Tax Audits and Their Distortionary Effects[*]

David Henning (UCLA) Joseph Okello Ayo (URA)

First version: October 28, 2024
This version: November 16, 2024

Please click here for the latest version

Abstract

Tax audits are essential for governments to raise revenue but they can create economic distortions. To avoid the financial burden of an audit, firms may remain small, move to the informal sector, or shut down. Leveraging detailed administrative tax data from the Ugandan Revenue Authority (URA), a novel linked survey, and a regression discontinuity design (RDD), we show that audits have a dual negative effect in our context: They reduce the tax revenue collected from audited firms *and* impose large economic distortions. Audited firms are 11 percentage points (p.p.) likelier to shut down, and those that remain operational reduce their output. The former result is driven by firms that must pay substantial back taxes and the latter by firms that believe they are likely to be audited again soon. Back-of-the-envelope calculations indicate that comprehensive audits lead to an overall revenue loss and, to the first order, an aggregate output loss. Our results demonstrate that comprehensive audits impose large costs on audited firms 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. We also extend special thanks to the Central Operations Office for 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 also benefited from many insightful discussions with Jonas Hjort. 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

A core challenge for countries is raising tax revenue while maintaining economic growth. Tax evasion is widespread (Alstadsaeter et al., 2023; Slemrod, 2019), suggesting that tax revenue could be raised through greater tax enforcement. However, tax enforcement measures may–at least in theory–induce distortions that have negative effects on the economy. This is a particular concern when the enforcement is targeted at firms, which are a key driver of economic growth (Aghion et al., 2014). To avoid enforcement measures, such as tax audits, firms may reduce their output, move to the informal sector, or in the extreme, shut down altogether. These responses cause firms to stay inefficiently small, potentially creating misallocation, which reduces aggregate output and economic growth (Bachas et al., 2019; Hsieh and Klenow, 2009). Despite their potentially negative impact on firm output and economic development, the distortionary effects of tax enforcement have received little attention in the literature (Jensen et al., 2024).

This potential tension is especially concerning in developing countries, where the need for revenue is greater and evasion is more widespread (Besley and Persson, 2013; Besley and Persson, 2014) but distortionary effects are likely to be more pronounced because of structural features of the economy. First, formal firms in developing countries are on average smaller than their counterparts in rich countries (Bento and Restuccia, 2021). More limited accounting and auditing capacity due to this smaller scale could make firms more vulnerable to intense enforcement measures. Second, the informal sector tends to be significantly larger in developing economies, offering firms an alternative operating environment where they can avoid all taxes and regulatory costs (Ulyssea, 2018). Finally, statutory tax rates are higher (Enache, 2023), creating a larger incentive for firms to adjust output in the hope of avoiding scrutiny by the tax authority.

In this paper, we investigate how receiving a comprehensive audit by the tax authority affects ex ante formal firms in a developing country, Uganda. Comprehensive audits are the most intense audit interventions. The median such audit takes three months and can investigate any part of a firm's accounts covering the past five years. We combine detailed administrative data with a novel linked survey and a regression discontinuity design (RDD) to document three new findings. First, leveraging the administrative tax data, we show that the potential revenue collected from the audited firms declines post-audit.[1] This reduction is driven by firms exiting the tax system and remaining firms reporting lower tax liabilities. A calculation of the marginal value of public funds (MVPF) for audits, in the style of Boning et al. (2023), yields a negative value. Second, we track down firms two years after a comprehensive audit and find that the audited firms not only change their tax filing behavior but also are likelier to shut down. Third, using our survey, we find that firms that remained operational – in either the formal or informal sector – reduce their output after the audit.

[1] In our data, we never observe how much taxes a firm paid – only how much it owes the tax authority. Thus, our results indicate the *potential revenue collected* or *tax liability* of the firm. We use both terms throughout the paper.

Receiving a tax audit can affect firm operations in several ways. First, the audit induces an immediate cost to the firm because of the hassle cost of dealing with the audit (such as time spent on the audit and accounting costs) and – if found liable – paying back taxes and associated penalties. If these costs are substantial, liquidity-constrained firms could be pushed toward insolvency. Second, an audit may induce changes in perceptions about the enforcement environment. If an audit raises the belief that the authority is able to detect evasion, the very success of this deterrence strategy implies an increase in the firm's expected effective tax rate. Under the higher expected effective tax rate, firms on the margin may be unprofitable. Finally, an audit may cause firms to change their beliefs about the extent to which audit rates depend on firm output. If firms increase their believe that having higher output makes them likelier to be targeted for an audit, it incentivizes them to lower their output for precautionary reasons.

Descriptive evidence from our survey supports the notion that taxation and tax enforcement impose a large economic burden on firms. The total costs of taxation range from 26 to 36% of *sales*, with firms in the lowest sales quintile reporting the highest shares – not because small firms pay more in taxes but because the enforcement costs represent a larger share of sales. An example of an enforcement cost is the cost of hiring an accountant to handle the audit. Furthermore, 63% of firms mention high taxes, tax administration costs or tax corrections as one of the three main obstacles to their performance. Indeed, this is the most frequently mentioned reason, with the next most frequent – competition – being mentioned by only 32% of firms.

Estimating the impact of a comprehensive tax audit is notoriously difficult because firms selected for such interventions differ from others in both observable and unobservable ways. We overcome this challenge by leveraging discontinuities in the audit selection process of the Ugandan Revenue Authority (URA), the Ugandan counterpart of the U.S. Internal Revenue Service. Firms are selected into comprehensive audits based on risk scores calculated for the universe of tax filing firms annually. The risk scores are calculated based on information submitted by firms in previous tax returns. The firms with the highest risk scores are assigned to comprehensive audits, creating a discontinuity in enforcement intensity, with similar firms just above and below the threshold. Importantly, the firms below the comprehensive audit threshold receive lighter-touch 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 potential tax revenue collected. Firms receiving a comprehensive audit are 30 percentage points (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. The effect is not driven by "fictitious" firms that only exist to create and sell fake invoices.[2] The reduction in filing leads to an overall decrease in CIT revenue collected of 5,463 USD per firm audited. We find similar, though less precise, results for the VAT, likely because the

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

VAT data cover only a short time period after the audit. While there is an increase in revenue from the correction of firms' past discrepancies, which highlights one reason why these audits might still be conducted, when we consider post-audit reporting behavior in the year after the audit, the overall effect of the marginal audit on revenue is negative. Specifically, our preferred estimates indicate the URA loses 7,400 USD per audit conducted. A back-of-the-envelope calculation suggests that the effect is not isolated to the marginal audit. Across all the comprehensive audits conducted in our year of analysis, the URA loses 3.6 million USD.

Characterizing the firms that exit the formal sector through a compliers analysis, we find that they tend to be relatively productive *marginally formal* firms. The firms that exit are smaller, have limited capital and are either in the service or wholesale sector. However, they incur substantial tax liabilities prior to the audit and have a high sales-to-cost ratio, suggesting they are important contributors to revenue collection and relatively productive. Lack of capital investments suggests that it is easier for these types of firms to relocate or shut down upon coming under pressure from the tax authorities. While this explains why they respond by exiting despite being productive, these small and productive firms are exactly the type of firms needed to improve employment opportunities and growth (Quinn and Woodruff, 2019).

While the administrative data are rich in detail, it is challenging to measure changes in firm output with administrative tax data. The information in tax returns submitted post-audit is influenced by both real production decisions 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 858 firms in the RDD sample. We implement the survey approximately two years after the audit and are able to verify whether the firm closed or remained open for 91% of the firms.

Leveraging the information from our novel survey, we find strong evidence that comprehensive audits cause firms to shut down. Our estimates show that firms on the margin of receiving a comprehensive tax audit are 11 p.p. likelier to shut down than the firms receiving a light audit intervention. We divide the coefficient by the change in the probability of filing taxes in the two years after the audit and find that slightly more than half of the firms that stop filing taxes (54%) shut down post-audit. The remaining 46% keep operating in the informal sector. Among the closed firms, 43% mention challenges with high taxes and/or the URA as a reason for exit. Crucially, less than half of the respondents that shut down open another firm, underscoring that the exit is long-lasting. The effect is entirely driven by firms that had to pay substantial back taxes. Based on our calculations, for the median firm receiving a tax correction, the total amount it had to pay back represents 86% of baseline gross profits.

We also find evidence that the firms that remain operational reduce their output after receiving a comprehensive audit. To improve our power for the sample for which we have survey data, we move away from the firms immediately around the discontinuity. Combining the survey and administrative data in a difference-in-difference (DD) strategy, we find that comprehensive audits cause operational firms to reduce their sales by 56%. This intensive distortion effect would not be detectable in the administrative tax data alone because the firms either do not file a tax return or report no sales on their tax return. We corroborate that the RDD result is qualitatively similar. Taken together, these estimates demonstrate that comprehensive audits reduce both the number of firms in the economy and the output of operational firms. A back-of-the-envelope aggregation exercise suggests that, to a first order, audits cause a reduction of 0.18 – 0.50% in formal firms' aggregate output.

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 VAT distorts firm-to-firm trade (Gadenne et al., 2022) and that VAT rebates affect firm export performance (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 making themselves appear small (Carrillo et al., 2017) or not filing taxes (Belnap et al., 2024). A common characteristic of this literature is that it relies on administrative tax data. While such data are rich in detail and ideal for studying compliance responses, they do not allow us to observe how much of a firm's response is driven by changes in reporting (evasion) or changes in firm sales (output). We overcome this challenge by linking administrative tax data with a survey on firms. The survey is novel in that it allows us to measure firm output while accounting for changes in evasion behavior. The paper closest to ours in this stream is Harju et al. (2024), which documents that bankruptcy rates increase after a tax audit in Finland. In contrast to these authors, we document that firms also respond in less extreme ways, such as by reducing output, which is crucial for understanding the overall implications for aggregate output. We are also the first to study this question in a low-income context, where the economic costs might be particularly severe. Indeed, to the extent that our measure of firm exit is comparable to the one in Harju et al. (2024), our estimates are 9 times larger.

Second, we contribute to the 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).[3] Estimates based on the MVPF approach suggest that audits have large positive welfare effects, at least in the US (Boning et al., 2023). We are among the first to document that tax audits can have a negative effect

[3] A large literature has investigated the effect of deterrence communication on tax evasion among firms (Almunia et al., 2022; Bergolo et al., 2023; Blumenthal et al., 2001; Boning et al., 2020; Brockmeyer et al., 2019; Carmen et al., 2022; Carrillo et al., 2017; Holz et al., 2023; Pomeranz, 2015; Slemrod et al., 2001) and tends to find positive or nil effects. In addition, see Slemrod (2019), Mascagni (2018) and Alm (2019) for recent surveys of the literature. Our work differs in that we consider the impact of receiving an enforcement *action*, not information.

on the potential revenue collected by the tax authority.[4] The novelty of our finding might be attributable 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 on revenue collection. In the case of the former, the reason for the null result is that the tax audits are simple cross-checks, and in the case of the latter, the paper focuses on a balanced panel of firms, shutting down extensive-margin responses. Thus, our work adds to the previous literature in similar contexts by drawing attention to the fact that a key margin of response from firms is to stop filing taxes altogether.

Third, this paper speaks to a broad literature on one of the key questions in development economics: Why are there so few large firms? An extensive literature in microeconomics investigates different reasons why firms may not grow. Possible explanations include credit constraints (de Mel et al., 2008; McKenzie and Woodruff, 2008), hiring barriers (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 formalization costs (McKenzie and Seynabou Sakho, 2010), and demand-side constraints (Bold et al., 2022; Hjort et al., 2020; Vitali, 2023). However, one of developing countries' defining features is that they have high statutory tax rates and high enforcement rates relative to firm size, potentially making these key obstacles to firm growth. Our results on exit are driven by firms that have to pay substantial back taxes, suggesting that the tax rates may be too high for the firms to operate. To the extent that one would want such firms to exist, Uganda might be on the wrong side of the Laffer curve. Tax enforcement has been recognized as a reason for misallocation in the structural and macroeconomics literature, reducing aggregate output (Bachas et al., 2019; Corbellini, 2024; Di Nola et al., 2021; Leal Ordóñez, 2014).[5,6] However, to the best of our knowledge, there is limited causal evidence on the extent to which tax audits – and, by extension, enforcement – induces firms to reduce their output. Our paper adds to this literature by highlighting that tax audits can distort firm output, and that it matters for aggregate output. Indeed, a back-of-the-envelope calculation suggests that comprehensive audits reduce GDP in Uganda by 0.05% – 0.13%.

The rest of the paper is structured as follows. In section 2, we describe the Ugandan economic context and tax system. Section 3 discusses comprehensive tax audits and the process whereby firms are selected for such audits. In section 4, we present the administrative data used in our

[4] A notable exception is DeBacker et al. (2015), who document that firms in the US evade more after an audit. A theoretical and experimental literature also documents the possibility that firms and individuals evade more in the immediate aftermath of an audit (Alm and Malézieux, 2021; Kasper and Alm, 2022; Kasper and Rablen, 2023). However, empirical evidence for this phenomenon remains limited.

[5] This is part of a broader literature documenting that firms lower their size to locate below regulatory thresholds, leading to misallocation (Boeri and Jimeno, 2005; Evans, 1986; Garicano et al., 2016; Gourio and Roys, 2014; Schivardi and Torrini, 2008).

[6] Microevidence of the distortionary effect of audits in public procurement has been documented by Gerardino et al. (2024).

