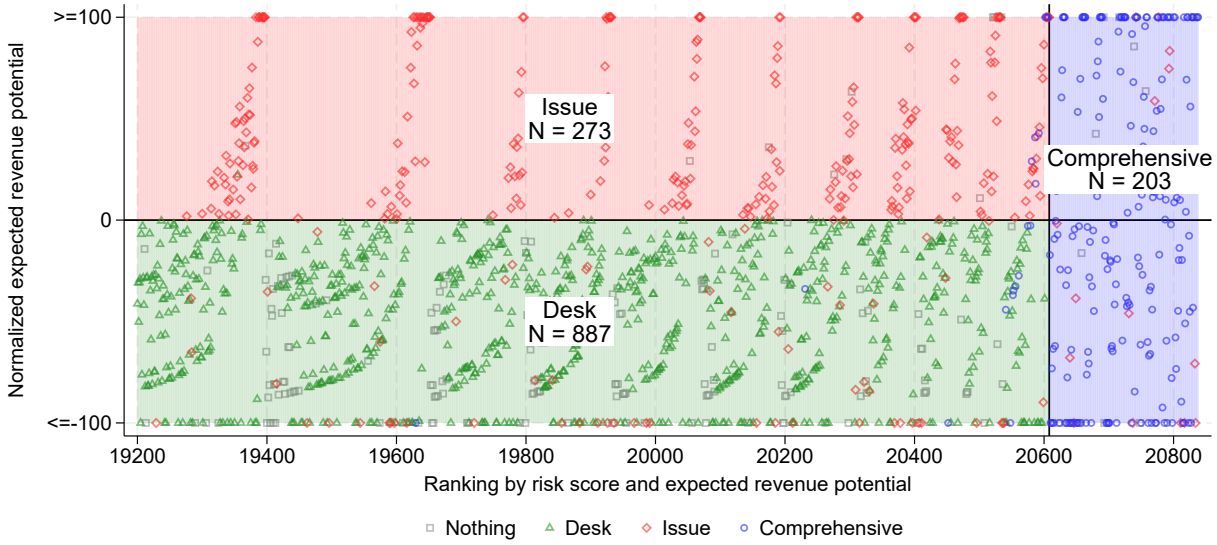


(b) What would encourage firm registration



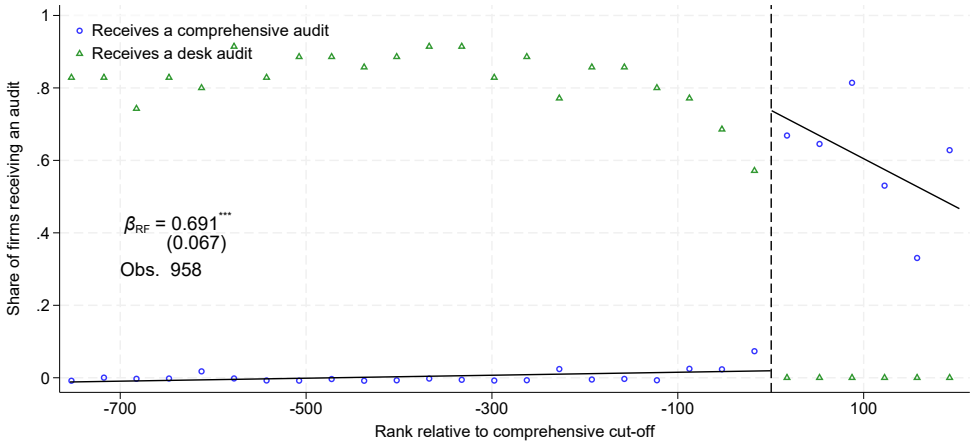
Notes: This figure displays the main reasons for why informal firms are not registered (Panel (a)) and what would encourage them to register (Panel (b)). The question concerned not being registered with Ugandan Registry Services Bureau, i.e. the firm was not registered at all. The question does not pertain to why the firm is not registered for taxes. In both Panel (a) and (b) reasons were not mutually exclusive, respondents could choose more than one option. In Panel (a) we combine "too expensive" and "bad for business" to capture the costs of becoming formal. N = 1951. Source: World Bank (2019b), authors' calculations.

Figure 4
Relation between ordinal rank, normalized expected revenue and audit assignment



Notes:

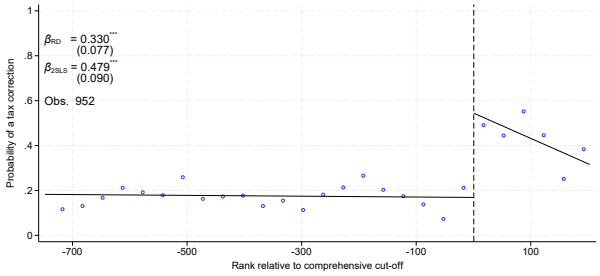
Figure 5
The effect of a firms' relative rank on the probability of receiving a comprehensive audit



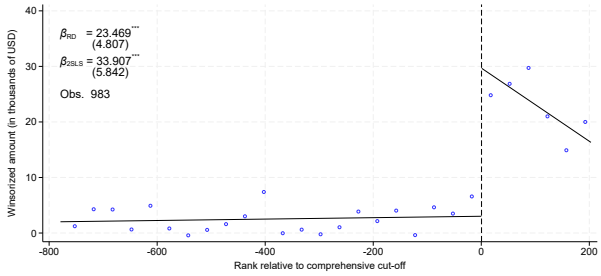
Notes: This figure presents the discontinuity leveraged for identification. The probability of comprehensive audit is residualized. The plots are manually constructed, each point on the graph includes approximately 35 firms. The blue circles show the share of firms receiving a comprehensive audit and the green triangles show the share of firms receiving a desk audit within each of the manually defined bins. The black lines show the linear best fit for the residualized data. The black dashed line presents the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). In the box we present the RD coefficient we retrieve from running equation 1, with heteroskedastic robust standard errors presented below the coefficient. "Obs." refers to the number of firms used in the regression. Source: Authors' calculations using URA tax data.

Figure 6
The impact of a comprehensive audit on increased tax liability through tax corrections

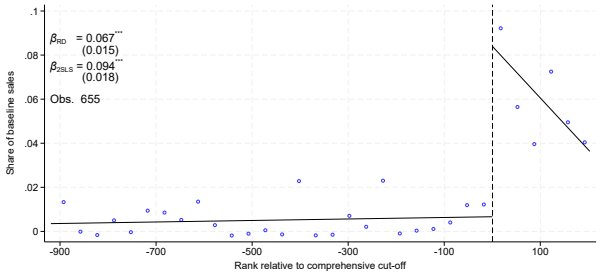
(a) Probability of an increase in tax liability



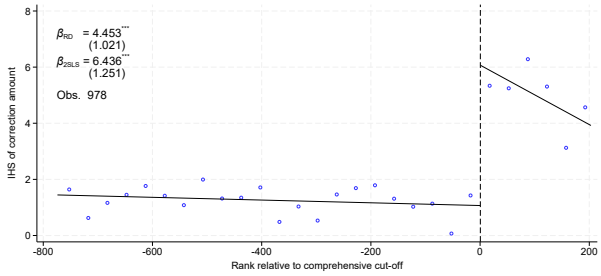
(b) Increase in tax liability (in thousands of USD)



(c) Increase in tax liability as a share of baseline sales



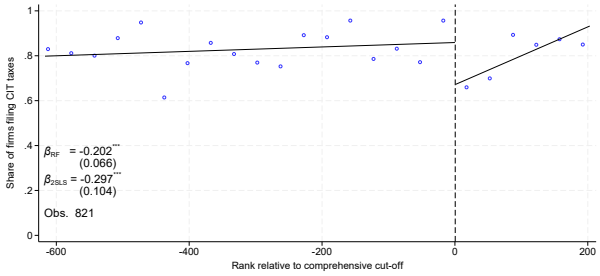
(d) IHS transformation of increase in tax liability



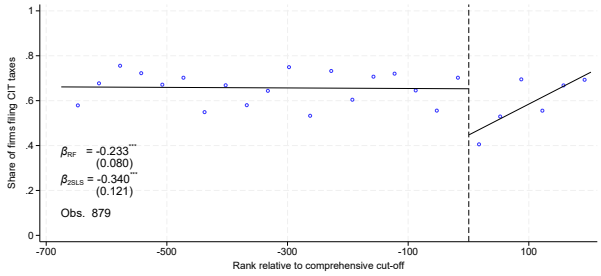
Notes: This figure presents results on the probability of filing after the audit. In Panel (a) we report the probability of filing either the last CIT or VAT return in the year after the audit. The last VAT return we observe is November 2022. In Panel (b) we report the probability of filing the CIT return in the second year post-audit, that is financial year 2022/23. The outcome variable shown in the figure is residualized. The plots are manually constructed, each point on the graph contains roughly 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). β_{RD} shows the reduced form effects – running equation 2 – and $\beta_{2\text{SLS}}$ shows the 2SLS effect – running equation 1 and 3. Heteroskedastic robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: Authors' calculations using URA tax data.

Figure 7
The impact of a comprehensive audit on subsequent tax filings

(a) Filed CIT one year post-audit



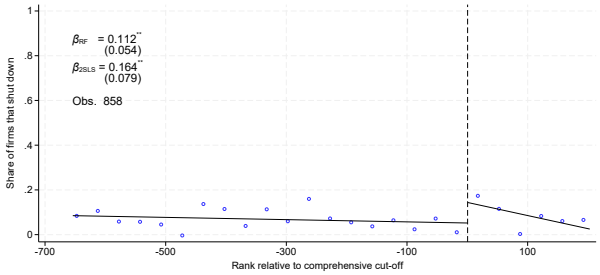
(b) Filed CIT two years post-audit



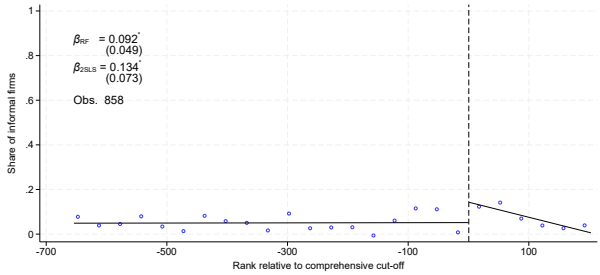
Notes: This figure presents results on the probability of filing after the audit. In Panel (a) we report the probability of filing either the last CIT or VAT return in the year after the audit. The last VAT return we observe is November 2022. In Panel (b) we report the probability of filing the CIT return in the second year post-audit, that is financial year 2022/23. The outcome variable shown in the figure is residualized. The plots are manually constructed, each point on the graph contains roughly 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). β_{RD} shows the reduced form effects – running equation 2 – and β_{2SLS} shows the 2SLS effect – running equation 1 and 3. Heteroskedastic robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: Authors' calculations using URA tax data.

Figure 8
The impact of a comprehensive audit on firm closure and informality

(a) Prob. of being closed



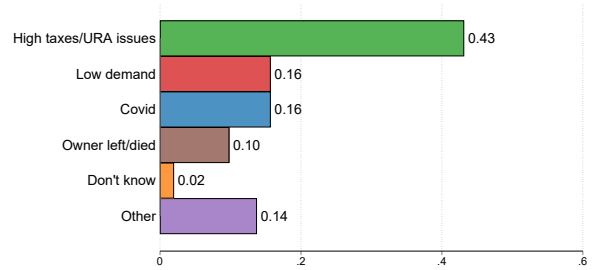
(b) Prob. of being informal



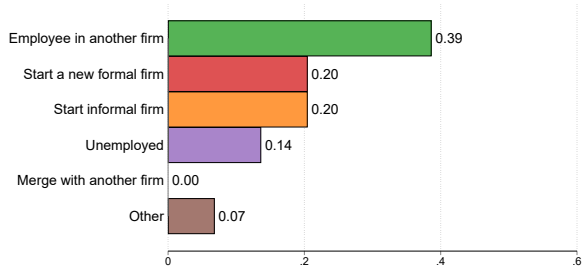
Notes: This figure present results on the probability of being closed or being informal rougly two years after the audit. ''"Prob of being closed" is an indicator for whether the firm closed or vanished *and* did filed neither the VAT nor CIT return in the immediate aftermath of the audit. "Prob. of being informal" is an indicator for whether we found the firm, but it did not filed either the VAT or CIT return in the immediate aftermath of the audit. The outcome variable shown in the figure is residualized. The plots are manually constructed, each point on the graph contains roughly 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). β_{RD} shows the reduced form effects – running equation 2 – and β_{2SLS} shows the 2SLS effect – running equation 1 and 3. Heteroskedastic robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: URA administrative tax data and authors' survey. Authors' calculations.

Figure 9
Reasons for closure and subsequent occupation

(a) Why did the firm close?



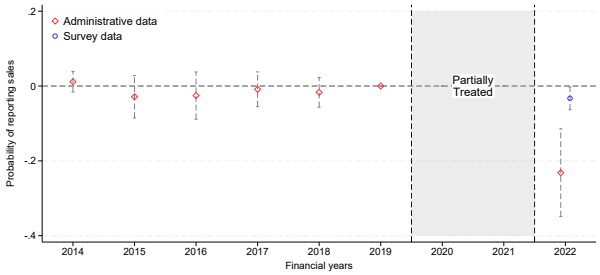
(b) What is the respondent doing now?



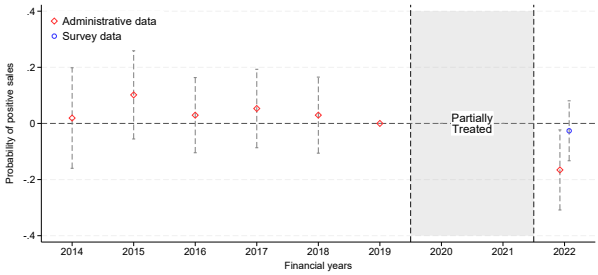
Notes: This figure presents reasons for closure and subsequent occupations of the respondent. In panel (a) we show the reasons for closure as reported by the respondent. This was an open ended question, we categorized question into the categories shown here. In panel (b) we show what the current occupation of the respondent. Occupations are mutually exclusive. For panel (a) N = 52, for panel (b) N = 44. Source: Authors' survey and calculations.

Figure 10
Event-study graphs

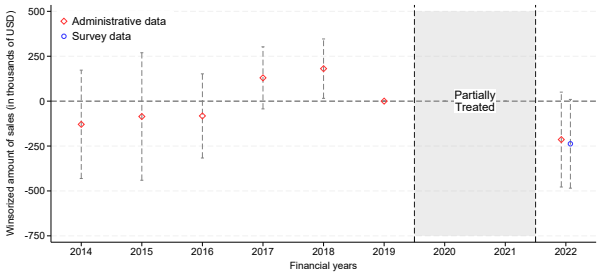
(a) Probability of reporting any sales



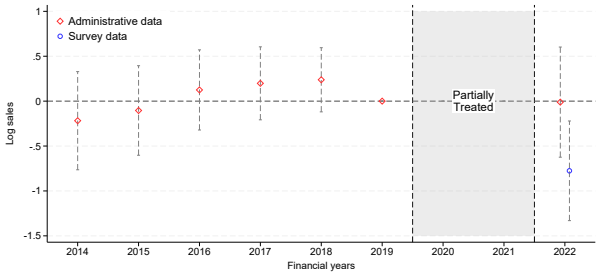
(b) Probability of reporting positive sales



(c) Winsorized amount of sales



(d) Logarithmic amount of sales



Notes: These figures presents the results from running equation 7 for the administrative data and a combination of the administrative and survey data for each of the outcomes shown in table 6. We restrict to firms that either reported sales in our survey or reported that their last active accounting year was before 2022, so we can infer that their sales in 2022 were 0. Panel (a) is an indicator for whether any sales information was reported (i.e. it was not missing), Panel (b) is an indicator for whether the firm reported positive sales, conditional on submitting any sales information. Panel (c) is the amount of sales reported winsorize at the top and bottom 1 percentile, and Panel (d) is the logarithmic transformation on the amount of sales. The shaded area indicates the two years we remove when running the DD-specification. Standard errors are robust to heteroskedasticity and clustered at the firm level.

Tables

Table 1
Summary statistics

	All	Comprehensive			
		Full		Optimal	
	Mean	Mean	P-value	Mean	P-value
	(1)	(2)	(3)	(4)	(5)
Registration					
Years since tin registration	5.95	6.46	0.000	7.67	0.000
VAT data					
Share of VAT returns filed	0.23	0.49	0.000	0.78	0.000
IHS Sales	10.44	10.95	0.000	12.04	0.000
IHS Purchases	9.44	9.98	0.000	11.18	0.000
Share with positive VAT	0.51	0.54	0.000	0.54	0.052
IHS VAT	4.49	4.74	0.011	4.72	0.374
Share with positive VAT due	0.44	0.45	0.020	0.44	0.773
IHS VAT Due	1.18	1.04	0.285	0.60	0.079
Number of suppliers	2.14	1.63	0.000	2.55	0.112
Number of buyers	1.17	1.12	0.008	1.58	0.000
Prob. caught evading	0.04	0.10	0.000	0.21	0.000
CIT data					
Prob. of filing CIT	0.56	0.73	0.000	0.77	0.000
IHS Turnover	6.32	9.00	0.000	9.85	0.000
IHS Cost of sales	4.88	7.14	0.000	8.49	0.000
Prob. of positive pre-tax profits	0.34	0.46	0.000	0.38	0.010
IHS Profit pre tax	0.51	0.62	0.172	-0.43	0.000
Prob. of positive CIT liability	0.32	0.43	0.000	0.36	0.014
IHS CIT liability	2.25	3.22	0.000	3.22	0.000
PAYE data					
Share of PAYE returns filed	0.28	0.50	0.000	0.69	0.000
Number of employees	245.41	146.07	0.000	304.15	0.336
Sectors					
Agriculture	0.05	0.04	0.000	0.03	0.017
Construction	0.14	0.13	0.005	0.13	0.254
Education	0.03	0.03	0.441	0.02	0.002
Financial	0.05	0.03	0.000	0.03	0.004
Health	0.03	0.03	0.798	0.06	0.000
Manufacturing	0.05	0.06	0.000	0.11	0.000
Missing sector	0.01	0.01	0.000	0.01	0.108
Other services	0.01	0.42	0.000	0.36	0.000
Retail trade	0.40	0.11	0.000	0.10	0.000
Wholesale trade	0.10	0.14	0.000	0.17	0.000
P-value of joint F-test			0.000		0.000
Number of firms	58,651	13,014		937	

Notes: All variables are based on information submitted in the returns for financial year 2019/20. Column (1) presents averages for the universe of firms that filed at least one tax return prior to the audit year. Column (2) and (4) present averages for the full comprehensive sample and the comprehensive sample that tend to fall within the optimal bandwidth. To derive at the latter, we select the firms within the median optimal bandwidth across all baseline covariates shown here. Column (3) shows the p-value of a t-test for whether the average of the comprehensive sample is significantly different from the average in the universe of firms. Column (4) repeats the same exercise, but comparing the average for firms in the optimal bandwidth to the average in the full comprehensive sample. Variables preceded by "IHS" are the inverse hyperbolic sine transformation of the variable in question. For the *VAT data* variables represent monthly averages over the 12 months from July 2019 to June 2020. There are two exceptions to this. "Prob. caught evading" is an indicator for whether any assessment was filed during the financial year that increased the tax liability of the firm. "Number of returns amended" is the total number of VAT amendments that were filed during the financial year, we count missings as 0's. Unless otherwise indicated, variables are conditional on filing at least one VAT return. For the *CIT data* variables show what is declared in the CIT return for financial year 2019/20. There is one exception, "Number of returns amended" is the total number of CIT amendments filed during the financial year. Unless otherwise indicated, variables are conditional on filing a CIT return. For the *PAYE data* "Number of employees" is the average number of employees reported conditional on filing at least one PAYE return. Sectors are aggregate versions of the sector classification used by the URA. Source: VAT, CIT, PAYE return from the URA. Authors' calculations.

Table 2
Summary statistics on tracking rates and survey completion rates

	N	Percent
All firms	839	
Closed shop	81	10%
Cannot be found (vanished)	75	9%
Exist	666	79%
Completed survey	298	45%
Refused participation	212	32%
Other	156	23%
NGO	17	2%

Notes: This table presents summary statistics for the tracking exercise and the share of survey done among the firms that were found to exist. We define a firm as "Closed shop" if we can confirm from neighbours of the former premises, former employees/owners that the firm closed. "Vanished" is defined as us not being able to reach the firm despite at least three phone call attempts for each number and a visit to each location we have on file. "Other" are firms that were found to exist and never officially gave us a refusal. However, they kept delaying and rescheduling the interview to the point where it was not completed. The sample is the set of firms that fall within the optimal bandwidth of the regression discontinuity for the probability of filing a tax return in the year after the audit.