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. Section 7 introduces the survey and presents results from the survey data. Section 8 discusses our results, including a back-of-the-envelope calculation and aggregation exercise, while section 9 concludes.

2 Context: Uganda's 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 is $3,040 in purchasing power parity (World Bank, 2023b). Its tax-to-GDP ratio – 12.2% in 2020/21 (OECD et al., 2023) – is slightly below the 15.6% average for sub-Saharan Africa (OECD et al., 2023) and substantially lower than the 34.1% average for OECD countries (OECD et al., 2023). However, similarly to other developing countries, the low tax-to-GDP ratio is not explained by low tax rates. The statutory CIT rate is 30%, 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).[7] Among businesses, 72% are estimated to operate informally, and the informal sector accounts for 51% of GDP (Sanday, 2023). These are large estimates by any measure but are similar to those for 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 approximately 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 push the government to enforce taxes more heavily to retrieve as much tax revenue as possible from the small pool of formal firms.

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 VAT and CIT. Below this threshold, a simple sales tax replaces both tax heads.[8] In contrast to VAT and CIT, PAYE must

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

[8] This restriction automatically excludes many of the smallest taxpayers in Uganda; however, from conversations with URA officers, it is clear that these taxpayers rarely experience any form of compliance intervention. Firms can choose to register for CIT and VAT even if they are below the threshold. Reasons they might do so are that they wish to deal with VAT firms or bid for government procurement contracts.

be paid for any employee with a monthly salary above 235 thousands UGX ($\sim$ 60 USD), regardless of firm sales.

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 specifying each of their costs and income sources in their CIT return. The Ugandan fiscal year run from July 1 to June 30, 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.[9] Once the return has been filed, firms can amend it – refile the return with updated information – indefinitely. Not filing a tax return on time costs 200,000 UGX (55 USD) or 2% of the tax liability for the period, whichever is higher.

VAT-registered firms have to submit monthly VAT declarations to the URA for the domestic part of their business within 15 days of month end. As for CIT, firms may file amendments for past returns indefinitely. When less than 5 million UGX ($\sim$ 1,400 USD), negative liabilities can be carried over to subsequent months. Higher amounts can be claimed as a refund, but this 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.

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

2.3 Descriptive evidence on 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 7), and a survey of informal firms conducted by the World Bank (for details on sample selection and methodology, see World Bank, 2019b).

Information from our survey suggests that high statutory tax rates translate to high tax-induced operating 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 rate (Bachas et al., 2023).[11] Using the results from our survey (details in section 7), we document relevant facts about the cost of filing taxes in Uganda, as reported by the firms.

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

[10] In Appendix Figure B1, we show how the tax rate varies with income. It quickly rises to 25%.

[11] The effective corporate income tax rate in Uganda is 15.1%, half the statutory rate (Bachas et al., 2023).

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 sales quintiles. There are three key takeaways. First, the costs of taxation as a share of sales are substantial, ranging from 26 to 36%. Second, the overall costs of taxation are substantially higher for the smallest firms in our survey and are relatively stable over the remaining quintiles, despite tax *payments* being lower for the lowest quintile. Third, enforcement costs to the firm – interactions with the URA and tax corrections – represent a large share of sales for the small firms in our survey. Overall, our survey indicates that taxation imposes a significant economic burden on firms and that the burden of tax enforcement – measured as a share of sales – disproportionately affects smaller firms in the formal sector. Note, however, that this does not mean these firms are small in general. The median firm in the smallest quintile reports sales of 24 thousand USD and has 4 permanent employees. These figures put it above the 40th percentile of the sales distribution among formal firms in Uganda in 2019.

Not only are taxation-induced operating costs high, firms also say that taxation and its associated costs are one of the main obstacles to their growth. In the survey, we asked firms about 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, making these the most frequently mentioned reason, with the second most important reason – competition – mentioned by only 32% of firms. Furthermore, firms seem not to have strong intrinsic motivations to pay taxes. In Panel (b) of Figure 2, we show firms' three main reported reasons for paying taxes. Only 39% mention that it contributes to economic development, and only 20% mention that it helps their business. These facts demonstrate that formal firms in Uganda consider taxes a major obstacle to their performance and do not see many benefits from paying them.

In addition to presenting a major obstacle for formal firms, high perceived costs of formal operations discourage informal firms from formalizing. 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, see World Bank, 2019b). In the questionnaire, nonregistered 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 have not formalized because it is too expensive or it "could be bad for business." These are the two most frequently mentioned reasons. Note that the question pertains to registering with local administrative authorities, *not* with the tax authority. The same set of informal firms are asked what would encourage them to formalize, and the top reason mentioned by 42% of firms is low fees, as shown in Panel (b). This underlines that formalization is seen as prohibitively costly or damaging to business operations by informal firms.

3 Intervention: Audit process and selection

The URA conducts three types of audits (ranked here from most to least intense): comprehensive, issue and desk. In this section, we introduce each type of audit intervention and describe the selection process for it. The focus of this study is the effect of a comprehensive audit relative to that of 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, with the median comprehensive audit taking 93 days.[12] 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 contrast, during a desk audit, the taxpayer receives a letter from the URA detailing a discrepancy detected in their tax return. The taxpayer is then asked to either amend the return to correct the discrepancy or provide documentation explaining it. A draft letter of the desk audit intervention can be found in Appendix Figure B2. Issue audits are a mix of the two prior types of audit interventions. Depending on the case, the audit may look more like a comprehensive audit or a desk audit, which is why we focus on the comparison between audits of these two types instead of issue audits.

A data-driven risk calculation and tax office capacity together determine which taxpayers receive audits of which type. Each year, the Central Operations Office (COO) at the URA picks risk categories to focus on. In our year of analysis, it picked 35 risks. It then takes the tax returns from two years prior and determines whether a given risk applies to a given taxpayer. If so, it calculates a *risk score* for that taxpayer-risk category combination and estimates 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, the COO aggregates the risk score and expected potential revenue of each underlying risk category to compute a taxpayer-specific *total risk score* and *total expected potential revenue.* While the URA avoids auditing the same taxpayer for two years after an audit, tax returns 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 center – the office responsible for the audits – can conduct. Equipped with these numbers, it creates an ordinal ranking of firms. This ranking prioritizes the total risk score over the total expected potential revenue. That is, firms with the highest risk score are always ranked higher. Only for firms with the same risk score does the expected potential revenue determine their rank. The firms with the highest ordinal rank are assigned a comprehensive audit until the capacity of the central audit center is reached.

Having concluded the comprehensive audit selection, the COO calculates the number of issue and desk audits each *local* taxpayer station can carry out. Taxpayers are generally registered with the local taxpayer station closest to their premises.[13] For each taxpayer station, the COO ranks the

[12] 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.

[13] Exceptions are large taxpayers, medium taxpayers, government entities, and oil and gas companies, which each have their own specialized taxpayer office. Taxpayers from the medium taxpayer office form part of our analysis, but the rest do not. These other specialized taxpayer stations do not follow the general compliance plan that the rest of the URA utilizes.

remaining taxpayers – those not selected for comprehensive audits – by their expected potential revenue. The highest-ranked firms within a local station are assigned issue audits, and the next set of firms are assigned desk audits, until the station's capacity 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, the earliest year for which we were able to pin this process down.

Because of the prioritization of 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 (horizontal axis) and their normalized total expected potential revenue (vertical axis). The total expected revenue is normalized by the station-specific cutoff that determines whether a firm receives issue or desk audits. There are two takeaways from Figure 4: First, as mentioned before, the comprehensive audit selection is entirely driven by the firm's ordinal ranking. The highest-ranked firms always receive comprehensive audits regardless of their normalized total expected potential revenue. Second, for firms near but below the comprehensive audit threshold, whether they receive issue or desk audits is driven by the firm's normalized expected potential revenue. Because issue audits are hard to interpret and there are three times more desk audits around the threshold than issue audits, we compare firms assigned comprehensive audits with those assigned desk audits. As a robustness check, we compare firms assigned comprehensive audits with the whole sample and find consistent results.

Upon selection, all audit cases are sent to the relevant taxpayer station, where the supervisor distributes the cases among its officers. The officer who receives the case verifies that the data-driven risks are still present in the tax return before moving forward with the audit.

For comprehensive audits, once the risk is verified, a notification is sent to the taxpayer informing her that she has been selected for an audit and asking her to prepare her 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, he can issue an assessment. An assessment re-issues the taxpayer's return with revised amounts.[14] The taxpayer then has 45 days to object to the assessment. If she objects, the case is sent to a new department at the URA, which will review the audit for accuracy.[15] 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 is found to have underdeclared its tax liability, the auditor is legally empowered to add interest to the outstanding liability and levy fines on the taxpayer. In practice, to reduce the economic burden on the taxpayer, fines are infrequently imposed. Appendix Figure B3 graphically illustrates the audit process.

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

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

The desk audit process is much shorter. Upon verification of the discrepancy by the auditor, a letter is sent to the taxpayer informing her about the exact discrepancy discovered. The letter requests that the taxpayer either amend her return to eliminate the discrepancy or provide documentation for why the apparent discrepancy is correct. The taxpayer has 7 days to respond to the request, after which the officer can issue an assessment. As in issue and comprehensive audits, the taxpayer can object to the assessment; if not, the audit is concluded.

4 Administrative 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 a comprehensive tax audit. In this section, we describe each dataset, discuss how we construct our main analysis sample, 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 returns with audit records and risk scores, all provided by the URA. We also introduce PAYE data from the URA and census data from the Ugandan Bureau of Statistics (UBOS), which we use to contextualize and describe our sample.

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 return includes a full balance sheet and profit–loss statement of the firm, with information on total sales and a detailed breakdown of each type of cost. Our analysis focuses 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.[16]

Administrative tax (VAT) The VAT data contain the universe of monthly returns filed between January 2013 and November 2022. The VAT data compliment the CIT data in two key ways. First, the VAT return is filed monthly, providing us with information at higher frequency. Second, the VAT 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 its risk calculations and corresponds with the baseline CIT data.

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

Administrative tax (PAYE) The PAYE data contain the universe of monthly returns filed between January 2013 and June 2022. The PAYE data complement the other administrative datasets by providing us with the number of employees in the firm. We have no 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 *assigned* from January 2013 to January 2024. This is important because even if an audit is assigned and sent to a taxpayer station, it is not necessarily executed. The data contain 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.[17] The data also record which supervisor was in charge of the audit and the audit outcome. For our analysis, we focus on audits 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 desk audit start date, it does include information on when the desk audit finished. Given the timeline of a desk audit, we assume that the desk audit started 7 days before its completion date. Furthermore, we assume the desk audit did not happen if there is no completion date.[18]

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 taxpayer's discrepancy across 35 different risk categories. We observe not only a firm's total risk score and expected potential revenue but also its risk score and expected potential revenue for each individual risk category. The aggregate risk score and expected potential revenue are an unweighted aggregate of the scores and expected potential 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 he believes 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 afterward is endogenous to the outcome of the first audit intervention.

[17] Based on conversations with auditors at the URA, this assumption seems reasonable. The numbers of executed audits form part of the auditors' annual evaluations; hence, they have every incentive to add every audit to the online system.

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

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 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 visualization of the time period covered by 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. Appendix Table A1 highlights how the sample changes when we impose each restriction.

We start with the universe of firms flagged for at least one risk in financial year 2021.[19] We proceed to remove firms registered at specialized taxpayer stations. This excludes firms in the large taxpayer office, oil and gas office and public sector office. These offices do not follow the compliance improvement plan of the rest of the URA, instead applying their own compliance strategies. Furthermore, administrative capacity at these stations differs from that of nonspecialized offices: The former typically have a higher officer-to-taxpayer ratio. We also remove mining-sector firms because 40% of firms in this are registered with the oil and gas office. The remainder would be a heavily selected sample.[20]

We remove firms audited the year before our year of analysis. Typically, taxpayers are audited two years in a row only if the URA needs to follow up on a prior issue. Including these audits could confound our estimates for two reasons. First, the firms were selected into audits for reasons not related to the risk scores, and hence, the audits are not induced by the instrument. Second, an audit in the previous year could already have changed the firm's behavior, leading us to identify a different treatment effect.

We restrict our sample 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 consider only stations where at least one firm received a comprehensive audit.

After imposing each restriction, 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.

[19] Firms are defined as taxpayers registered as nonindividuals and private entities. This excludes entities registered as clubs, estates or trusts, government bodies, international organizations, and nongovernmental organizations (NGOs).

[20] The sector with the highest percentage of firms registered with one of the specialized offices is manufacturing at 8%.

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 among both formal firms and the universe of firms in Uganda. We leverage information from tax filing records to compare our sample to the universe of tax filing firms and the Census of Business Establishments from the UBOS to compare it to the universe of all firms in Uganda.

Table 1 presents summary statistics for our samples and p-values for the differences in means across the samples. In Column (1), we present averages for all firms that filed CIT or VAT returns at some point prior to the audit year, including firms that did not receive a risk score. Column (2) restricts to the full comprehensive sample, whereas Column (4) restricts to the firms in the comprehensive sample that are used to identify our treatment effects, namely, those that tend to fall within the optimal bandwidth.[21]

The firms in the comprehensive sample – shown in Column (2) – are on average larger than the firms in the universe filing CIT or VAT returns. The former 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, higher tax liabilities. One caveat is that they tend to report approximately half as many employees. Note that the firms in the comprehensive sample are far likelier to file tax returns across each of the three tax bases presented here. This is consistent with the URA's leveraging information from firms' tax returns when calculating the risk scores. Finally, we note that the sector decomposition is very similar across the two samples. Column (3) presents the p-values for the differences in the averages of the full comprehensive sample and the universe of firms. The differences are significant at the 5% level for 24 of the 26 variables.[22]

These differences are more pronounced for the subset of firms in the comprehensive sample that tend to fall within the optimal bandwidth. In this case, the firms are larger across all measures to which we have access, even the number of employees. Despite being larger, they report VAT liabilities lower than and CIT liabilities similar to those of the universe of firms. These characteristics reflect the URA's audit objectives. Comprehensive audits are the heaviest-touch tax compliance intervention that the URA has at its disposal. These interventions are targeted toward large taxpayers not contributing enough (or at all) to revenue collection. Overall, Table 1 demonstrates that the subset of firms we study is large formal firms in Uganda. The median firm within the optimal bandwidth is at the 80th percentile of the size distribution for formal firms.

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 using the information in the PAYE returns and divide the number of firms in our samples of interest for each category by the number of all firms

[21] The optimal bandwidth changes across outcomes. We restrict to firms that fall within the median optimal bandwidth across all baseline covariates used in the balance test, shown in Table A2.

[22] The p-value of a joint F-test on all variables is 0.000, as shown in the penultimate line.

reported in the census. This procedure allows us to indirectly assess the share of firms in each size category.[23] In Appendix Figure A2, we show the share of firms filing taxes and the share of firms in the comprehensive sample for each size category. As expected, both the set of firms filing taxes and the set of firms in the comprehensive sample are heavily skewed toward the upper tail of the size distribution. The number of firms that file PAYE taxes adds up to 95% 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% 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-touch enforcement interventions on firms that are relatively large in the Ugandan context.

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) that there is no manipulation of audit assignment, 2) that observed differences around the discontinuity are due only to the difference in audit assignment, and 3) that changes in audit assignment have large effects on firms' likelihood of receiving a comprehensive audit. We provide evidence that each assumption holds.

Our first-stage equation – capturing the increase in firms' 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 assignment to a comprehensive audit, 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 assignment to a comprehensive audit – the ordinal rank of the last firm *not* assigned to a comprehensive audit. We recenter the threshold by $l + 0.5$ such that the discontinuity is measured midway between the last firm not assigned 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 sets of controls are included to mimic the URA's selection procedure. We include station fixed effects because the counterfactual of treatment – receipt of a desk audit – is conducted by officers at the local stations, not the central audit center. While the central operations office assigns desk audits to an initial set of firms, officers at the stations

[23] There is no unique identifier to match firms across the tax data and the census.

have some leeway to adjust the selection. The station fixed effects thus control for any differential selection into desk audits. We include two indicators for whether the firm is in the VAT or CIT sample because the URA targets comprehensive audits at firms in more than one tax base.

Thresholds While we know how firms are selected for audits, we do not know the exact assignment cutoff. This information could not be retrieved, probably because the URA did not keep a record of it.[24] 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 we impose each sample restriction and rank them from highest to lowest based on their risk score (first) and expected potential revenue (second). That is, we rank firms by their risk score, and for firms that share the same score, we rank them by their expected potential revenue. Once the firms are ranked, we conduct a structural break test to identify the rank at which we observe the single largest change in the likelihood of the firm being *assigned* a comprehensive audit (distinct from *receipt* of an audit). Because we know the initial number of firms assigned a comprehensive audit, 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 assigned to comprehensive and desk audits. Once we derive the threshold, we again rank firms by 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 this rank in the sample of firms assigned to desk and comprehensive audits, we avoid having gaps in the data from our exclusion of firms assigned to issue audits or receiving no intervention. As a robustness check, we include every firm assigned any enforcement measure.

Counterfactual As discussed extensively in section 3, we compare the effects of comprehensive audits with those of desk audits – both because issue audits are hard to interpret and because there are three times as many desk audits as issue audits around the comprehensive audit threshold, as illustrated in Figure 4.

To capture the reduced-form effect of a firm being above the comprehensive audit threshold on an outcome of interest, we run the following reduced-form RD equation:

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

[24] We received the risk scores from the supervisor in charge of audit selection during our year of analysis. The risk scores were stored on his work laptop, and his was the only existing copy of them. He did not have a record of where the threshold was set, meaning that it is unlikely to have been recorded.

[25] Other studies with similar setups have used a score standardized around the cutoff (Busso et al., 2023). We prefer the ordinal ranking because it drives selection into comprehensive audits.

The only difference between this equation and 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 a firm *receiving* a comprehensive audit on the outcome of interest, not just the effect of it being above the threshold. To derive the effect of receipt of 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:

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

Throughout our results, we present both the reduced-form RD estimates and the 2SLS estimates. β_{RF} indicates the former and β_{2SLS} the latter.

Bandwidth We select the bandwidths using a data-driven approach implemented through the `rdrobust` package by Calonico et al. (2014). We use the universal optimal bandwidth but also show that the main results are robust to changes in the bandwidth.

5.1 Increase in the likelihood of receipt of a comprehensive audit

The discontinuity created by the audit selection process creates large differences in the likelihood that a firm receives a comprehensive audit, the URA's most intensive tax enforcement intervention. As outlined in more detail in section 3, comprehensive audits last 3 months on average and can involve examination of all records in relation to all tax bases from up to 5 years in the past. As such, they represent an in-depth examination of a firm's accounts.

Figure 5 shows that crossing the risk score threshold increases a firm's likelihood of receiving a comprehensive audit by 69 p.p., which is 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 receipt of a comprehensive audit, and the share of firms receiving a desk audit, along the ordinal ranking of firms. We construct residualized outcomes to match the variation shown graphically to the regression results.[26]

Figure 5 highlights three important aspects of the audit selection process. First, firms' 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, which underlines the importance of our including 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

[26] The residualized outcome is constructed by running a regression of the outcome variable against a constant, α, station fixed effects, and two indicators for whether the firm is 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 that a firm is assigned a comprehensive audit. Every firm above the threshold was assigned a comprehensive audit.

executed, the firm was not downgraded to a desk audit. Instead, the audit was carried over to the next financial year or dropped. Finally, the probability that a firm receives 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. Auditors' performance is measured in the number of audits they execute and the amount of revenue they collect, incentivizing them to avoid difficult and long-lasting audits.

5.2 Instrument validity

We have shown that the discontinuity causes large changes in firms' likelihood of receiving a comprehensive audit. The two remaining assumptions for the validity of our RD design are that there is no manipulation of audit assignment and that any observed differences around the discontinuity are due only to the difference in audit assignment. We perform a series of balance and validation tests to verify that both assumptions hold.

Manipulation by firms The first concern is that firms might manipulate their risk scores to avoid being assigned a comprehensive audit. To successfully manipulate selection into comprehensive audits, a firm would have to do the following. First, it would need to find out which risk parameters are used to conduct the risk analysis in a given year. This is hard to ascertain. The group tasked with risk analysis is small, there were 35 risk parameters involved in the calculation, and both the officers in charge of the 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 no single officer knows how every risk parameter is computed. Second, a firm seeking to manipulate its risk score would need to find out how its risk profile compares to that of other firms. What matters for selection is not just a firm's own risk score but this score relative to the scores of other firms. Finally, the firm would have to adjust its tax returns from two years prior, as the risk calculations are retrospective. Even if a firm could gain access to someone who knew how each parameter was calculated in the current year, it would have to file an amendment of 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. The assignment process is therefore unlikely to be affected by corruption. Corruption may influence whether a firm *receives* an audit, but this does not affect the validity of our instrument.

A density test is not appropriate in this context because the running variable is the ordinal rank of firms. To alleviate concerns that the ordinal ranking hides 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. (2020) and Cattaneo et al. (2018) for an alternative running variable designed to capture the relative importance of the risk score and expected potential revenue in selection. The alternative running variable is defined as the risk score value plus the expected potential revenue of the firm relative to the maximum revenue of all firms with the same risk score.[28] The density is highly nonnormal toward 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 0.232. A histogram with the density plot and function estimated with the `rddensity` command is shown in Appendix Figure A5.

Balance One might be worried that the URA selection procedure 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 replace the indicator for whether a firm was audited in the regression. The results can be seen in Appendix Table A2. We cannot reject the null hypothesis of no difference for any of the 26 baseline covariates. Second, we conduct a test for the joint significance of all baseline covariates combined and cannot reject the joint null hypothesis of balance for all baseline covariates across various specifications. We conduct this test varying both the bandwidth and whether we impute zeroes for missing values. As seen in the last rows of Appendix Table A2, the p-value of an F-test of joint significance is 0.151 and 0.531 when we use the minimum and maximum optimal bandwidth, respectively. For robustness, we also conduct a dimension reduction exercise, discussed in more detail in section 6.3.

Manipulation by URA One might be concerned that URA officers change the threshold. The central audit center might, for example, adjust the threshold to include or exclude firms with certain characteristics. We show that this is unlikely to be the case. First, if the firms right below the threshold are systematically different from the ones right above, this difference should appear as a violation in the balance tests. As shown in Table A2, we find no evidence of such violations. Second, the difference in the threshold that we determine through our data-driven method and what would be implied by the initial assignment of comprehensive audits is minimal. We discuss robustness to the use of the assigned threshold in section 6.3.

6 Results from administrative tax data

The URA's purpose in conducting an audit is to raise revenue by correcting past evasion and in the post-audit periods through specific and general deterrence. In this section, we show that while audits do detect evasion, post-audit compliance drastically falls both because audited firms exit the tax system and because the remaining firms report lower tax liabilities. As such, audits do not induce specific deterrence. We discuss general deterrence in section 8.

[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 to which f belongs, and r_f be the expected potential revenue of firm f. Then, formally, $zn_f = risk_{j(f)} + \frac{r_f}{\max(r_{j(f)})}$.

6.1 Tax corrections

We first investigate whether comprehensive audits differentially increase revenue potential through detecting and correcting past evasion. Here, we consider the outcome of the audit itself, not firms' filing responses in the aftermath of the audit. From the audit data, we know what corrections of previous tax returns were filed by the URA as a result of the audit. By merging the audit corrections with the original tax returns, we can discern to what extent the corrections increased the firms' tax liability. In contrast to comprehensive audits, desk audits do typically not lead to corrections; instead, the taxpayer is requested to correct her own 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) with 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 (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 zeroes, we also use the inverse hyperbolic sine (IHS) transformation of the outcome variable. All three variations are standard 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 32 p.p. likelier to experience an increase in their tax liability from corrections. The implied effect of receipt of an audit from the 2SLS is an increase of 46 p.p. In Panel (b), we show that the corrections lead to an increase in the tax liability of 29 thousand USD, significant at the 1% level. Scaling the tax liability increase by average baseline sales (as shown in Panel (c)) suggests that the correction amount is approximately equal to 7.3% of baseline annual sales.

Our estimates do not imply that the URA actually collected 29 thousand USD from the marginal comprehensive audit. We do not have information on the actual payment made by each firm; 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 by correcting past evasion, we next investigate the tax filing behavior of the audited firms in the financial years after the audit. We primarily focus on CIT because we have information on filing behavior for two post-audit years for the CIT but only 5 months of post-audit data for the VAT.

6.2.1 Post-audit tax filing

In the period immediately after the audit (June 2022 to June 2023), firms are less likely to file a CIT return. Figure 7(a) shows that firms above the discontinuity are 20 p.p. (35% relative to the control mean) less likely to file a CIT return in the year after the audit. The implied effect of audit receipt, given by the 2SLS, is a decrease of 30 p.p. Both the reduced-form and 2SLS effects are statistically significant at 1%. The effect persists and becomes slightly larger for 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. The result is driven by firms that were marginally selected for a comprehensive audit. That the effects persists for two years indicates that it is driven by firms leaving the formal sector in the aftermath of an audit.

We do two additional checks to verify that the results can be interpreted as real firms exiting the formal sector. In Appendix Figure A6, we show that the result holds when we use an indicator for whether a firm filed either a VAT or CIT return. The reduced-form coefficient suggests a reduction of 16 p.p. and remains statistically significant at 1%. Audited firms stop filing not only CIT returns but also returns for other taxes such as VAT. Nor are the results driven by fictitious firms. It could be the case that the "firms" leaving in fact exist solely to create and sell fake invoices (see Carrillo et al. (2023) and Waseem (2023) for two recent papers exploring this phenomenon in developing countries). To verify that this is not the case, in column (4) of Table A5, we drop firms 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.

Taken together, we interpret our results as strong evidence that audits induce real firms to exit the tax system. We discuss robustness in section 6.3.

6.2.2 Post-audit tax liability

The effects on tax filing are strikingly large but may not have implications for the tax liability amounts submitted by firms in the aftermath of the audit. If the firms that stop filing taxes are small and do not pay substantial taxes at baseline, removing them from the tax system would have little implication for the revenue gathered by the tax authority. Furthermore, if the firms that kept filing taxes submitted substantially higher liability amounts post-audit, tax liability amounts may actually increase despite some firms exiting the tax system, as in Harju et al. (2024). We again focus primarily on CIT but show that the results hold for VAT, as well. Because we are interested in the overall change in tax liability, we impute zeroes for firms that did not submit a tax return.

When we focus on CIT returns for the year after the audit (but filed two years after) to effectively account for business operations being affected by the audit, we find strong negative effects on CIT

[29] We received information on all firms identified as "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 "suspicious."

liability. Note that 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 2, we find that the probability of a positive tax liability for firms that received a comprehensive audit falls by 27 p.p. (53%), which is statistically significant at the 1%. We present three variations of the CIT liability amount variable, all of which point to a negative effect. The winsorized 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 IHS-transformed variable also indicates a large negative effect. RD figures with the equivalent outcomes can be found in Appendix Figure A7. Overall, we take these estimates as strong evidence that comprehensive audits negatively impact post-audit CIT liability amounts.