Table 3
Effect of a heavy tax audit on CIT liability in the year after the audit

	Prob. pos.	Amount		
		Winsor.	Scaled	IHS
	(1)	(2)	(3)	(4)
RD coefficient	−0.189*	−3.714**	−0.015**	−1.864***
	(0.106)	(1.761)	(0.007)	(0.661)
2SLS coefficient	−0.265*	−5.463**	−0.022**	−2.729***
	(0.146)	(2.670)	(0.010)	(0.997)
N	664	816	497	861
Mean in control	0.51	3.98	0.02	3.01
Bandwidth	800	613	658	658

Notes: This table presents reduced form and two stage least squares estimates for the amount of CIT revenue collected in the first (Panel (a)) and second (Panel (b)) year post-audit. In Column (1) and (5) we present the probability of filing a tax return with a positive tax liability. In Column (2) and (6) we present the amount of tax liability submitted in thousands of USD winsorized at the top and bottom 1%. In Column (3) and (7) we present the amount of tax liability submitted scaled by baseline sales. In Column (4) and (8) we present the inverse hyperbolic sine transformation of the tax liability. The estimating equations are equation 2, and 3. All regressions include station fixed effects and indicators for whether the firm was in the VAT or CIT sample at baseline. Standard errors are robust to heteroskedasticity.

Table 4
Effect of a heavy tax audit on VAT liability

	Prob. pos. (1)	Amount (2)	Scaled amount (3)	IHS amount (4)
RD coefficient	−0.007	−4.272**	−0.199	−1.371
	(0.071)	(2.067)	(0.159)	(1.526)
2SLS coefficient	−0.010	−6.176**	−0.301	−1.991
	(0.097)	(2.942)	(0.236)	(2.171)
N	799	974	853	931
Mean in control	0.40	-1.43	-0.28	-0.07
Bandwidth	833	771	926	728

Notes: This table presents reduced form and two stage least squares estimates for the amount of VAT revenue collected in the first five months of the financial year the audit. In Column (1) we present the probability of filing a tax return with a positive tax liability. In Column (2) we present the amount of tax liability submitted in thousands of USD winsorized at the top and bottom 1%. In Column (3) we present the amount of tax liability submitted scaled by baseline sales. In Column (4) we present the inverse hyperbolic sine transformation of the tax liability. The estimating equations are equation 2, and 3. All regressions include station fixed effects and indicators for whether the firm was in the VAT or CIT sample at baseline. Standard errors are robust to heteroskedasticity.

Table 5
Characteristics of compliers

	All	Compliers			
		Audited		Not filing 21/22	
	Mean	Mean	P-value	Mean	P-value
	(1)	(2)	(3)	(4)	(5)
Years since tin registration	7.68	7.63	0.93	6.25	0.19
Share of VAT returns filed	0.75	0.87	0.01	0.76	0.94
IHS Sales	9.72	11.47	0.02	9.87	0.91
IHS Purchases	9.10	11.50	0.00	8.80	0.85
Prob of positive VAT Payable	0.54	0.53	0.84	0.50	0.78
IHS VAT Payable	3.69	3.50	0.89	3.90	0.94
Prob of positive VAT Due	0.44	0.44	1.00	0.58	0.35
IHS VAT Due	0.41	0.47	0.98	4.81	0.21
Number of suppliers	2.39	2.56	0.84	0.66	0.02
Number of buyers	1.59	2.52	0.00	0.64	0.01
Prob caught evading VAT	0.20	0.19	0.92	0.37	0.31
Prob of filing CIT	0.77	0.80	0.64	1.03	0.00
IHS Turnover	7.63	8.42	0.50	5.82	0.41
IHS Cost of sales	6.56	8.66	0.07	5.40	0.58
Prob of positive pre-tax profits	0.39	0.25	0.10	0.16	0.09
IHS Pre-tax profits	-0.33	-3.09	0.10	-4.51	0.11
Prob of positive tax liability	0.36	0.33	0.69	0.26	0.51
IHS Tax liability	2.52	2.51	0.99	2.22	0.81
IHS Value of plant/machines	2.72	1.66	0.30	0.55	0.06
Sales/Cost of sales	22.23	1.46	0.30	1.95	0.31
Prob of filing PAYE	0.69	0.79	0.11	0.34	0.04
Number of employees	24.77	34.40	0.48	1.17	0.00
Agriculture	0.03	0.01	0.47	0.02	0.80
Construction	0.14	0.14	0.98	0.04	0.23
Education	0.03	0.00	0.00	0.01	0.00
Financial	0.03	0.00	0.00	-0.01	0.00
Health	0.05	0.05	0.96	-0.01	0.00
Manufacturing	0.07	0.24	0.03	-0.07	0.00
Missing sector	0.01	0.03	0.45	0.11	0.40
Other services	0.35	0.29	0.50	0.55	0.27
Retail trade	0.10	0.12	0.63	0.13	0.75
Wholesale trade	0.17	0.15	0.76	0.25	0.58

Notes: All variables are based on information submitted in the returns for financial year 2019/20. Column (1) presents average baseline characteristics for the sample within the optimal bandwidth of the variable in question. In Column (2) we present the baseline characteristics for the firms that get induced into comprehensive audits. In Column (4) we present baseline characteristics for firms that get induced into not filing in the immediate aftermath of the audit. Column (5) and (6) present the p-value of a difference in means between the averages of the compliers and the sample. Variables preceded by "IHS" are the inverse hyperbolic sine transformation of the variable in question. We impute 0's for missing values throughout. For the *VAT data* variables represent monthly averages over the 12 months from July 2019 to June 2020. There are two exceptions to this. "Prob. caught evading" is an indicator for whether any assessment was filed during the financial year that increased the tax liability of the firm. "Number of returns amended" is the total number of VAT amendments that were filed during the financial year. For the *CIT data* variables show what is declared in the CIT return for financial year 2019/20. There is one exception, "Number of returns amended" is the total number of CIT amendments filed during the financial year. For the *PAYE data* "Number of employees" is the average number of employees reported. Sectors are aggregate versions of the sector classification used by the URA. Source: VAT, CIT, PAYE return from the URA. Authors' calculations.

Table 6
Comparing the effect of heavy tax audits on sales using tax data vs survey data

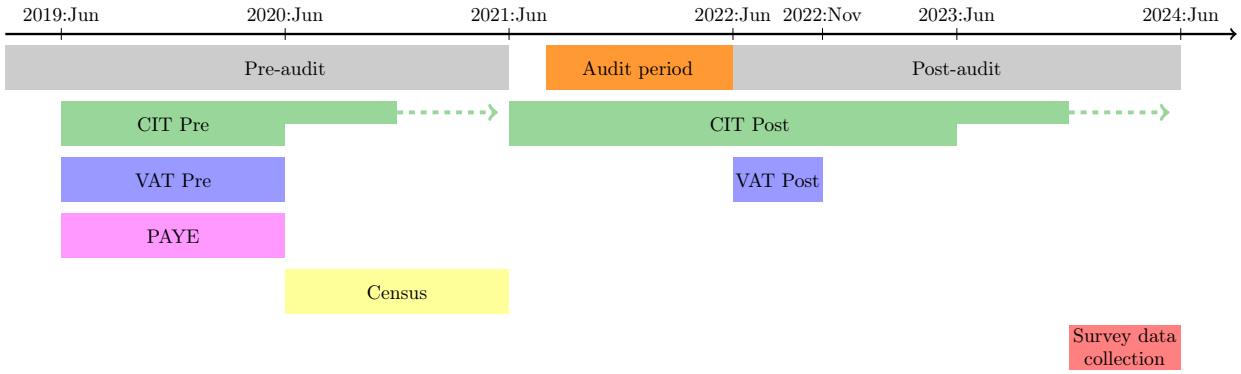
	Panel A: Admin					*Panel B: Admin & survey*				
	Probability		Sales amount			Probability		Sales amount		
	any info	Sales>0	Log	Winsor.	Winsor.	any info	Sales>0	Log	Winsor.	Winsor.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DD coefficient	−0.208***	−0.193**	0.017	−118.171	−258.861**	−0.022	−0.062*	−0.842***	−93.690	−265.958**
	(0.060)	(0.078)	(0.337)	(211.747)	(131.683)	(0.018)	(0.033)	(0.270)	(144.726)	(127.295)
N	4549	2856	2342	2856	4549	4549	3686	3158	3686	4549
No. of firms	782	503	478	503	782	782	654	630	654	782
Mean of Dep. in Control	0.952	0.836	19.831	386.696	347.433	0.952	0.838	19.800	357.250	353.101

Notes: This table presents results from running equation 6 for the administrative data (Panel A) and a combination of the administrative and survey data (Panel B). In Panel B we replace the administrative with the survey data in the post-audit period (2022). Throughout we include years 2013-2019 for the pre-period and 2022 for the post-period. We restrict to firms that either reported sales in our survey or reported that their last active accounting year was before 2022, so we can infer that their sales in 2022 were 0. Column (1) is an indicator for whether any sales information was reported (i.e. it was not missing), Column (2) is an indicator for whether the firm reported positive sales, conditional on submitting any sales information. Column (3) is the amount of sales reported winsorize at the top and bottom 1 percentile, and Column (4) is the logarithmic transformation on the amount of sales. Column (5) – (8) repeat the same outcomes for the combination of the administrative and survey data. Standard errors are robust to heteroskedasticity and clustered at the firm level. Event study-specifications are shown in appendix figure 10.

A Robustness

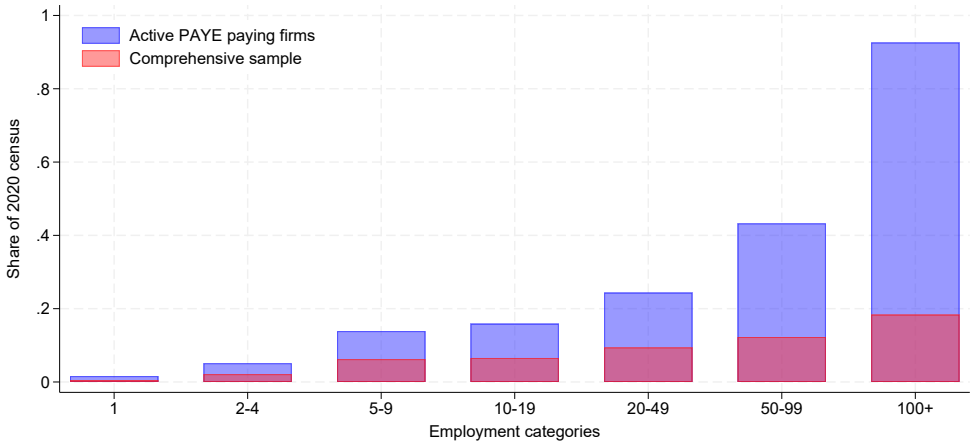
A.1 Figures

Figure A1
Timeline of data sources



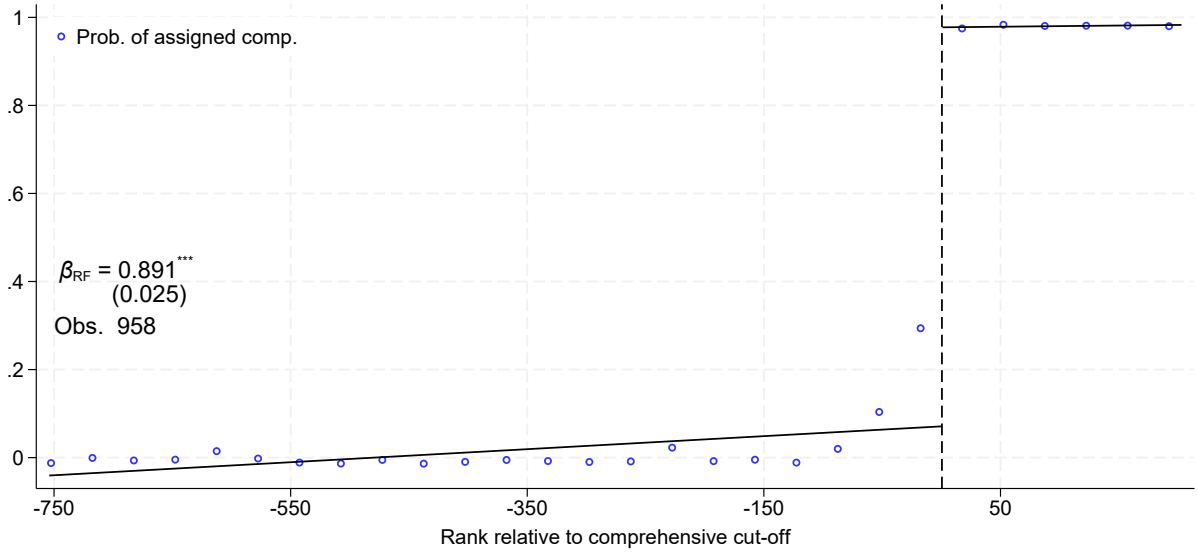
Notes: This figure displays the timeline of each of the data sources. The CIT data is filed annually, and is due 6 months after the close of a firms' accounting year. 79% end their accounting year in June, 6% in December and the remainder are scattered across the year. The dashed green line illustrates the time period where the last set of firms are filing the return for the previous financial year. Given the delay in filing by firms, the last unaffected CIT return before the audit intervention is the CIT return for financial year 2019. VAT returns are filed monthly, but we use VAT information that coincides with the coverage of the CIT return. The census of business establishments was conducted in financial year 2020/21. We started survey data collection in January 2024 and it is still ongoing.

Figure A2
Share of tax filing firms in each employment category



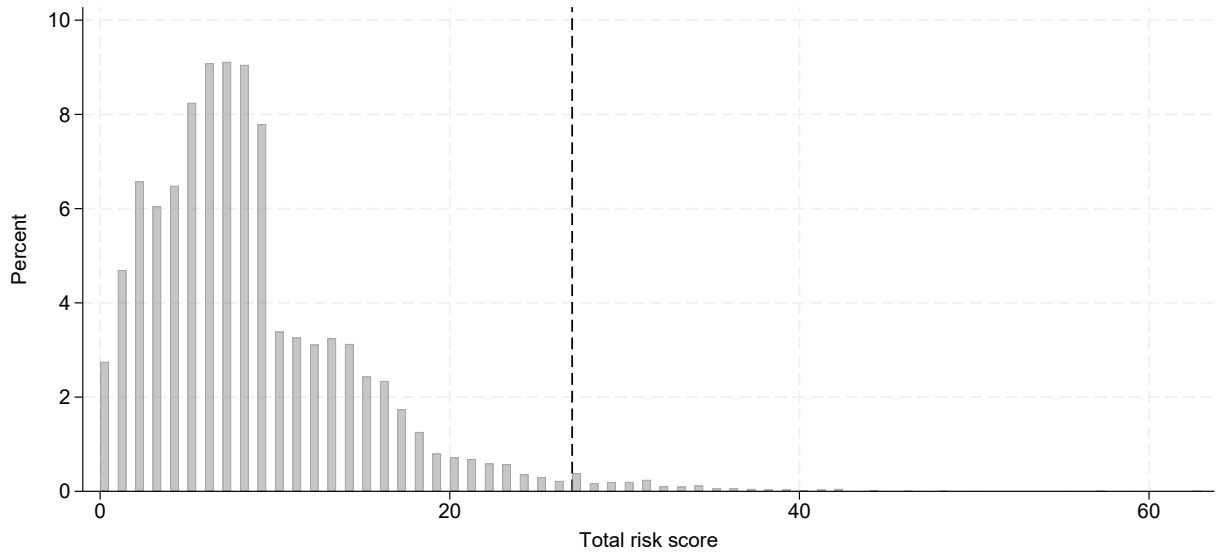
Notes: This figure presents the share of firms filing taxes and the share of firms in the comprehensive sample in each employment category. The employment categories are defined by the census, and we construct equivalent categories using information on number of employees submitted in the PAYE tax return. We restrict to firms that have filed at least one VAT or CIT before the audit year. We take a conservative appraoch and define the number of employees as the minimum number of employees reported in a PAYE return over the financial year 2020/21, which is the year the census was conducted. Thus, if anything, we are undercounting the number of tax-filing firms in the larger employment categories.

Figure A3
RD graph for whether comprehensive audit was assigned



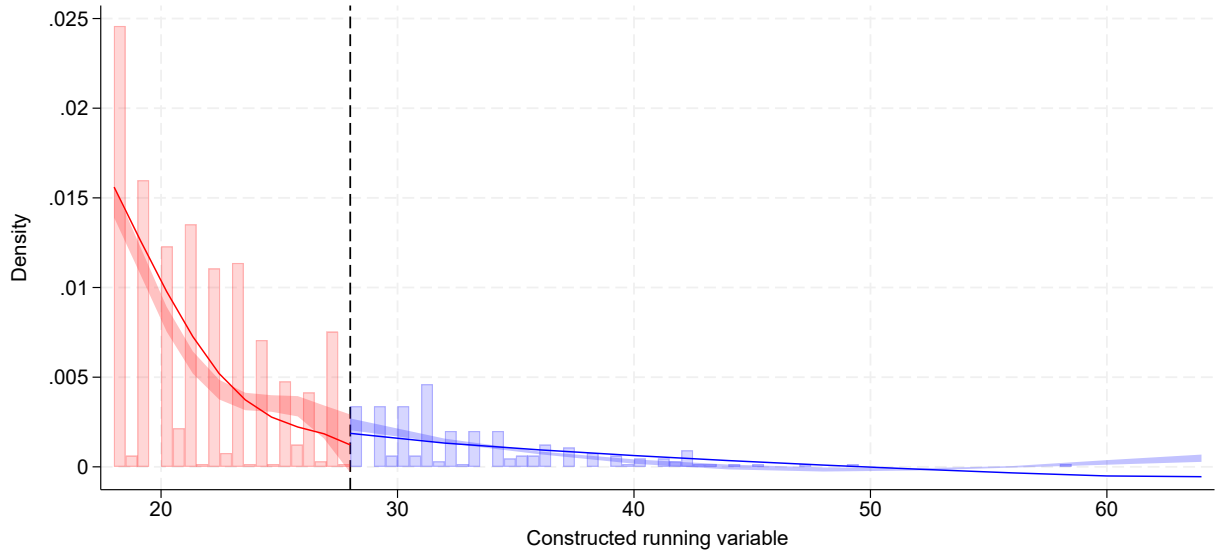
Notes: This figure present results on the probability that a comprehensive audit was assigned. The outcome variable shown in the figure is residualized. The plots are manually constructed, each point on the graph contains roughly 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). β_{RD} shows the reduced form effects – running equation 2 – and β_{2SLS} shows the 2SLS effect – running equation 1 and 3. Heteroskedastic robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: URA administrative tax data. Authors' calculations.

Figure A4
Risk score density distribution



Notes: This figure presents the percent of firms at each discrete risk score in the comprehensive sample. Source: URA administrative tax data. Authors' calculations.