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 A3. The coefficients are small and statistically insignificant, probably because most comprehensive audits occur toward the end of the financial year and hence are unlikely to affect the operations of the firm. Any differences would therefore be 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, not tax filing behavior.

In Appendix Table A4, we break down the overall decrease in tax liability into the extensive and intensive filing margins. We leverage the reduced-form coefficients for this exercise. As shown in column (2), firms are 23 p.p. less likely to file a tax return in the second year post-audit. Conditional on filing, firms are 19 p.p. less likely to file a tax return with a positive liability, although this estimate is only marginally significant (column (3)). Conditional on filing a return with a positive liability, firms report liability amounts that are approximately 15% smaller, but this estimate is statistically insignificant (column (4)). We conclude that our results are primarily driven by firms not filing taxes., The reduction in the probability of filing a CIT return with positive liability also plays a minor role.

As a final check, we verify that our results are not driven only by the CIT base. Leveraging information on 5 post-audit months from the VAT data, we investigate whether the VAT liability amount also declines. We use the same outcomes as for the CIT analysis and take the cumulative sum across the five months.[30] The results are shown in Table 3. RD figures with the equivalent outcomes can be found in Appendix Figure A8. While the estimates for all variations of the outcome variable indicate a negative effect, the coefficient is significant only in column (2), where the outcome is the winsorized amount of VAT liability submitted. The coefficient suggests that the total VAT liability amount submitted in the first five months post-audit falls by 30,833 USD. While this is a far larger magnitude than what we find for CIT, given that only one of the specifications is significant, we interpret this as suggestive evidence of a negative effect.

[30]We will extend the time frame of the analysis once more recent data become available.

We conclude that comprehensive audits have a negative effect on post-audit tax liabilities for the marginally audited firms through two channels: 1) causing firms to leave the tax system, and 2) reducing their probability of reporting positive liabilities. However, the former force drives the aggregate revenue decrease. We discuss robustness in the section below and then investigate which firms stop filing taxes in the aftermath of the 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 A9. 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 drive the results.

Alternative specifications Next, we show that our results are robust to our using alternative controls. We fix the bandwidth to that used in our main specification and change the controls we include. In Table A5, we focus on whether a firm received a comprehensive audit and whether it filed tax returns. In Table A6, we focus on firms' tax liabilities. In column (1), we include no controls; in column (2), we add station fixed effects; column (3) presents results from our baseline specification; and column (4) removes fictitious firms. 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 that control for the risk score and expected revenue of each of the 35 underlying risk parameters. Finally, in column (7), we run a LASSO specification to select the controls. In both Tables A5 and A6, the effect sizes are stable across a large number of specifications and are almost always statistically significant.

Alternative threshold As discussed in section 5, we can identify the threshold in two ways. Our baseline approach is to conduct a structural break test to find the largest change in firms' 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 A7, we show that our main results are generally robust to us using this alternative threshold.

Alternative sample In our main specification, we restrict the sample to firms assigned comprehensive or desk audits. We then calculate the ordinal ranking of firms based on the restricted sample. One may be worried that because this moves some firms closer to the threshold on the ordinal ranking, our results rely on assignment of different weights to some firms. Alternatively, one may be worried that by comparing firms with different normalized expected potential revenue (see Figure 4), we introduce imbalances in our estimation (though we do not find any evidence of this when we conduct our balance tests). To address both concerns, we run a specification where we include all firms assigned any intervention. That is, we run an RD on all the firms shown in Figure 4. In Table A8, we show that our main results are generally robust to us using this alternative threshold.

Dimension reduction exercise We conduct a dimension reduction exercise where we predict our outcome variables using all 26 baseline covariates. We then run the RD on the predicted outcome. The results are displayed in Table A9 and confirm that imbalances in baseline covariates do not explain the large impacts that we find.

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: Which firms 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 of Pinotti (2017). We conduct the analysis with two outcomes: the probability of a firm receiving a comprehensive audit, and the probability of a firm not filing CIT returns for two years following the audit (our proxy for exit of the tax system). This analysis relies on the monotonicity assumption – that is, that a firm's being above the threshold does not decrease its probability of receiving an audit and does not increase its probability of filing taxes. Given our analysis up to this stage, both assumptions seem plausible on average. However, we recognize that some firms may be likelier to file taxes after being audited because they believe they are under greater scrutiny from the tax authority (there may be some defiers). To the extent that this is the case, our estimates should be interpreted as suggestive.

We run a 2SLS 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 the firm not filing CIT returns for two years following the audit. The characteristics of the compliers are given by θ. Similarly to in 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 or CIT sample, l is the threshold for comprehensive audits, and z_f is the firm's ordinal ranking.

Figure 8 presents results on key variables of interest. The first column ("Sample") in each figure presents the baseline average for the sample that falls within the optimal bandwidth for that variable. The second column ("Audited") is the baseline average of the firms induced into comprehensive audits (the comprehensive audit complier mean). Finally, the third column ("Not filing") is the baseline average for the firms induced to stop filing taxes (the nonfiling compliers mean). Appendix Table A10 shows results for all baseline variables and p-values for the differences between the complier means and sample averages.

We first focus on the set of firms induced into being audited. Firms induced into comprehensive audits have slightly higher sales at baseline (Panel (a)), a lower value of plants and machines (Panel (b)), a lower sales-to-cost ratio (Panel (c)), and higher VAT and CIT liability amounts (Panels (d) and (e)). Among these, only the sales-to-cost ratio is significantly different at the 5% level. They are also less likely to be in the service sector (Panel (f)) and, as can been seen in Table A10, significantly likelier to be in the manufacturing sector. This is consistent with the URA targeting larger firms that pay several tax heads, which are typically firms not in the service sector.

Interestingly, the pattern changes when we focus on the firms induced into not filing taxes. They report lower sales, use less machines and plants and are likelier to be in the service sector. However, the latter is statistically insignificant. This helps explain why they are able to disappear: They do not have significant amounts of capital that would be forgone if they closed shop. However, crucially, they have a high sales-to-cost ratio, especially in comparison to the set of firms induced into audits, and report significantly higher VAT liabilities at baseline. The firms induced into not filing taxes are relatively profitable and important contributors to revenue collection. Again, it is important to highlight that these are not small firms as such. IHS sales of 5.40 put a firm near the 45th percentile of the formal firm sales distribution for 2019.

Our evidence suggests that comprehensive audits induce the *marginal formal* firm – smaller formal firms that do not require a great deal of capital – to exit. These appear to be profitable firms that are important contributors to revenue collection at baseline. This is exactly the type of firm that drives growth (Quinn and Woodruff, 2019) and that tax authorities should want to keep in the tax system.

6.5 Discussion: Net revenue and MVPF

Our 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 amounts of approximately 29,000 USD by correcting evasion among marginally audited firms. This estimate does not necessarily imply that the URA collects 29,000 USD from each marginal comprehensive audit. We do not have information on what each individual firm paid, but leveraging information from audit reports for our year of analysis, we know that the URA collected 60% of the total amount of all the tax corrections it issued. Under the assumption that the URA collects the same share from each firm and that conducting the audit was costless, the marginal audit raised approximately 17,400 USD.

However, post-audit reporting behavior demonstrates that this revenue boost comes at the cost of a reduction in revenue collection post-audit. Our estimates suggest the URA loses 31,000 USD from firms reducing their VAT liability amounts in the first five months for the year after the audit, and 5,400 USD from reductions in CIT liability amounts. This highlights a trade-off. Clamping down on evasion does raise revenue in the short term, but does so at the cost of reducing revenue in the following year(s).

The net effect of the marginal audit on revenue once compliance behavior in the following year is considered is unambiguously negative. Assuming that the URA eventually collects the full amount of tax corrections, that the cost of the audit is zero and that the treatment effects we identify persist only for the period in our analysis, the net effect is $29,000 - 31,000 - 5,400 = -7,400$ USD. That is, even the *upper bound* on revenue collected from the marginal audit is negative. Under the more realistic assumptions that the URA collects 60% of the tax correction amounts and that the VAT treatment effects persist for the entire year post-audit, the net effect is $17,400 - 74,400 - 5,400 = -62,400$ USD.

Estimates from a MVPF approach following Boning et al. (2023) 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 taxpayer's willingness to pay (WTP) to avoid an audit and the net revenue gain from the audit. Formally,

$$MVPF^{\text{audit}} = \frac{\text{WTP to avoid audit}}{\text{net government revenue raised}}$$

As discussed above, the revenue raised for the government from the marginal audit is negative (even before we consider the cost to the government of the audit itself). Because the WTP to avoid an audit is always positive, it is immediately clear that the marginal value of the audit is negative.

This calculation speaks to the effect of the marginal comprehensive audit and does not consider general deterrence – the idea that auditing one taxpayer may affect other taxpayers' behavior – or the overall effect of all comprehensive audits. We return to this in section 8.

7 Tracking marginal firms through a survey

We have shown that firms exit the tax system post-audit and that tax liability amounts decrease substantially in the post-audit period. While the effect on the tax liability is unambiguously negative, this does not imply that the output of audited firms changes. Firms might instead move to the informal sector or maintain the same production levels but evade more. Changes in output are hard to measure with administrative tax data because the information in tax returns is influenced by both real production decisions and changes in reporting behavior. This is especially the case for the group of firms we study: those designated as likely tax evaders by the URA.

To overcome this challenge, we designed and implemented a novel survey of firms on the margin of receiving a comprehensive tax audit. Our survey was designed to determine whether each firm existed, and if it did, what its financial situation looked like. We divided our data collection into two parts. The first, intended to determine whether the firms exist, was an extensive tracking exercise where we tracked down every firm on the margin of receiving a comprehensive audit (that is, falling within the optimal bandwidth). The second component was the survey itself. The survey included detailed questions about each firm's financial situation, such as sales, costs, profits, number of employees, and prices and quantities of goods sold. Because we are 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) 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, these items were designed to capture the full monetary burden imposed on firms by taxation. Because of our close collaboration with the URA, we can link the survey with the administrative tax data.

The survey compliments the administrative tax data in two ways. First, it provides us with a measure of firm sales that is not affected by tax evasion decisions. Second, the survey provides an in-depth look 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 to firms of interacting with the tax authority and the cost of tax corrections in developing countries.[31]

7.1 Data collection

We implemented the survey from November 2023 to July 2024.

Selection We selected 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 RD cutoff for comprehensive audits *and* the various station-specific cutoffs for issue audits. Later on, given budgetary constraints, we redirected our effort to focus on firms within the optimal bandwidth for our measure of tax system exit, that is, the probability that the firm files a CIT return in the two years 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 and to calculate the descriptives and contextual information in section 2. We contacted a total of 858 and 3,323 in the narrow and total samples, respectively. Among the contacted firms, 35% and 26% were surveyed in the narrow and broad samples, respectively.

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

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.[32] 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 URSB moved this registry to an online database, facilitating access to the registry information.[33] The contact information from the URSB and the URA contains details on both firms' addresses and phone numbers. The 2020 Census of Business Establishments contains information only on the firms' geographic location.

Once we had combined all sources 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 avoided the phone number registered with the URA because we knew from a pilot conducted the year prior that some firms know which numbers are registered with the URA. To avoid inciting bias among respondents, we used these latter phone numbers only as a last resort. Second, if we could not reach the firm through a phone number, we sent enumerators to each location on file that was within the Kampala metropolitan area.[34] Finally, if we did not find the firm at any of the locations and neighbors did not know of the firm or what had happened to it, we called firms at the phone numbers registered with the URA.

7.2 Descriptive statics

Table 4 presents summary statistics for the tracking exercise and survey completion rates. We focus on the narrow sample. Of the 858 firms that we attempted to track down, we found evidence on whether the firm existed for 91.0% of them. That is, we either reached the firm or learned from people familiar with it – such as neighbors or previous employees – that it had closed. We label the remaining 9.0% "vanished" and consider them closed in the analysis.[35] To our surprise, 17 firms on our list turned out to be NGOs. Because they registered as a business with the URA, it is unclear to what extent they behave as NGOs or enterprises. We therefore include them in our analysis.

We completed interviews with 35% of all firms in the sample (N = 858) and 44% of the firms that existed at the time of the survey (N = 683). This response rate is slightly below the rates in the World Bank Enterprise Surveys for Uganda, Kenya and Rwanda, which focus on similarly

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

[33] One important limitation is that the online database holds information only for firms that updated their registry after the switch to the online system. Because of resource limitations, it was not possible to look through the manually added information on firms for years prior to the switch to the online system.

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

[35] To verify that assuming the vanished firms had closed is reasonable, we compute the tax filing rates for each of our categories. In Appendix Table A11, we show that the filing rates of the closed and vanished firms are remarkably similar. Among the closed firms, 39% file a CIT in the latest financial year, and 39% of the vanished firms do. The p-value for a t-test on whether the two averages are different is 0.992.

sized firms (World Bank, 2023a). This could be explicable by our sample of firms being harder to reach because of recent (probably negative) experiences with the tax authority, which would make them more suspicious. They also tend to be larger than the firms in the World Bank Enterprise Surveys and therefore busier, making them more likely to refuse a survey. Importantly, we find no differential survey response rates – defined as whether the firm agrees to be interviewed – above and below the discontinuity. We show in Appendix Figure A10 that the coefficient on the probability of completing a survey is small (-0.018) and statistically insignificant.