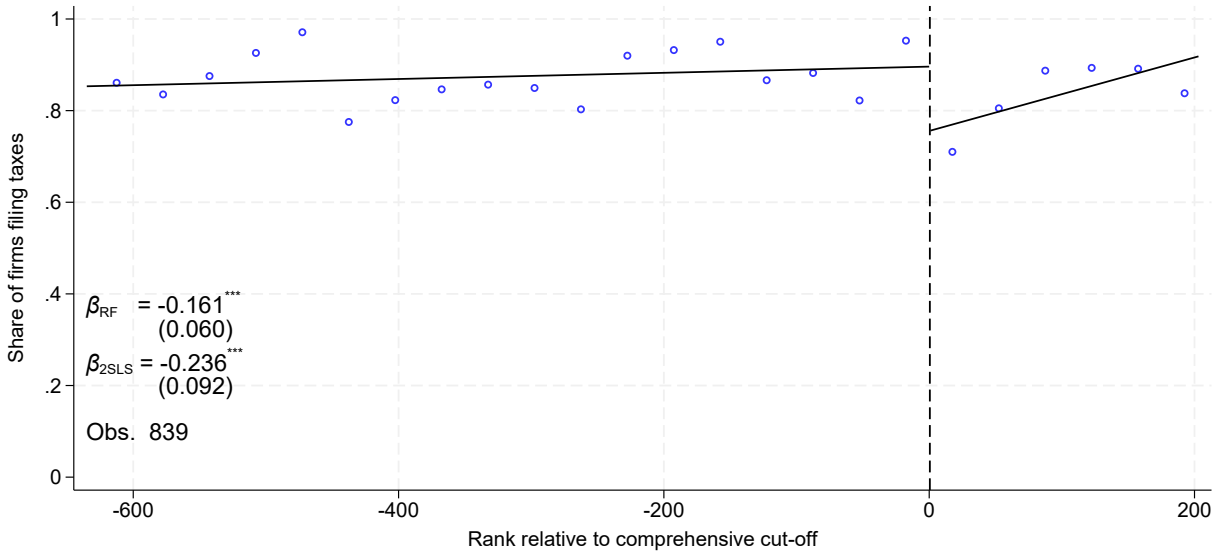
Figure A5
Density plot for alternative running variable



Notes: This figure density plot for the alternative running variable. The figure is produced using the "rddensity" package developed by **cattaneoSimpleLocalPolynomial2020 cattaneoManipulationTestingBased2018**. Due to left-skewed distribution, shown in figure A4 we restrict the rddensity command to run within the optimal bandwidth for our main outcome of interest, filing a return in the immediate aftermath of the audit. Source: URA administrative tax data. Authors' calculations.

Figure A6
Prob. of filing CIT or VAT tax return one year post-audit

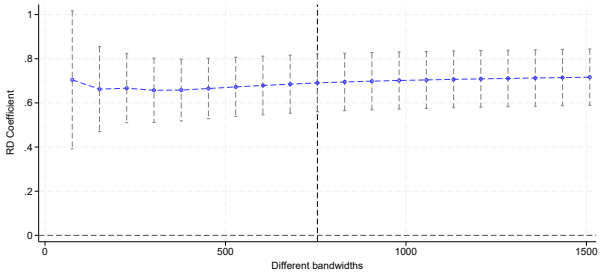
(a) Filed CIT or VAT one year post-audit



Notes: This figure present results on the probability of filing a CIT return – in panel (a) – and the last VAT return – in panel (b) – separately for the year after the audit. The last VAT return we observe is November 2022, whereas the last CIT return is simply the CIT return for financial year 2021/22. The outcome variable shown in the figure is residualized. The plots are manually constructed, each point on the graph contains roughly 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). β_{RD} shows the reduced form effects – running equation 2 – and β_{2SLS} shows the 2SLS effect – running equation 1 and 3. Heteroskedastic robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: URA administrative tax data. Authors' calculations.

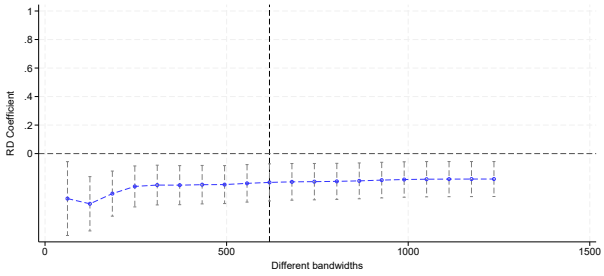
Figure A7
Key outcomes with different bandwidths

(a) Likelihood of receiving a comprehensive audit

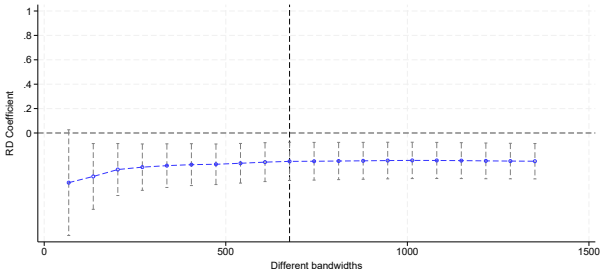


(b) Increase in tax liability from tax corrections

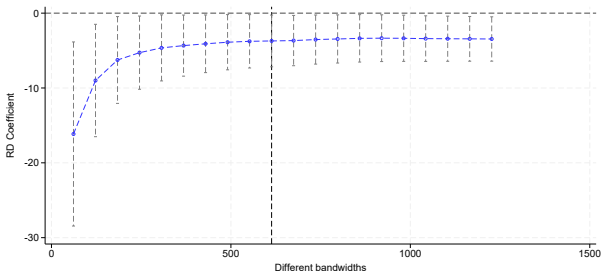
(c) Filing CIT tax one year post-audit



(d) Filing CIT tax two years post-audit



(e) CIT liability one year post-audit

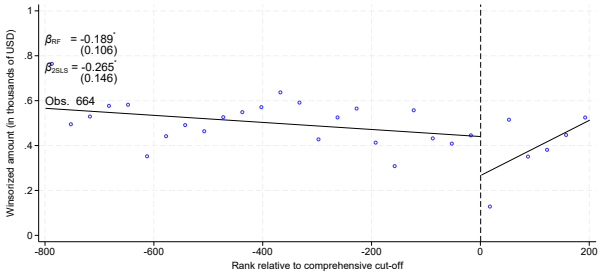


(f) Average VAT liability during 5 months post-audit

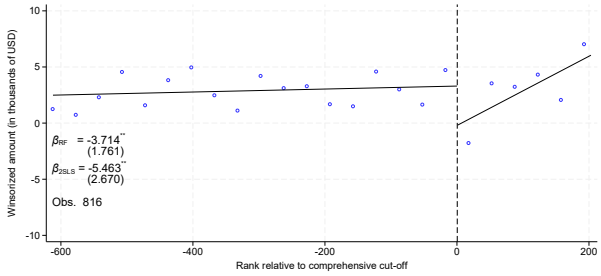
Notes: This figure presents the coefficient on the probability of filing either a CIT or the last VAT return in the year after the audit (Panel (a)) and on the probability of filing the latest CIT return (Panel (b)) across different bandwidths. The vertical dashed line represent the optimal bandwidth selected by Calonico et al. (2014). We restrict this analysis to the reduced form effect, i.e. we run equation 1 with the outcome of interest replacing D_j^c. 95% confidence intervals are presented in the figure.

Figure A8
RD graph for post-audit CIT liability in the year after the audit

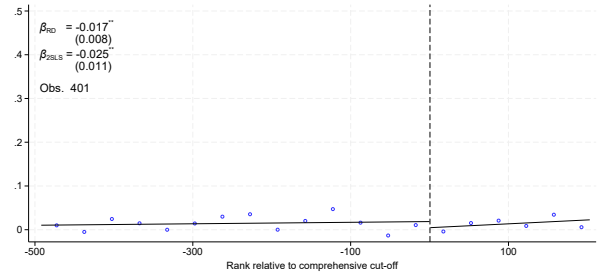
(a) Prob of reporting positive CIT liability



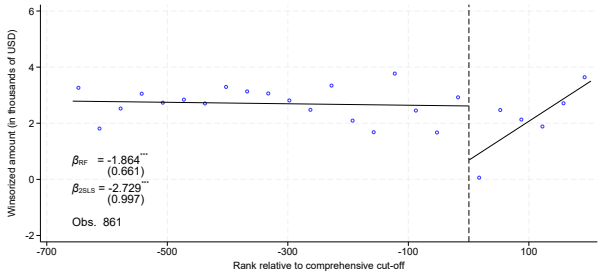
(b) Amount of CIT liability



(c) Scaled CIT liability



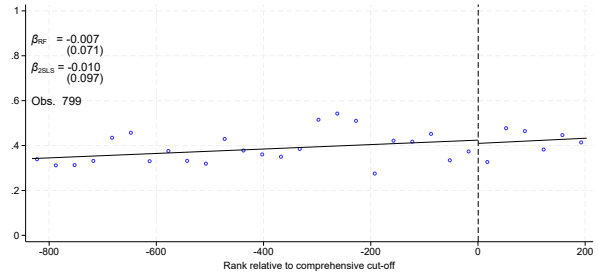
(d) IHS amount of CIT liability



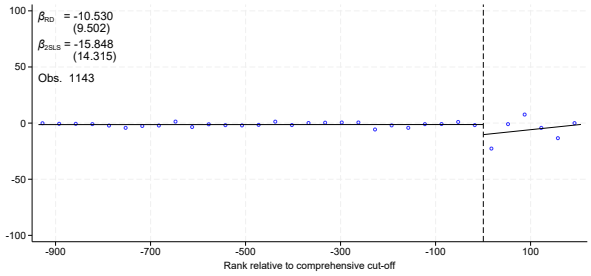
Notes: This figure presents the coefficient on the amount of CIT revenue declared in the aftermath of the audit. In Panel (a) we present the probability of filing a tax return with a positive tax liability, in panel (b) we present the amount of tax liability submitted in thousands of USD winsorized at the top and bottom 1%, in panel (c) we present the amount of tax liability submitted scaled by baseline sales, and in panel (d) we present the inverse hyperbolic sine transformation of the tax liability. The left-hand side of the figure (subfigure (a), (c), (e), (g)) pertain to tax returns for the year of the audit. The remaining figures repeat the exercise for the year after the audit. The outcome variable shown in the figure is residualized. The plots are manually constructed, each point on the graph contains roughly 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). β_{RD} shows the reduced form effects – running equation 2 – and β_{2SLS} shows the 2SLS effect – running equation 1 and 3. Heteroskedastic robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: URA administrative tax data. Authors' calculations.