To check whether firms believed – despite our best efforts – that we were from the URA, we asked enumerators to indicate whether the respondent did so in the survey. Based on qualitative work and our pilot, it was apparent whether the respondent thought we were affiliated with the URA. Of the 308 firms that we interviewed from the narrow sample, 25 (8%) thought we were from the URA. For the broader sample, the equivalent number is 43 (6%). We therefore think that most firms did not have an incentive to misreport their information in the survey.

7.3 Results on firm shutdown

Our first key outcome is firm shutdown. We define a firm as having shut down if it does not file CIT returns for two years post-audit *and* either someone informs us that it closed or it "vanishes." This is important because firms might become "dormant," shutting down production and awaiting a profitable large contract – for example, from the government – to restart. In such cases, firms would keep filing taxes despite not operating. Combining the administrative tax and the survey data allows us to identify real shutdowns. 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 of informality – as opposed to whether the firm is tax registered – because firms do not have an incentive to change their tax registration status when they leave the formal sector. Thus, nonfiling is the definition that most closely resembles the one used in the informality literature (Ulyssea, 2020).

Figure 9 suggests that the majority of firms that stop filing taxes close. In Panel (a), we show that firms above the threshold are 22 p.p. likelier than those below it to have shut down when we track them down approximately two years after the audit. The implied effect of receipt of an audit is 27 p.p. (395%). Both the reduced-form and 2SLS estimates are significant at 5%. In Panel (b), we show the effect on informality. The coefficient on the probability of the firm being informal is also large, suggesting an increase in informality for firms above the threshold of 7 p.p., but this estimate is only marginally significant. The coefficient on the firm's probability of filing a CIT return during the two years post-audit is -20 p.p. Dividing the coefficient on shutdown and informality by the coefficient on nonfiling suggests that slightly more than half of the firms that stop filing taxes (54%) shut down whereas the remaining 46% keep operating informally. Using non-filing as a proxy for shut down, would have cause us to severely overestimate firm the effect of tax audits on shut downs.

Robustness We conduct the same robustness checks for the survey data as we did in the analysis with administrative tax data. In Appendix Table A12, we show that the point estimates are similar when we use different controls, and in Appendix Figure A11, we show that the results are stable over narrow bandwidths.[36]

7.4 Causes of shutdown and subsequent occupations

Firm closures are identified in various ways. Sometimes, neighbors of a firm's former premises told us that the firm had shut down, in which case retrieving information on the cause of the shutdown was difficult. However, in cases where we could reach the former owner or a former employee, we asked the respondent for the cause of the shutdown. We display the reasons mentioned in Panel (a) of Figure 10. While the number of observations is not high (and hence this result should be interpreted with caution), note that the number of answers represents 63% of the firms that we confirmed closed. The most commonly cited reason for shutdown was taxes or troubles with the URA: 43% of the firms mentioned one of these issues as a reason the firm closed. The next most cited category was low demand at 16%.

Finally, we asked the respondent what she had been doing since the firm closed (or, if we could not reach the firm, if the respondent knew what the firm owners were doing now). We find that 39% of the respondents reported being an employee in another firm. Crucially, less than half of respondent (40%) said that they had started a new firm, suggesting that the majority of the firms that closed did not open again. A significant percentage of the respondents also reported being unemployed (14%), indicating that firm closure has devastating consequences for some people. We view this as suggestive evidence that audits not only lead firms to close but also discourage owners from starting new firms in their aftermath.

7.5 Results on output reductions

Despite the lack of differential response rates, we have limited statistical power to assess whether firms reduce their sales post-audit because we interviewed only 44% of the firms in the narrow sample. To gain power, we move away from the firms in the vicinity of the discontinuity and include all 13,014 firms for which we can infer sales in the survey. In other words, we include all firms that reported their sales in the survey or firms that closed or "vanished" and did not report any sales in their tax returns. These include "dormant" companies as well.[37]

In moving away from the discontinuity, we lose the ability to use RD as an identification strategy. We therefore leverage a DD specification, combining the administrative and survey data. Specifically, we use the administrative data to test for pretrends and the survey data to capture the outcome in

[36]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, which leads us to conclude that the results are robust.

[37]This restriction leads us to remove 133 firms (14%) from the sample. These include firms where the respondent was afraid or not allowed to disclose financial information (100) and firms mistakenly labeled as closed (33).

the post-audit period. We run two regressions: one where we use only administrative tax data and one where we substitute these with the survey data for the last financial year. The DD specification is formally written below:

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

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 comprehensive audit selection, and l is the threshold for comprehensive audit assignment. α_f and δ_t are firm and year fixed effects, respectively.

If the parallel trends assumption holds such that, 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 this assumption holds, we run a two-way fixed effect (TWFE) version of equation 6:

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

Here, β_k captures the dynamic treatment effect of the firm being above the threshold on the outcome of interest. All event-study results are shown in Figure 11. We find no evidence of differential pre-trends for any of the specifications. The red squares are the coefficients based on the administrative data, whereas the blue circles indicate the survey data. We discuss below why the results from the two data sources diverge.

A separate concern is what constitutes the post-period. As discussed 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 2022. We thus exclude 2020 and 2021 from the analysis altogether. These years may be affected by the audit, but to a lesser extent, which would mechanically drive the treatment effect toward zero.

Table 5 presents results on sales from the DD specification. In columns (1)–(5), we show results using only the administrative data, whereas in columns (6)–(10), we show results from the estimation replacing the post-audit administrative information with the survey data. Throughout, we use the same set of firms.

Focusing on the results using only the administrative tax data, we find a 20.8 p.p. drop in the firms' likelihood of reporting any sales to the tax authority. Note that this estimate is broadly in line with the reduced-form coefficient on the probability that the firm files 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 their probability of reporting positive sales of 19.3 p.p. (column (2)), in line with the reduction in firms' 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 2022 tax return, 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 zeroes 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 average sales of 259 thousand USD. Note that over half of this reduction is driven by the firms not filing taxes. If we exclude these firms, the decline shrinks to 118 thousand USD, and the coefficient becomes statistically insignificant (column (4)).

The interpretation changes significantly when we leverage the survey data for the same set of firms. Because we restricted to firms for which we can infer sales throughout the analysis, there are no significant differences in the firms' likelihood of reporting sales (column (6)). More interestingly, the reduction in the firms' 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, a result that highlights the importance of us 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 estimate is entirely based on the survey data and hence should not be affected by evasion. The results are significant at the 1% level. This suggests that firms reduce their output after an audit but, instead of reporting an output reduction in their tax returns, they stop filing taxes or reporting sales altogether. Finally, we include firms that exited and estimate the overall effect on sales. Interestingly, the effect on overall sales estimated from the survey data is similar to what we find when we use the administrative data. In 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 in the administrative data (column (5)). This suggests that the extensive-margin effect on sales identified in the administrative data closely resembles the intensive-margin effect in the survey data in terms of overall magnitude. Mechanically, this need not be true.

To verify that our output results are not just due to using a different identification strategy, in Appendix Table A13, we show results for sales leveraging the RD specification. As in Table 5, we include exiters and dormant firms imputing zeroes and leverage both the winsorized and log transformation of the amounts. The standard errors are large. The reduced-form coefficient on the logarithmic transformation of sales (column (3)) suggests a reduction of 49%, which is similar to what we found in the DD specification (57%), though the estimate is imprecise. We also include other measures of firm size, such as costs and total number of employees in our analysis, shown in columns (4)–(6). Our reduced-form estimates suggest a reduction in costs of 49% and in the number of employees of 5.7 workers. However, none of these estimates are statistically significant. Overall, the RD analysis corroborates what we find in the DD analysis: firms that remain operational reduce their size.

7.6 Mechanisms

We have documented that comprehensive tax audits have a negative effect on tax liabilities and firm output. A key remaining question is what causes firms to respond in such a drastic manner. In this section, we explore potential mechanisms behind the firm responses.

7.6.1 A simple firm problem and potential mechanisms

Several mechanisms could drive 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 that there is an output tax τ and that firms produce output y at a strictly convex and differentiable cost $c(y)$.[38] They can choose to engage in evasion by underdeclaring their output $e = y - z$, where z is the amount of output declared to the tax authority. Firms face a convex and differentiable cost $g(e)$ of evading that increases in e. If they are audited, they have to pay back the evaded tax liability at interest rate r. Firms also face a fixed cost of being audited c_a that is independent of the amount evaded and represents the hassle cost of the audit for the firm. 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. Firms have idiosyncratic beliefs about both the audit rate and the extent to which it varies with firm output. Formally,

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

This yields the following optimal evasion and output decision:

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

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

Intuitively, firms will set the marginal cost of evasion, $g'(e)$, equal to the expected marginal gain from evasion (equation 9). In the absence of audits, the gain from evasion would be τ. With audits, the expected gain from evasion decreases in the idiosyncratic belief $\hat{p}(y)$ and interest rate r. In other words, the expected gain from evasion is lower the higher the firm's perceived probability of being audited and the higher the interest rate.

The intuition behind equation 10 is the usual profit maximization condition, setting marginal cost equal to marginal revenue. In the absence of taxes or audits, the optimal input choice would set $c'(y) = 1$. In this firm problem, marginal cost are lower than optimal because of two mechanisms. First, the output tax reduces input purchases. This is analogous to the classic result that tax rates

[38]This is a simplification, but the framework can be easily modified to include a parameter for what portion of expenses are deductible. The logic presented here holds. See Best et al. (2015) and Basri et al. (2021) for recent papers including the additional parameter.

distort labor supply (Mirrlees, 1971), in this case applied to firms.[39] The higher the tax rate, the further away is the input choice from optimality. Second, there is an additional incentive to lower input purchases because audit rates depend on firm output. The more the firm believes the audit rate to depend on firm size, the more the firm will reduce output to avoid being audited.

How might an audit affect the firm's decision? The theoretical framework highlights two broad channels. First, if a firm is audited, the cost of the audit could turn profits negative. That is, while in expectation $\pi_f > 0$, if a firm is audited ($p(y) = 1$), $\pi_f < 0$. There are two parts of the cost: One is the back taxes the firm must pay ($(1 + r)\tau e$) and the other the cost of the audit c_a.

The second mechanism points to the central role of firms' beliefs about their probability of being audited and how it varies with firm output. Receiving an audit could alter $\hat{p}$. An upward revision of $\hat{p}$, according to equation 9, would reduce evasion, increasing the effective tax rate. Under the new effective tax rate, the firm may no longer expect to be profitable. Receiving an audit could also alter $\hat{p}'(y)$. Firms may adjust their beliefs about the extent to which the audit rate varies with firm output. If $\hat{p}'(y)$ is revised upward following an audit, following equation 10, firms will reduce their output to avoid future audits.

While tax morale influences firms' compliance (Luttmer and Singhal, 2014), we do not think it plays a major role in Uganda. It has been documented that tax morale in Uganda is low (Kakumba, 2024). Among Ugandans, 83% say they find it hard to know how the government uses its tax revenue, and less than half (46%) believe that the government uses taxes for the well-being of citizens. Six out of ten (59%) say that most or all tax officials are corrupt, and only 35% say they trust the URA somewhat or a great deal. Finally, one-third of citizens say that other people in the country "often" or "always" avoid paying taxes. Given the overall low levels of tax morale in the country, we deem it unlikely that changes in tax morale induced by audits explain the results.

7.6.2 Empirical evidence

Having discussed the various channels, we combine descriptive and causal evidence to understand which ones may drive the results.

If the key mechanism is the amount of back taxes the firm has to pay, we should observe a greater increase in shutdowns among firms with larger relative tax corrections. To test this, we divide the sample of firms above the threshold into two groups: those above and those below the median ratio of tax corrections over baseline sales. Some firms filed sales of 0 on their tax returns at baseline; we include these in the category of firms above the median. If a firm above the threshold did not file a tax return at baseline, we remove it from the analysis. We run a separate regression for each group, while keeping the firms below the threshold constant.

Figure 12 presents results on firm shutdown for these subsamples. In Panel (a), we present the

[39] Note that under a pure profit tax, τ would not distort the firm's input decision.

results for firms above the median and in Panel (b) the results for firms below it. We observe that the effect is driven *entirely* by firms above the median.[40] The reduced-form coefficient suggests that firms with an above-median tax correction share are 21 p.p. likelier to shut down after the audit. This estimate is statistically significant at the 5% level. For firms below the median, the same coefficient suggests a 0.5 p.p. *decrease* in the probability of shutting down and is statistically insignificant. A key driver of firm shutdown appears to be that firms cannot afford to pay back taxes owed. If a firm does not pay its tax correction, the URA typically does not issue a tax clearance certificate, which is essential for firms to import goods and bid for government contracts. Not paying a tax correction can therefore have significant negative effects on business operations.

While the tax correction share is a key driver of firm shutdown, this does not explain whether the firm response is "rational" – that is, whether, based on the firm's past profits, it could reasonably expect to pay back the amount owed over time. Leveraging information from firms' tax returns prior to the audit, we document that, for the median firm above the threshold that received a tax correction, the correction is worth 20% of annual sales and 86% of gross profit. If we reduce the importance of outliers by winsorizing at the 5% level, the average is 99% of annual sales and 3.7 times gross profits. This is a large economic shock for a firm to absorb, making it plausible that a rational response is for the firm to close shop even if it is productive at baseline. Our results support the notion that Uganda may be on the wrong side of the Laffer curve.