Figure A9
RD graph for post-audit VAT liability

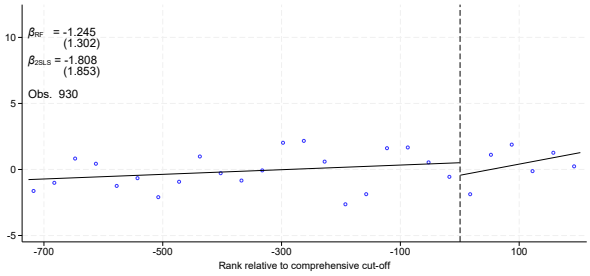
(a) Prob. of reporting positive VAT liability



(b) Amount of VAT liability



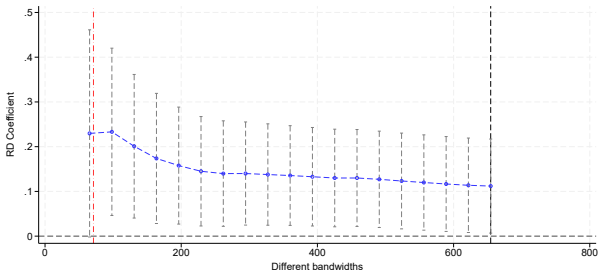
(c) IHS amount of VAT liability



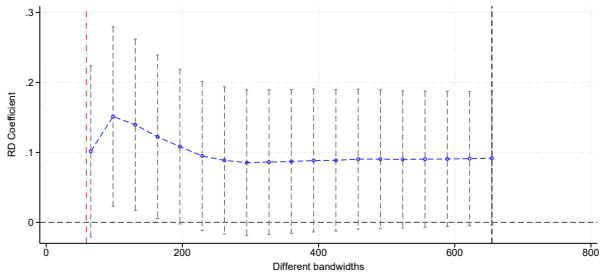
Notes: This figure presents the coefficient on the amount of VAT revenue declared in the first five months of the financial year after the audit. In figure (a) we present the probability of filing a tax return with a positive tax liability, in figure (b) we present the amount of tax liability submitted in thousands of USD winsorized at the top and bottom 1%, in figure (c) we present the amount of tax liability submitted scaled by baseline sales, and in figure (d) we present the inverse hyperbolic sine transformation of the tax liability. The outcome variable shown in the figure is residualized. The plots are manually constructed, each point on the graph contains roughly 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). β_{RD} shows the reduced form effects – running equation 2 – and $\beta_{2\text{SLS}}$ shows the 2SLS effect – running equation 1 and 3. Heteroskedastic robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: URA administrative tax data. Authors' calculations.

Figure A10
Prob. of being closed or informal post-audit

(a) Prob. of being closed

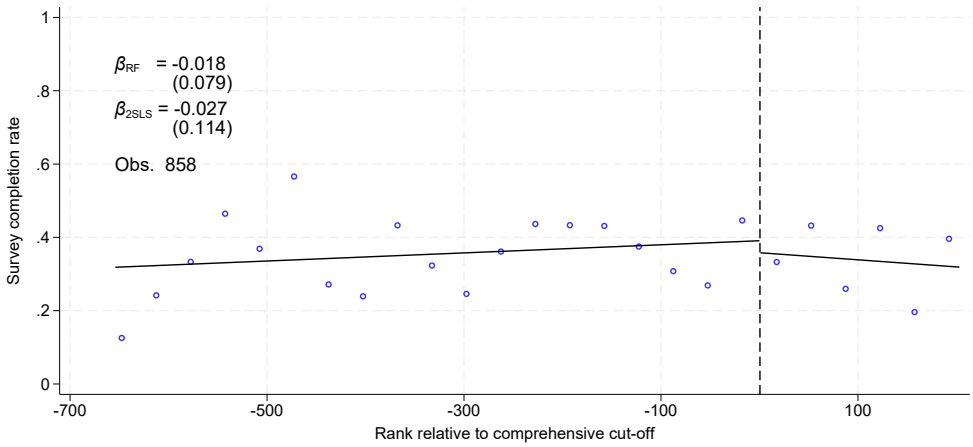


(b) Prob. of being informal



Notes: This figure presents the coefficient on the probability being closed (Panel (a)) and informal (Panel (b)) across different bandwidths. The vertical dashed line represent the optimal bandwidth selected by Calonico et al. (2014). We restrict this analysis to the reduced form effect, i.e. we run equation 1 with the outcome of interest replacing D_j^c. 95% confidence intervals are presented in the figure.

Figure A11
Survey completion rate



Notes: This figure present results on the probability of completing a survey. The outcome variable shown in the figure is residualized. The plots are manually constructed, each point on the graph contains roughly 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated using the "rdrobust" package by Calonico et al. (2014). β_{RD} shows the reduced form effects – running equation 2 – and β_{2SLS} shows the 2SLS effect – running equation 1 and 3. Heteroskedastic robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: URA administrative tax data and authors' survey. Authors' calculations.

A.2 Tables

Table A1
Sample restrictions

	N	Comprehensive	Issue	Desk	Nothing
All firms with at least one risk	29,082	276	1,411	16,058	11,337
Removing specialized stations and mining sector	28,205	243	1,301	15,459	11,202
Removing taxpayers audited previous year	27,579	233	1,231	15,053	11,062
Ever filed a CIT or VAT return before audit period	23,592	231	1,217	14,528	7,616
Removing stations with no comprehensive audits	20,838	231	1,041	12,783	6,783
Focus on firms assigned to Comprehensive and desk audits	13,014	231		12,783	

Notes: This table presents the number of firms that remain after each restriction is imposed. In Column 1 we present the total number of firms in the sample, whereas in Column 2 - 4 we show how many firms were assigned to each of the enforcement interventions. Specialized stations include the Large Taxpayer Office, Oil and Gas office and the Public Sector Office. The mining sector is excluded because 40% of firms in the mining sector are registered with the Oil and Gas office.

Table A2
Balance table

	Comprehensive			
	(1) Mean	(2) Coef	(3) p-value	(4) N
Years since tin registration	7.52	-0.73	0.15	813
Share of VAT returns filed	0.76	0.04	0.24	871
IHS Sales	11.91	-0.09	0.84	814
IHS Purchases	10.98	0.44	0.39	776
Prob of positive VAT Payable	0.54	0.00	0.96	818
IHS VAT Payable	4.63	-0.79	0.59	806
Prob of positive VAT Due	0.44	0.02	0.74	826
IHS VAT Due	0.57	1.11	0.54	806
Number of suppliers	2.49	-1.11	0.20	713
Number of buyers	1.48	0.11	0.69	833
Prob caught evading VAT	0.19	0.01	0.88	959
Prob of filing CIT	0.75	0.04	0.49	886
IHS Turnover	9.84	-1.00	0.32	731
IHS Cost of sales	8.29	0.13	0.90	715
Prob of positive pre-tax profits	0.39	-0.13	0.13	716
IHS Pre-tax profits	-0.34	-2.24	0.22	720
Prob of positive tax liability	0.37	-0.04	0.61	660
IHS Tax liability	3.26	-0.56	0.51	624
Prob of filing PAYE	0.66	-0.07	0.32	817
Number of employees	270.36	-47.05	0.79	677
Agriculture	0.03	-0.01	0.73	954
Construction	0.13	-0.01	0.82	938
Education	0.03	-0.01	0.02	4650
Financial	0.02	0.00	0.95	833
Health	0.07	0.03	0.42	775
Manufacturing	0.09	0.05	0.42	963
Missing sector	0.01	0.03	0.36	831
Other services	0.36	-0.04	0.61	913
Retail trade	0.10	-0.03	0.49	900
Wholesale trade	0.16	0.01	0.86	913
P-value of joint f-test (min. bandwidth)			0.13	310
P-value of joint f-test (med. bandwidth)			0.48	360
P-value of joint f-test (max. bandwidth)			0.24	1182
P-value of joint f-test (min. bandwidth) imputing 0			0.58	775
P-value of joint f-test (med. bandwidth) imputing 0			0.63	936
P-value of joint f-test (max. bandwidth) imputing 0			0.12	4650

Notes: In this table we present results from conducting balance tests on a wide range of covariates measured at baseline. We run equation 1 with the outcome replacing D_j^c and the optimal bandwidth selected using the "rdrobust" package by Calonico et al. (2014). In the last 6 rows we conduct a test for whether the baseline covariates are jointly different from 0. We run the specification with 3 different bandwidths, the minimum, median and maximum bandwidth selected across all the baseline covariates and separately for when we impute 0's for missing values and when we do not. All variables are based on information submitted in the returns for financial year 2019/20. Variables preceded by "IHS" are the inverse hyperbolic sine transformation of the variable in question. For the *VAT data* variables represent monthly averages over the 12 months from July 2019 to June 2020. There are two exceptions to this. "Prob. caught evading" is an indicator for whether any assessment was filed during the financial year that increased the tax liability of the firm. "Number of returns amended" is the total number of VAT amendments that were filed during the financial year, we count missings as 0's. Unless otherwise indicated, variables are conditional on filing at least one VAT return. For the *CIT data* variables show what is declared in the CIT return for financial year 2019/20. There is one exception, "Number of returns amended" is the total number of CIT amendments filed during the financial year. Unless otherwise indicated, variables are conditional on filing a CIT return. For the *PAYE data* "Number of employees" is the average number of employees reported conditional on filing at least one PAYE return. Sectors are aggregate versions of the sector classification used by the URA. Source: VAT, CIT, PAYE return from the URA. Authors' calculations.

Table A3
Dimension reduction exercise: results when using predicted outcomes

	One year post-audit	Two years post-audit				
	Prob of filing any return	Prob of filing CIT return	Prob. of filing Positive amount	CIT aomunt Winsorized		
RD coefficient	−0.002	−0.028	−0.011	−0.433	−0.402	−0.068
	(0.005)	(0.030)	(0.028)	(0.492)	(0.801)	(0.355)
N	958	821	879	983	816	974
Mean in control	0.01	0.86	0.67	3.58	3.98	-1.43
Bandwidth	755	618	676	780	613	771

Notes:

Table A4
Effect of a heavy tax audit on CIT liability, year of audit

	Prob. pos.	Amount		
		Winsor.	Scaled	IHS
	(1)	(2)	(3)	(4)
RD coefficient	0.100	−0.375	−0.007	−0.069
	(0.096)	(2.751)	(0.009)	(0.827)
2SLS coefficient	0.133	−0.553	−0.010	−0.102
	(0.123)	(3.973)	(0.013)	(1.189)
N	798	794	543	832
Mean in control	0.50	4.97	0.02	3.78
Bandwidth	737	591	735	629

Notes: This table presents reduced form and two stage least squares estimates for the amount of CIT revenue collected in the first (Panel (a)) and second (Panel (b)) year post-audit. In Column (1) and (5) we present the probability of filing a tax return with a positive tax liability. In Column (2) and (6) we present the amount of tax liability submitted in thousands of USD winsorized at the top and bottom 1%. In Column (3) and (7) we present the amount of tax liability submitted scaled by baseline sales. In Column (4) and (8) we present the inverse hyperbolic sine transformation of the tax liability. The estimating equations are equation 2, and 3. All regressions include station fixed effects and indicators for whether the firm was in the VAT or CIT sample at baseline. Standard errors are robust to heteroskedasticity.

Table A5
Effect on different lines of the CIT return, year of audit

	Turnover	Cost of sales	Other income	Expenses			Profits
				Operational	Admin	Financial	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Amounts							
2SLS coefficient	286.636	280.303	−44.090*	−19.605	−58.107	−0.816	32.887
	(415.418)	(355.174)	(23.871)	(24.136)	(41.253)	(12.203)	(32.370)
N	798	851	714	753	741	770	856
Mean in control	696.45	536.92	19.97	42.66	60.59	14.43	6.39
Bandwidth	595	648	511	550	538	567	653
Panel B: Amounts scaled by baseline sales							
2SLS coefficient	−0.055	0.152	−0.031	−0.197**	−0.280**	−0.031	−0.000
	(1.217)	(0.779)	(0.108)	(0.094)	(0.120)	(0.030)	(0.104)
N	519	490	534	516	518	558	497
Mean in control	1.88	1.15	0.07	0.22	0.23	0.04	0.00
Bandwidth	697	649	719	689	692	757	658
Panel C: Log transformation of amounts							
2SLS coefficient	0.360	0.475	1.405	−0.086	−0.394	0.261	0.025
	(0.357)	(0.398)	(1.246)	(0.421)	(0.537)	(0.544)	(0.539)
N	631	604	305	625	598	576	487
Mean in control	12.59	12.09	9.67	9.99	10.12	8.18	9.60
Bandwidth	706	763	986	680	679	705	825

Notes: This table presents two stage least squares results on more granular lines shown in the CIT for information submitted that pertains to the year of the audit. That is, we run equation 3. Panel (A) shows the results for winsorized amounts, Panel (B) for values scaled by baseline saes, and Panel (C) presents the inverse hyperbolic sine transformation of the amount. In Column (1) we show total revenue, in Column (2) total cost of sales, and in Column (3) we show other forms of income. Column (4) – (6) show different forms of expenses and Column (7) shows total profits before taxes are applied. In theory Column (7) is supposed to be (1) + (3) - (2) - (4) - (5) - (6). All regressions include station fixed effects and indicators for whether the firm was in the VAT or CIT sample at baseline. Standard errors are robust to heteroskedasticity.

Table A6
Effect on different lines of the CIT return, one year after the audit

	Turnover	Cost of sales	Other income	Expenses			Profits
				Operational	Admin	Financial	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Amounts							
2SLS coefficient	1516.679	1511.825	−120.558*	−48.553	−22.707	51.275	−24.605
	(1057.910)	(996.856)	(65.007)	(34.668)	(56.071)	(36.671)	(73.015)
N	975	1023	659	747	860	921	848
Mean in control	657.69	529.07	31.44	40.08	53.43	12.25	6.22
Bandwidth	772	820	456	544	657	718	645
Panel B: Amounts scaled by baseline sales							
2SLS coefficient	0.863	0.936	−0.012	−0.066	−0.203*	−0.027	−0.005
	(1.130)	(0.891)	(0.045)	(0.135)	(0.111)	(0.025)	(0.122)
N	543	535	481	545	490	571	509
Mean in control	1.75	1.17	0.02	0.19	0.17	0.03	-0.00
Bandwidth	736	722	632	738	649	773	678
Panel C: Log transformation of amounts							
2SLS coefficient	0.573	0.316	1.708	−0.277	−1.328*	−0.321	0.389
	(0.463)	(0.508)	(1.045)	(0.453)	(0.792)	(0.800)	(0.730)
N	470	454	187	439	477	467	313
Mean in control	12.61	12.23	9.77	9.89	10.13	8.17	9.76
Bandwidth	714	764	1008	642	722	755	730

Notes: This table presents two stage least squares results on more granular lines shown in the CIT for information submitted that pertains to the year after the audit. That is, we run equation 3. Panel (A) shows the results for winsorized amounts, Panel (B) for values scaled by baseline saes, and Panel (C) presents the inverse hyperbolic sine transformation of the amount. In Column (1) we show total revenue, in Column (2) total cost of sales, and in Column (3) we show other forms of income. Column (4) – (6) show different forms of expenses and Column (7) shows total profits before taxes are applied. In theory Column (7) is supposed to be (1) + (3) - (2) - (4) - (5) - (6). All regressions include station fixed effects and indicators for whether the firm was in the VAT or CIT sample at baseline. Standard errors are robust to heteroskedasticity.

Table A7
Intensive vs extensive margin, reduction in CIT liability

	Amount Winsorized (1)	Prob. of filing (2)	Prob. of positive amount cond. on filing (3)	Log of CIT amount (4)	
RD coefficient	-3.663^{**}	-0.233^{***}	-0.182^{*}	0.058	-0.209
	(1.711)	(0.080)	(0.108)	(0.086)	(0.514)
N	879	879	578	578	280
Mean in control	4.10	0.67	0.49	0.16	8.42
Bandwidth	676	676	676	676	676

Notes: This table presents reduced form results for different margins of the behavioral responses among firms. That is, we run equation 1 with the outcome of interest replacing D_j^c. In Column (1) we show the inverse hyperbolic sine transformation of CIT revenue imputing 0's for missing values, in Column (2) we show the probability of a filing a return, the first extensive margin. In Column (3) we show the probability of a filing positive return conditional on filing a return, the second extensive margin. In Column (4) we show the logarithmic transformation of the liability submitted conditional on submitting a tax return with a positive CIT liability. To make the results comparable we fix the bandwidth to be the optimal bandwidth for the outcome in Column (1). All regressions include station fixed effects and indicators for whether the firm was in the VAT or CIT sample at baseline. Standard errors are robust to heteroskedasticity.

Table A8
First stage and reduced form effect on filing returns with different set of controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Prob of receiving a comprehensive audit						
RD coefficient	0.719***	0.686***	0.691***	0.690***	0.684***	0.680***	0.681***
	(0.067)	(0.067)	(0.067)	(0.067)	(0.066)	(0.068)	(0.066)
N	958	958	958	958	958	958	958
Bandwidth	755	755	755	755	755	755	755
	Panel B: Prob of filing a CIT return one year post-audit						
RD coefficient	−0.192***	−0.210***	−0.202***	−0.197***	−0.185***	−0.173***	−0.188***
	(0.067)	(0.065)	(0.066)	(0.065)	(0.063)	(0.065)	(0.061)
N	821	821	821	821	821	821	821
Bandwidth	618	618	618	618	618	618	618
	Panel C: Prob of filing a return two years post-audit						
RD coefficient	−0.207**	−0.225***	−0.233***	−0.226***	−0.226***	−0.280***	−0.234***
	(0.081)	(0.080)	(0.080)	(0.078)	(0.079)	(0.078)	(0.075)
N	879	879	879	879	879	879	879
Bandwidth	676	676	676	676	676	676	676
Station FE	No	Yes	Yes	Yes	Yes	Yes	No
Sample FE	No	No	Yes	Yes	Yes	Yes	No
Fictitious FE	No	No	No	Yes	No	No	No
Sector FE	No	No	No	No	Yes	Yes	No
Risk controls	No	No	No	No	No	Yes	No
Lasso selected controls	No	No	No	No	No	No	Yes

Notes: This table presents reduced form results on the probability of receiving a comprehensive audit – in Panel A – and the probability of filing either a VAT or CIT return in the first year after the audit – Panel B. That is, we run equation 1 with the outcome of interest replacing D_j^c, and varying $\mathbf{X}_f$. In Column (1) we do not use any controls, in Column (2) we only include station fixed effects, in Column (3) we add controls for whether the firm was in the VAT or CIT sample at baseline (our main specification), in Column (4) we include a control for whether the firm is "fictitious" defined as being labeled a suspicious firm by the URA, in Column (5) we add sector fixed effects, in Column (6) we add risk controls. The risk controls are 70 separate controls which control for the risk score and expected revenue of each of the underlying 35 risk parameters. Finally, in Column (7) we run a lasso specification to select the controls. In the specification we do not penalize station fixed effects because – based on conversations with the URA – they are an important part of selection. We fix the bandwidth to be the optimal bandwidth selected when running the specification in Column (3). Heteroskedasticity robust standard errors are reported in parentheses.

Table A9
Estimates on changes in tax liabilities with different set of controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Increase in tax liability from tax corrections						
RD coefficient	27.178***	23.847***	23.469***	23.241***	23.448***	21.524***	22.448***
	(4.992)	(4.812)	(4.807)	(4.786)	(4.796)	(4.703)	(4.620)
N	983	983	983	983	983	983	983
Bandwidth	780	780	780	780	780	780	780
	Panel B: CIT liability one year post-audit						
RD coefficient	−2.457	−3.278*	−3.714**	−3.658**	−3.834**	−4.454**	−3.705**
	(1.698)	(1.727)	(1.761)	(1.760)	(1.727)	(1.810)	(1.572)
N	816	816	816	816	816	816	816
Bandwidth	613	613	613	613	613	613	613
	Panel C: Average VAT liability during 5 months post-audit						
RD coefficient	−4.536**	−4.426**	−4.272**	−4.303**	−4.189**	−2.613	−4.396**
	(2.038)	(2.066)	(2.067)	(2.073)	(2.010)	(1.861)	(1.866)
N	974	974	974	974	974	974	974
Bandwidth	771	771	771	771	771	771	771
Station FE	No	Yes	Yes	Yes	Yes	Yes	No
Sample FE	No	No	Yes	Yes	Yes	Yes	No
Fictitious FE	No	No	No	Yes	No	No	No
Sector FE	No	No	No	No	Yes	Yes	No
Risk controls	No	No	No	No	No	Yes	No
Lasso selected controls	No	No	No	No	No	No	Yes

Notes: This table presents reduced form results on the probability of receiving a comprehensive audit – in Panel A – and the probability of filing either a VAT or CIT return in the first year after the audit – Panel B. That is, we run equation 1 with the outcome of interest replacing D_j^c, and varying $\mathbf{X}_f$. In Column (1) we do not use any controls, in Column (2) we only include station fixed effects, in Column (3) we add controls for whether the firm was in the VAT or CIT sample at baseline (our main specification), in Column (4) we include a control for whether the firm is "fictitious" defined as being labeled a suspicious firm by the URA, in Column (5) we add sector fixed effects, in Column (6) we add risk controls. The risk controls are 70 separate controls which control for the risk score and expected revenue of each of the underlying 35 risk parameters. Finally, in Column (7) we run a lasso specification to select the controls. In the specification we do not penalize station fixed effects because – based on conversations with the URA – they are an important part of selection. We fix the bandwidth to be the optimal bandwidth selected when running the specification in Column (3). Heteroskedasticity robust standard errors are reported in parentheses.