Next, we investigate whether firms' perceptions about their probability of being audited could explain why some firms reduce their output. In our survey, we asked firms what percentage of firms similar to theirs they believe will be audited in the next financial year.[41] We use equation 6 and split the firms above the threshold into two groups: firms whose belief about the probability of being audited was above the median and firms whose belief about this probability was below the median.[42] The results are presented in Figure 13. Firms perceiving their probability of being audited as high drive the reduction in log sales (as seen in Panel (a)). The coefficient on log sales for firms perceiving this probability to be low is statistically insignificant (Panel (b)). In other words, the firms that were audited two years ago and now believe there is a high likelihood that firms similar to theirs will be audited are the ones driving the reduction in sales post-audit. This suggests that changes in beliefs play a role in determining output among operational firms.

This mechanism discussion has implications for tax enforcement policy. First, we show that firms are unable to pay their back taxes. From a policy perspective, this suggests that tax rates are too high for these firms. Under a scenario with no evasion, these firms would not survive. Second, firms overestimate the probability of receiving a comprehensive audit (the median firm believes that 50% of firms similar to itself will be audited in the next financial year). We also observe that the firms

[40] The median is 0. However, for the firms above the median, the median penalty as a share of sales is 13 – a substantial amount.

[41] The exact phrasing of the question was as follows: "Finally, imagine firms similar to yours in size and sector. In your opinion, what percentage of these firms do you think will be audited in the upcoming financial year?"

[42] The median perception of this probability was 50% – significantly above the actual audit rate.

believing this probability to be highest are the ones reducing their output. This suggests that the URA could provide guidance about true audit rates, alleviating some fear among firms.

8 Discussion

This paper has shown that audits not only reduce firms' tax liability in the post-audit period but also affect the output reported by firms. Three central questions remain. What is the effect of all comprehensive audits on the revenue collected by the URA, as opposed to the marginal audit? Does the reduction in output have implications for aggregate output? And, finally, what would the general deterrence effect have to be for comprehensive audits to be effective? We discuss each question in turn.

8.1 Back-of-the-envelope calculation for overall revenue collected

To understand the impact of all comprehensive audits on revenue collection, we conduct two back-of-the-envelope calculations. In the first, we include all comprehensive audits assigned and started in 2021/22 (our year of analysis). Thus, the only audits excluded are comprehensive audits assigned in the year prior or started after the financial year closed. In the second, we restrict to the subset of comprehensive audits of firms registered at regular taxpayer stations. The firms in the specialized taxpayer offices are the largest firms in Uganda. Our analysis thus far has excluded them because tax enforcement strategies differ for these firms.[43]

Firms receiving comprehensive audits are together liable for 42.2 million USD in tax payments because of tax corrections. While we do not know whether each individual firm paid its liability, leveraging information from audit reports for our year of analysis, we know that the URA collected 60% of the total tax correction amount. The revenue collected from comprehensive audits is therefore approximately 25.3 million USD.

However, we know that the tax liability of marginally audited firms declines in the aftermath of the audit. To account for this, we aggregate the CIT liability amount submitted by audited firms at baseline, in 2019/20, and compare it to the amount submitted in 2022/23. On aggregate, the combined CIT liability amount of these firms declined by 7.3 million USD. Conducting a similar exercise for the VAT, comparing the average amount submitted in 2019/20 against the average amount submitted during the first five months of 2022/23, we derive a reduction in average monthly VAT liability of 1.8 million USD. Given that this calculation is based on estimates over 5 months, the total reduction in aggregate VAT liability is $1.8 \times 5 = 9$ million USD.

Putting these calculations together, we find that the URA gains $25.3 - 7.3 - 9 = 9$ million USD. However, bear in mind that the VAT liability is based on estimates from only 5 months post-audit. If we assume that the average reduction in VAT liability stays the same over the entire financial

[43] See section 4 for a discussion about the differences between firms registered with regular and specialized taxpayer offices.

year, the URA gain shrinks to $25.3 - 7.3 - 21.6 = -3.6$ million USD. That is, when we account for firms' behavioral response post-audit, under reasonable assumptions, the URA lost 3.6 million USD from comprehensive audits conducted in 2021/22. These calculations do not account for the cost to the URA of conducting a comprehensive audit and consider firms' behavioral response for only one year post-audit. If we assume that the behavioral responses of firms persist for a second year post-audit, the loss increases to 32.5 million USD.

Restricting our calculations to firms registered at regular taxpayer stations alters the analysis results slightly. Together, audited firms in this subset are liable for 18.7 million USD. Assuming that 60% of their liability was collected, the URA gains 11.2 million USD from conducting comprehensive audits on these firms. On aggregate, the CIT liability of these firms falls by 0.6 million USD and their VAT liability by an average of 0.7 million USD. Assuming that these effects hold for the entire financial year, the URA gains $11.2 - 0.6 - 0.7 \times 12 = 2.2$ million USD. In contrast to the calculations for all firms, the URA does seem to gain revenue from auditing these firms. However, if we assume that the behavioral responses of firms persist into the second year post-audit, the URA would make a loss of 6.8 million USD.

Overall, we conclude that while comprehensive audits do raise money through tax corrections, under realistic assumptions, the URA experiences an aggregate revenue loss from conducting comprehensive audits.

8.2 Back-of-the-envelope aggregation

The results from our survey demonstrate that comprehensive audits not only cause the firms' tax liability to decline but also drive them to reduce output. A central remaining question is whether these impacts have implications for aggregate output.

If we assume that 1) the production technology has constant returns to scale, 2) all firms sell to final consumers or firms not in group g, 3) inputs are perfect complements, and 4) the economy was at a constrained efficient equilibrium before the intervention, the effect of an intervention on the changes in aggregate output for a group of firms g is, to the first order, approximated by the following:

$$\Delta Y_g \approx \sum_{i \in g} [s_i - c_i] \times \% \Delta l_i \tag{11}$$

where s_i is total sales of firm i at baseline, c_i is total costs of firm i at baseline, and $\%\Delta l_i$ is the percentage change in labor induced by the intervention (in this case, the firm having a risk score above the threshold). The full derivation appears in Appendix C. We do not observe quantities of labor; instead, we use changes in the wage bill. By doing so, we impose the additional assumption that labor markets are competitive.

Note that the approximation assumes homogeneous effects across all firms. That is a strong assumption in our context since we know that some firms exited, indicating quite different responses. Closure of a firm has, naturally, a distinct and much stronger response on the amount of inputs used. To account for the heterogeneity in firms' responses, we note that $\%\Delta l_i = (1-\Delta p_e)\%\Delta l_i + \Delta p_e \%\Delta l_i$, where Δp_e is the increase in the probability of a firm exiting induced by the treatment. Firms that exit do not use any labor; thus, $\Delta l_i = -1$ for firms that exit. Plugging this into equation 11, we derive a first-order approximation that allows separate estimates for firms that exit and firms that remain operational.

$$\Delta Y_g = \sum_{i \in g} \Delta \pi_i \approx \sum_{i \in g | operating} [s_i - c_i] \times \Delta \log l_i - \sum_{i \in g} [s_i - c_i] \Delta p_e \tag{12}$$

because we condition on firms existing $\%\Delta l_i \approx \Delta \log l_i$. The full derivation appears in Appendix C.

To estimate equation 12, we note that all components are either readily identifiable in the data or can be estimated with the DD regression from equation 6. We observe firms' sales (s_i) and costs (c_i) at baseline (financial year 2019) in the CIT data. Under the standard DD assumption that firms outside the treatment group (firms below the threshold) are unaffected by the audit, $\Delta \log l_i$ and Δp_e can be predicted from a DD regression. We restrict to the set of firms for which we have survey data (the same sample used in Table 5 and Figure 11), because we can only identify firm shut down for this sample. The outcome variable that we use to estimate Δp_e is the probability that the firm reports sales. Given the restrictions imposed on the sample, the primary reason a firm would not have reported sales is that it closed.

We present results from the DD regression in Figure 14. Panel (a) presents results for the probability that the firm reports sales, whereas Panel (b) report results for the logarithmic transformation of the wage bill. We estimate that audits increase the probability of exit by 21.9 p.p. and that they induce a reduction in the wage bill of 0.606 log points. Both estimates are statistically significant at the 1% level.

With the estimates derived from the DD regression and the data points from the CIT data, we have all the components to calculate equation 12. Note that our estimates represent the average treatment effect on the treated, which is why we estimate equation 12 for the firms above the threshold: the treated group. Multiplying the estimate for $\Delta \log l_i$, 0.606, by the aggregate output of the firms above the threshold that remained operational, we find that comprehensive audits induced a 7.4 million USD decline in aggregate output. A further 3.5 million USD in aggregate output was lost from firms exiting, for a total aggregate output loss of 10.9 million USD. If we assume that the treatment effects we estimate hold for *all* firms above the threshold (not just those on which we managed to collect survey data), the effect increases to 19 million USD for the firms that remain operational and 11.2 million for exiting firms. This would make the total aggregate output loss 30.2 million USD.[44]

[44] Since we do not have survey data for all firms, we assume in this exercise that firms that stop filing taxes have exited.

To derive a sense of the magnitude of the reduction, we compare it to pre-audit aggregate output for different samples. Starting with all 13,014 firms in the comprehensive sample, our estimates suggest a reduction in aggregate output of 1.4–4.0%, depending on whether we assume the treatment effect applies to all firms above the threshold. Repeating the same exercise for all formal firms gives us a decline of 0.18–0.50% in aggregate output for these firms. Finally, our own calculation suggests that formal firms accounted for 26% of GPD in Uganda in 2019. Scaling our results by this GDP share of formal firms, we calculate a range of 0.05–0.13%.

Given the small number of firms affected by audits, these estimates are substantial. While the assumptions imposed are strong, the calculations underscore that audits induce economic distortions that are sufficiently large to be relevant for aggregate output in the economy. Also note that the largest effect on audit is coming from firms that remain operational reducing their size, a margin that was not explored in previous literature.

8.3 General deterrence

One core reason that governments conduct tax audits is to promote *general* deterrence – the idea that auditing one taxpayer may affect induce compliance among taxpayers' that are not audited – as opposed to specific deterrence. What we estimate in this paper is the specific deterrence effect, because estimating general deterrence effects is difficult. However, we can use our estimates from the specific deterrence computation to bound how large the general deterrence effects would have to be for audits to break even.

Recall that our preferred estimate suggests that the URA is losing 3.6 million USD from comprehensive tax audits. Dividing the reduction in revenue by the amount of tax liability filed in 2019 of all firms *not* audited across both the CIT and VAT, suggests that revenue collection among the remaining tax base would have to increase by 1% to make up for the revenue lost from audits.

However, this ignores the output implications – private cost – of audits. The URA is losing 3.6 million USD, but our aggregation exercise suggests that the economy is losing an additional 10.9 million USD. To make up for the 14.5 million USD lost to audits, revenue collection among the remaining tax base would have to increase by at least 4% on average across all firms.

We note that this is a conservative estimate for several reasons. First, it does not account for the cost to the URA of conducting audits. Second, it restricts to losses for one year post-audit, these losses are likely to compound over time. Third, we use the aggregate output estimate for the firms we have survey data on, the treatment effect is likely to at least partially also apply to other firms above the threshold. Finally, we are putting the same welfare weight on money held by the government and money held by private individuals. In a welfare analysis one would put lower weight on the government, because the revenue is raised from distortive taxation (Keen and Slemrod, 2017).

Our discussion illustrates that general deterrence effects are unlikely to render comprehensive audits optimal.

9 Conclusion

In this paper, we investigate the impact of comprehensive tax audits on firms' tax liability and output. We show that the marginal comprehensive audit backfires for the revenue authority. The upper bound on the revenue collected from the marginal audit – when post-audit compliance behavior is accounted for – is $-7,400$ USD. An MVPF estimate for the marginal audit would be negative. These findings starkly contrast with those in the recent literature on tax audits in rich countries, which finds large positive effects for each dollar spent on audits (Boning et al., 2023). A back-of-the-envelope calculation suggests that, across all comprehensive audits, assuming the audit is costless to the government and under realistic assumptions about taxpayer behavior one year after the audit, the URA loses 3.6 million USD.

However, comprehensive audits reduce not only the potential tax revenue collected but also firm output. Firms are more likely to shut down after an audit, and audited firms that remain operational reduce their sales. This underscores that the reduction in revenue is driven not only by changes in evasion but also by adjustments in firms' operational decisions. A back-of-the-envelope aggregation exercise suggests that the comprehensive audits in our year of analysis reduced aggregate output by 10.9 million USD for the audited firms that we could survey. This represents a reduction in aggregate output for all firms in the formal sector of 1.4%.

Accounting for the loss in revenue and aggregate output, the general deterrence effect on the remaining tax base would have to lead to an increase in revenue collection of at least 4%, which we think is unlikely to be the case.

Our results have implications for policy. First, negative tax revenue from the marginal comprehensive audits indicates that the URA should either conduct less comprehensive audits or adjust their selection algorithm. Implications of changes in selection algorithms is something we will explore in future work. Second, the shut down result is entirely driven by firms with substantial back taxes to pay, indicating that the amount of taxes firms have to pay is too high. This suggests tax rates may be too high for firms to operate. Finally, firms are overestimating the probability of being audited, which induces them to lower their output. Credible guidance from the revenue authority on audit rates may induce firms to adjust their beliefs.

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).

Aghion, Philippe, Ufuk Akcigit, and Peter Howitt (2014). "What Do We Learn From Schumpeterian Growth Theory?" In: *Handbook of Economic Growth*. Ed. by Philippe Aghion and Steven N. Durlauf. Vol. 2. Handbook of Economic Growth. Elsevier, pp. 515–563. DOI: 10.1016/B978-0-444-53540-5.00001-X. (Visited on 11/06/2024).

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 (2019). "What Motivates Tax Compliance?" In: *Journal of Economic Surveys* 33.2, pp. 353–388. ISSN: 1467-6419. DOI: 10.1111/joes.12272. (Visited on 09/27/2024).

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.

Alstadsaeter, Annette, Sarah Godar, Panayiotis Nicolaides, and Gabriel Zucman (2023). "Global Tax Evasion Report 2024". In: *PSE Working Papers* halshs-04563948. (Visited on 11/06/2024).

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).

Bergolo, Marcelo, Rodrigo Ceni, Guillermo Cruces, Matias Giaccobasso, and Ricardo Perez-Truglia (2023). "Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment". In: *American Economic Journal: Economic Policy* 15.1, pp. 110–153. ISSN: 1945-7731. DOI: 10.1257/pol.20200321. (Visited on 10/21/2024).

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).

– (2014). "Why Do Developing Countries Tax So Little?" In: *Journal of Economic Perspectives* 28.4, pp. 99–120. ISSN: 0895-3309. DOI: 10.1257/jep.28.4.99. (Visited on 10/17/2024).

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).

Blumenthal, Marsha, Charles Christian, and Joel Slemrod (2001). "Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota". In: *National Tax Journal* 54.1, pp. 125–138. ISSN: 0028-0283. DOI: 10.17310/ntj.2001.1.06. (Visited on 09/27/2024).

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., John Guyton, Ronald Hodge, and Joel Slemrod (2020). "Heard It through the Grapevine: The Direct and Network Effects of a Tax Enforcement Field Experiment on Firms". In: *Journal of Public Economics* 190, p. 104261. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2020.104261. (Visited on 09/27/2022).

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).

Brockmeyer, Anne, Spencer Smith, Marco Hernandez, and Stewart Kettle (2019). "Casting a Wider Tax Net: Experimental Evidence from Costa Rica". In: *American Economic Journal: Economic Policy* 11.3, pp. 55–87. ISSN: 1945-7731. DOI: 10.1257/pol.20160589. (Visited on 11/07/2024).

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).

Carmen, Giselle Del, Edgardo Enrique Espinal Hernandez, and Thiago Scot (2022). *Targeting in Tax Compliance Interventions: Experimental Evidence from Honduras*. Tech. rep. 9967. Washington, DC: World Bank.

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).

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma (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).

– (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).

Holz, Justin E., John A. List, Alejandro Zentner, Marvin Cardoza, and Joaquin E. Zentner (2023). “The $100 Million Nudge: Increasing Tax Compliance of Firms Using a Natural Field Experiment”. In: *Journal of Public Economics* 218, p. 104779. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2022.104779. (Visited on 04/11/2023).

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.

Kakumba, Makanga Ronald (2024). *Most Ugandans Struggle to Access Information on Taxes They Owe and How Revenues Are Used*. Tech. rep. 886.

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).

Keen, Michael and Joel Slemrod (2017). “Optimal Tax Administration”. In: *Journal of Public Economics* 152, pp. 133–142. ISSN: 0047-2727. DOI: 10.1016/j.jpubeco.2017.04.006. (Visited on 03/28/2022).

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).

Mascagni, Giulia (2018). “From the Lab to the Field: A Review of Tax Experiments”. In: *Journal of Economic Surveys* 32.2, pp. 273–301. ISSN: 1467-6419. DOI: 10.1111/joes.12201. (Visited on 03/28/2022).

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).

Mirrlees, J. A. (1971). “An Exploration in the Theory of Optimum Income Taxation12”. In: *The Review of Economic Studies* 38.2, pp. 175–208. ISSN: 0034-6527. DOI: 10.2307/2296779. (Visited on 10/24/2024).

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).

Slemrod, Joel, Marsha Blumenthal, and Charles Christian (2001). “Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota”. In: *Journal of Public Economics* 79.3, pp. 455–483. ISSN: 0047-2727. DOI: 10.1016/S0047-2727(99)00107-3. (Visited on 09/27/2024).

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#advancedDownloadOptio (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.

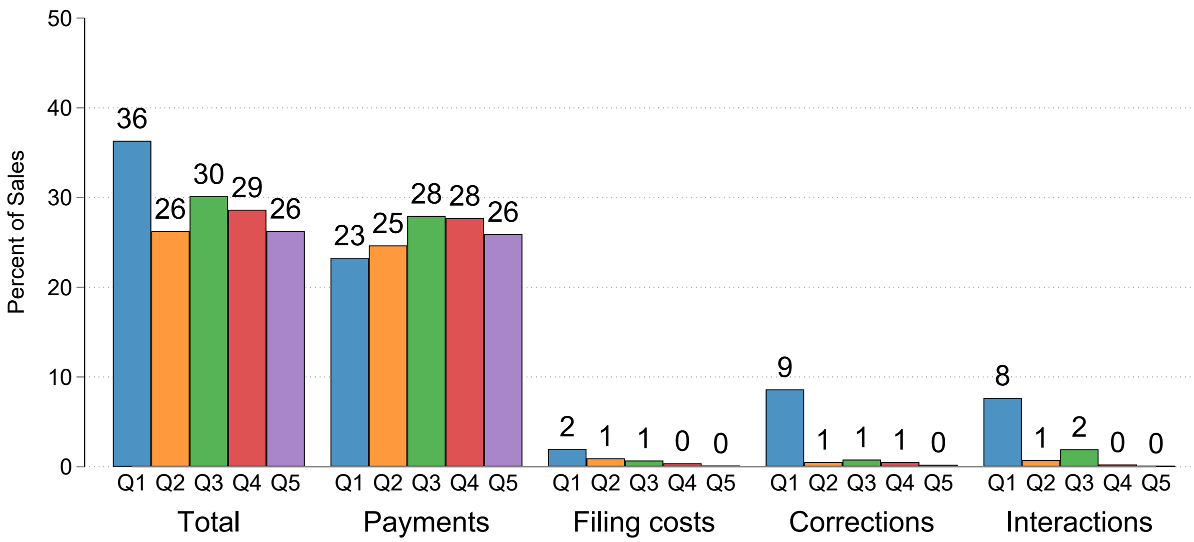
– (2023a). *Enterprise Surveys*.

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

Tables & Figures

Figures

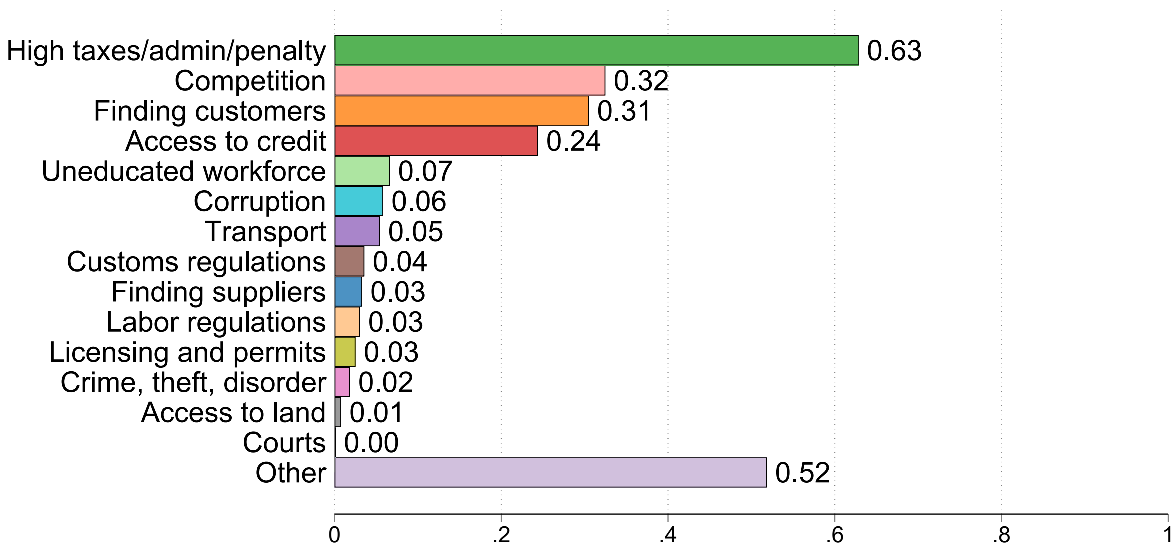
Figure 1
Costs of tax filing as share of sales



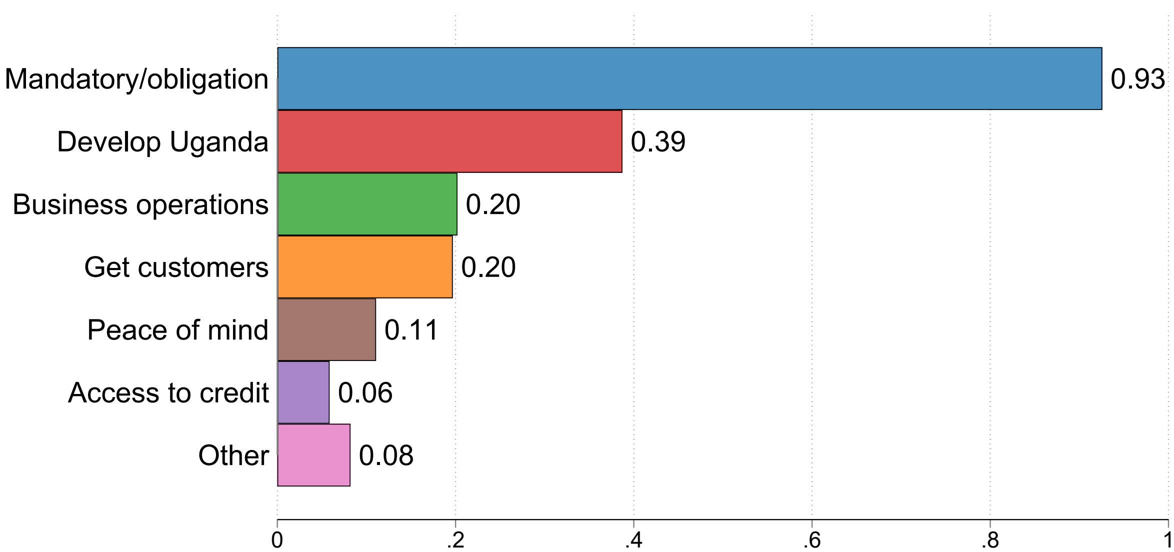
Notes: This figure displays the average costs associated with filing taxes as a share of sales across the quintiles of firm sales. "Payments" is total tax payments across all tax heads. "Filing costs" is the total administrative costs (internal and outsourced) of filing taxes. "Corrections" is the average value of the tax corrections filed by the URA over the last two accounting years. "Interactions" is all costs to a firm of interacting with the URA, including in audits. "Total" is the total costs across all the categories. Each term is presented as a share of firm sales. We restrict the sample to firms that reported sales and 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