Table A10
Alternative threshold: Results when using the threshold from the initial assignment of audits

	One year post-audit	Two years post-audit					
	Prob of filing any return	Prob of filing CIT return	Prob. of filing Positive amount	CIT aomunt Winsorized			
RD coefficient	0.611^{***}	-0.168^{***}	-0.190^{**}	19.596^{***}	-2.744	-0.010	-3.967^{**}
	(0.068)	(0.062)	(0.078)	(4.609)	(1.781)	(0.007)	(1.988)
N	849	850	924	907	895	555	1020
Mean in control	0.01	0.85	0.67	3.54	4.10	0.02	-1.46
Bandwidth	641	642	716	699	687	749	812

Notes:

Table A11
Filing rates across tracking categories

	File vat or cit	File cit 22/23
Closed	0.71	0.39
Vanished	0.64	0.39
Exist	0.96	0.72
NGO	0.94	0.72

Notes: This table presents the share of firms filing a vat or cit tax return in the immediate aftermath of the audit (Column (1)) and filing a CIT tax return in the second year post-audit (Column (2)) across each of the tracking categories.

Table A12
Reduced form effect on being closed and informal with different set of controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Prob of being closed						
RD coefficient	0.108**	0.115**	0.112**	0.109**	0.098*	0.080	0.112**
	(0.052)	(0.053)	(0.054)	(0.053)	(0.051)	(0.050)	(0.053)
N	858	858	858	858	858	858	858
Bandwidth	655	655	655	655	655	655	655
	Panel B: Prob of being informal						
RD coefficient	0.083*	0.096**	0.092*	0.090*	0.088*	0.095*	0.092*
	(0.049)	(0.048)	(0.049)	(0.049)	(0.049)	(0.053)	(0.048)
N	858	858	858	858	858	858	858
Bandwidth	655	655	655	655	655	655	655
Station FE	No	Yes	Yes	Yes	Yes	Yes	No
Sample FE	No	No	Yes	Yes	Yes	Yes	No
Fictitious FE	No	No	No	Yes	No	No	No
Sector FE	No	No	No	No	Yes	Yes	No
Risk controls	No	No	No	No	No	Yes	No
Lasso selected controls	No	No	No	No	No	No	Yes

Notes: This table presents reduced form results on the probability of being closed – in Panel A – and the probability of being informal – Panel B. That is, we run equation 1 with the outcome of interest replacing D_j^c, and varying $\mathbf{X}_f$. In Column (1) we do not use any controls, in Column (2) we only include station fixed effects, in Column (3) we add controls for whether the firm was in the VAT or CIT sample at baseline (our main specification), in Column (4) we include a control for whether the firm is "fictitious" defined as being labeled a suspicious firm by the URA, in Column (5) we add sector fixed effects, in Column (6) we add risk controls. The risk controls are 70 separate controls which control for the risk score and expected revenue of each of the underlying 35 risk parameters. Finally, in Column (7) we run a lasso specification to select the controls. In the specification we do not penalize station fixed effects because – based on conversations with the URA – they are an important part of selection. We fix the bandwidth to be the optimal bandwidth selected when running the specification in Column (3). Heteroskedasticity robust standard errors are reported in parentheses.

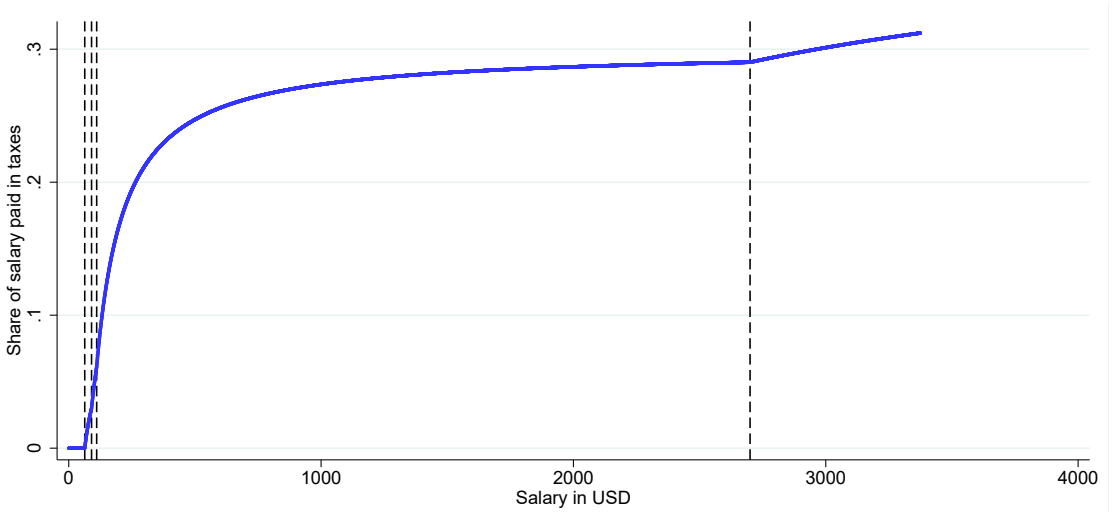
Table A13
Effect of a heavy tax audit on sales, cost and number of employees

	Sales			Cost		
	Prob.	Amount		Amount		Number of
	Positive	Winsor.	Log	Winsor.	Log	Employees
	(1)	(2)	(3)	(4)	(5)	(6)
RD coefficient	−0.082	−238.853	−0.687	−121.005	−0.684	−5.776
	(0.119)	(376.564)	(0.650)	(254.942)	(0.645)	(20.709)
2SLS coefficient	−0.104	−303.348	−0.828	−157.610	−0.876	−7.316
	(0.145)	(459.692)	(0.793)	(319.344)	(0.847)	(25.136)
N	445	393	236	370	215	443
Mean in control	0.65	523.05	5.42	422.37	5.13	25.97
Bandwidth	636	636	636	636	636	636

Notes: This table presents reduced form and two stage least squares estimates for revenue, costs and number of employees reported by firms. In Column (1) we present an indicator for whether the firm reported having positive sales. In Column (2) and (3) we present the amount of sales, winsorized at the top and bottom 1st percentile and the logarithmic transformation. In Column (4) and (5) we present the same transformations, but for total costs. Finally, in Column (6) we present results on the number of employees. Throughout we fix the bandwidth to be the optimal bandwidth for the probability of filing any tax return in the year after the audit. The estimating equations are equation 2, and 3. All regressions include station fixed effects and indicators for whether the firm was in the VAT or CIT sample at baseline. Standard errors are robust to heteroskedasticity.

B Context

Figure B1
PAYE Tax schedule



Notes: This figure shows how the tax rate for PAYE evolves as the income of a person increases.

Figure B2
Desk audit notification letter

RE: TAX COMPLIANCE ADVISORY FOR THE PERIOD ENDED DECEMBER 31, 2020.

We appreciate your continued support and contribution to revenue mobilization for the development of Uganda.

We have reviewed ABC Limited's tax returns for the period ended December 31, 2020 and noted the following tax compliance issues in the company's declarations that require your immediate attention.

1. **Risk 1.**
2. **Risk 2.**

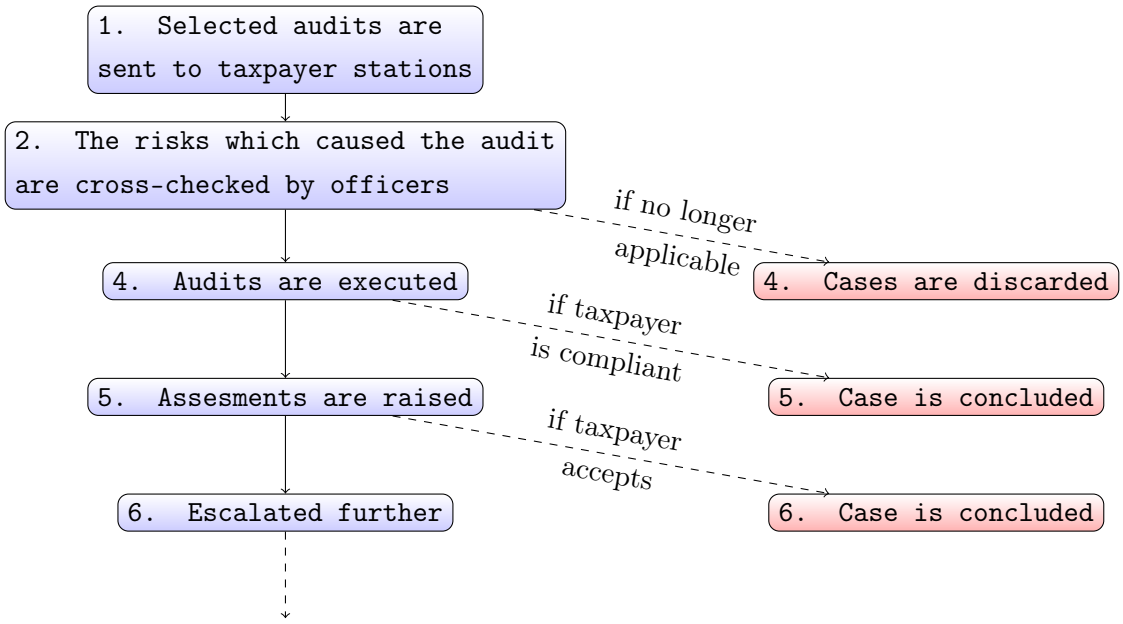
You are therefore advised to amend the company's tax returns to rectify the above issues and pay resultant tax liability.

Provide your response to the above issues by **October 20, 2022.** Note that failure to respond to this communication within the stated timeline may lead to raising of additional assessments and penalties thereof as **per Sec 23(2) and Sec 50 of the Tax Procedures Code Act**, **2014** respectively without further reference to you.

For further guidance and clarification, please do not hesitate to contact the undersigned on Tel; **0772-xxx-xxx** or email xxx@ura.go.ug

We thank you and look forward to your continued cooperation as We Develop Uganda together.

Figure B3
Audit process



Notes: This figure graphically presents the audit process after firms were selected for an audit.