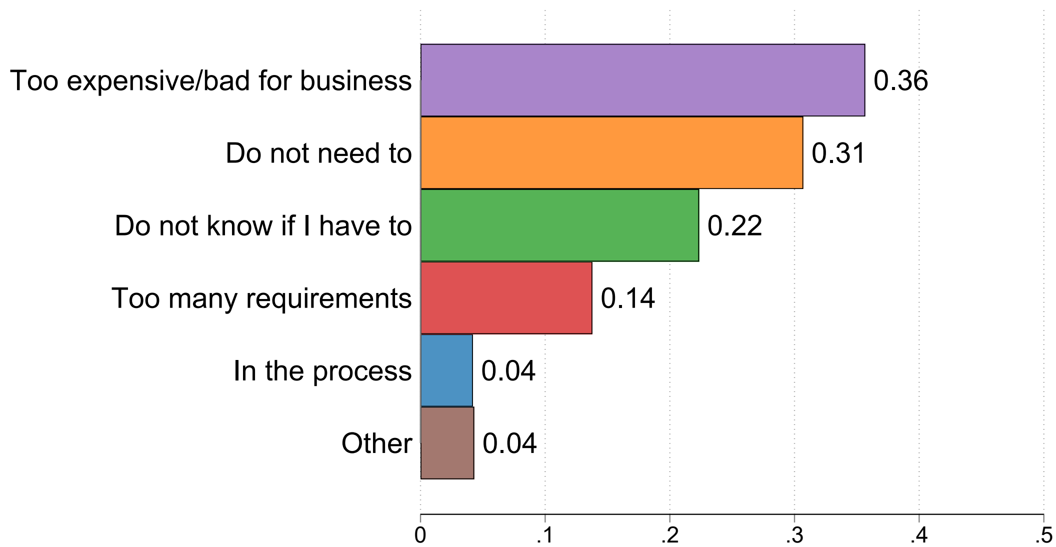
(b) Top three reasons for paying taxes



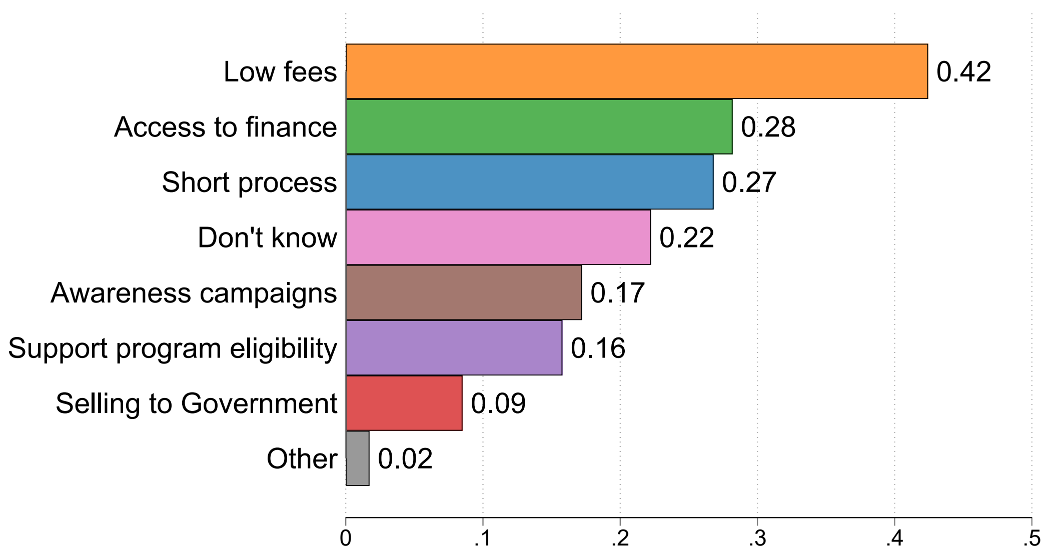
Notes: This figure displays the top three obstacles to firm performance (Panel (a)) and the top three reasons why firms pay taxes (Panel (b)). In Panel (a), the question is a multiple-selection question including the response option "other, specify," where the enumerator could type in the respondent's answer. The exact phrasing of the question is as follows: "We will now ask you about the three most important obstacles to the firm's performance in the last accounting year." The question was asked before discussion of the firm's financials and any questions related to taxes. We aggregate across three response options related to taxes and the URA, namely, "High taxes," "High administrative costs of taxes," and "Tax assessments." N = 754. In Panel (b), the question is a multiple-selection question including the response option "other, specify," where the enumerator could type in the respondent's answer. The exact phrasing of the question is as follows: "What is the most/second most/third most important reason you pay taxes?" N = 767. Source: Authors' survey and computation.

Figure 3
Reasons informal firms do not formalize

(a) Why is the firm not registered?

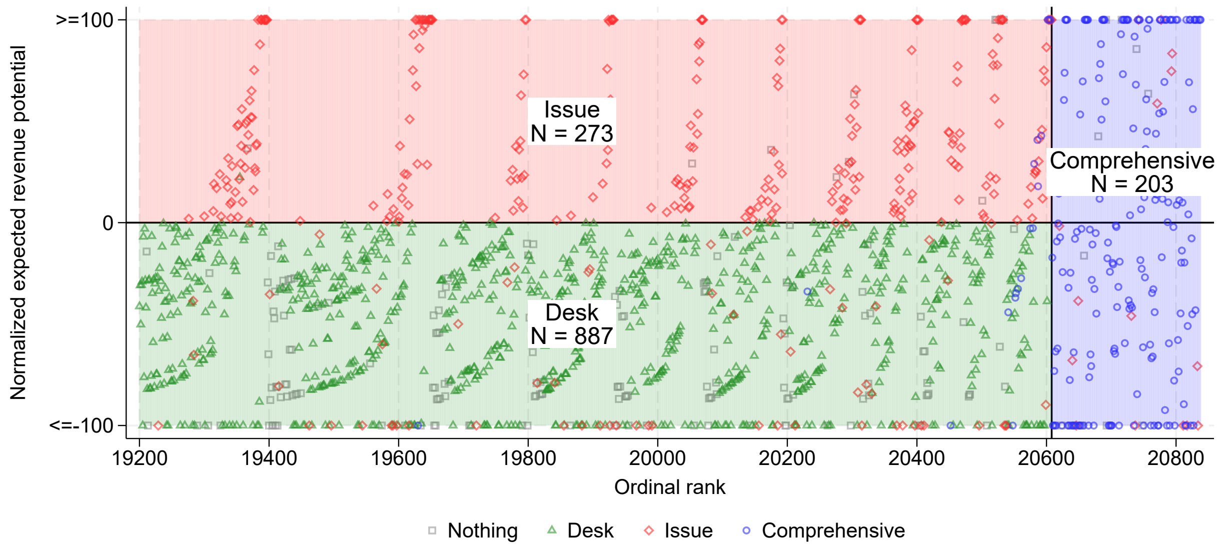


(b) What would encourage firm registration?



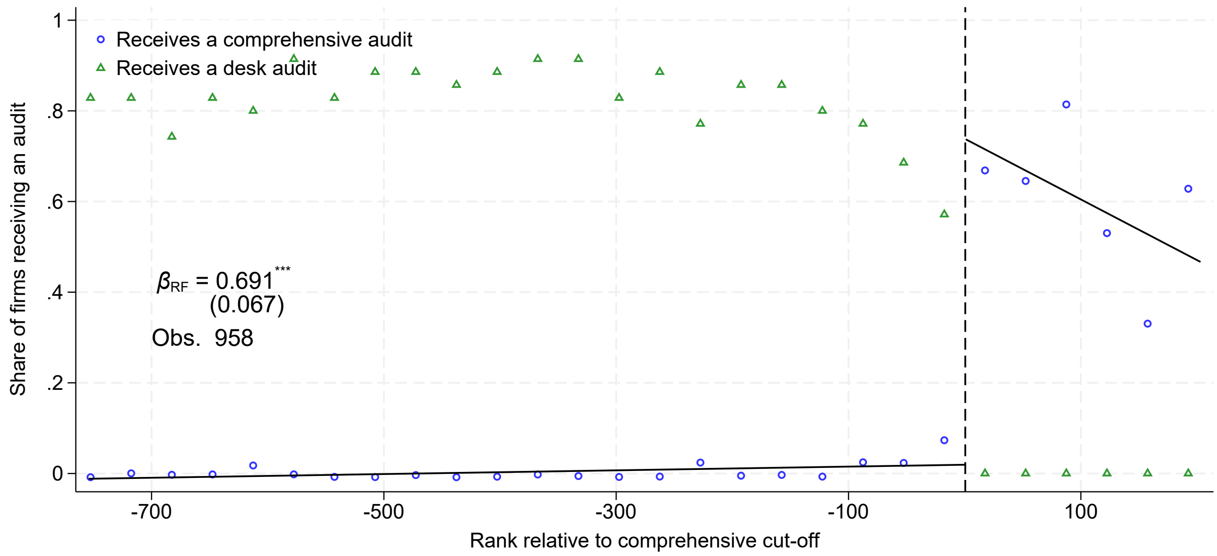
Notes: This figure displays the main reasons why informal firms are not registered (Panel (a)) and what would encourage them to register (Panel (b)). The question is about registration with the Ugandan Registry Services Bureau, not the Ugandan Revenue Authority. In both Panels (a) and (b), the reasons are 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 formalization. N = 1951. Source: World Bank (2019b), authors’ calculations.

Figure 4
Relationship between firm ordinal rank, normalized expected potential revenue, and audit assignment



Notes: This figure presents the discontinuity leveraged for identification. On the horizontal axis is the ordinal rank of the firm, which depends on the discrete risk score (first) and the expected potential revenue (second). On the y-axis is the expected potential revenue of the firm normalized by the expected potential revenue of the last firm receiving a desk audit within the firm's local jurisdiction. Each point represents a firm. We stop the distribution at ordinal rank 19,200 to make it readable. In the main specification, we compare firms assigned comprehensive audits with firms assigned desk audits. Source: Authors' calculations from URA tax data.

Figure 5
Effect of firm relative rank on probability of receipt of a comprehensive audit

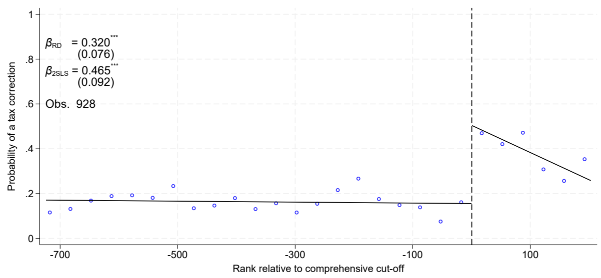


Notes: This figure presents the discontinuity leveraged for identification. The probability of comprehensive audit is residualized. The plots are manually constructed; each point on the graph includes approximately 35 firms. The blue circles show the share of firms receiving a comprehensive audit and the green triangles 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 from Calonico et al. (2014). We present the reduced-form coefficient we retrieve from running equation 1, with heteroskedasticity-robust standard errors presented in parentheses below the coefficient. “Obs.” refers to the number of firms used in the regression. Source: Authors’ calculations from URA tax data.

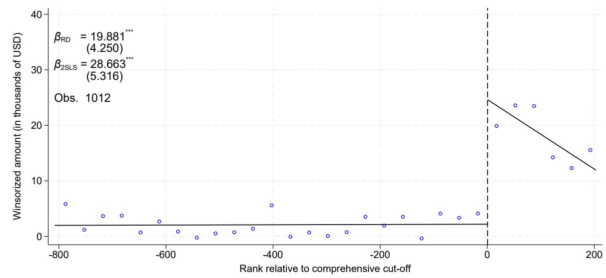
Figure 6

Effect of comprehensive audits on tax liability amount due to tax corrections

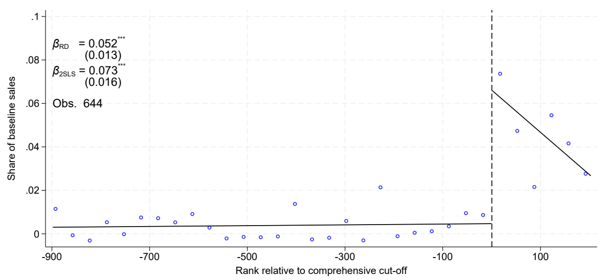
(a) Probability of an increase in tax liability



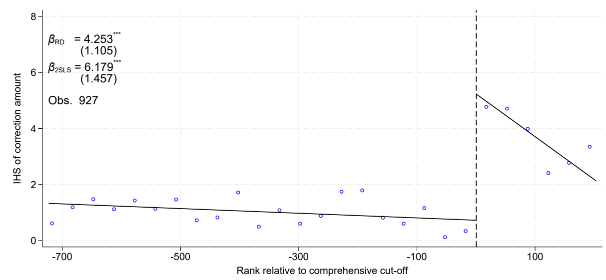
(b) Winsorized increase in tax liability



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



(d) IHS transformation of increase in tax liability

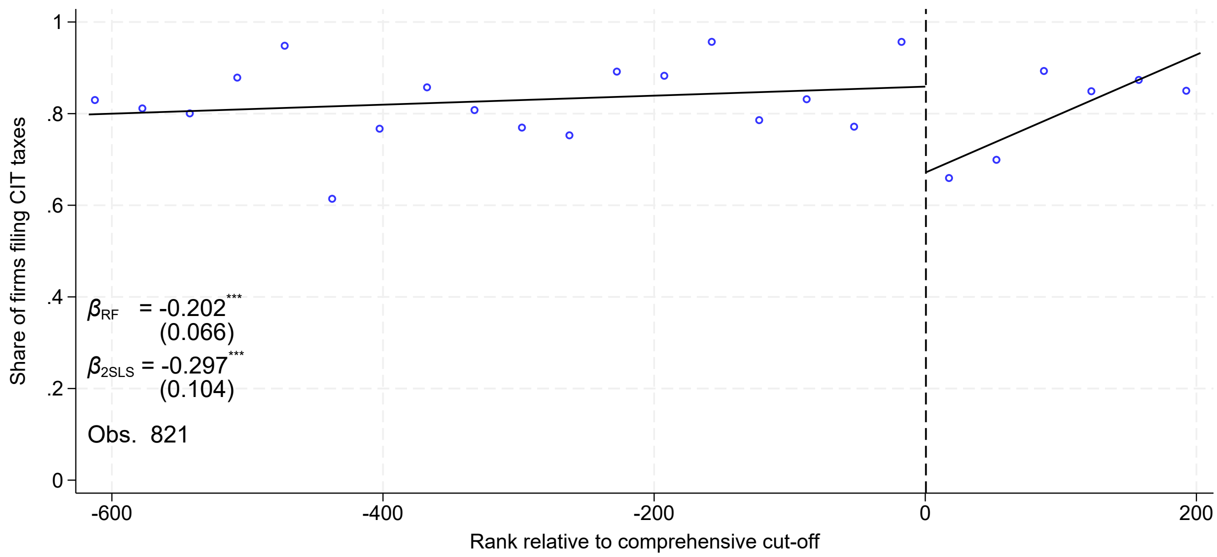


Notes: This figure presents results on the differential increases in tax liability due to tax corrections issued as a consequence of an audit. We compare the increase in tax liability from tax corrections (among firms receiving comprehensive audits) against the increase in tax liability from self-corrections (among firms receiving desk audits). In Panel (a), we report the probability of the tax liability increasing. In Panel (b), we report the increase in tax liability in thousand USD, winsorized at 1%. In Panel (c), we present the increase in tax liability as a share of baseline sales. In Panel (d), we present the inverse hyperbolic sine transformation of the amount of tax liability. The outcome variables shown in the figures are residualized. The plots are manually constructed; each point on the graph contains approximately 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 with the `rdrobust` package from Calonico et al. (2014). β_{RF} shows the reduced-form effects—from equation 2—and β_{2SLS} the 2SLS effect—from equations 1 and 3. Heteroskedasticity- robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: Authors' calculations from URA tax data.

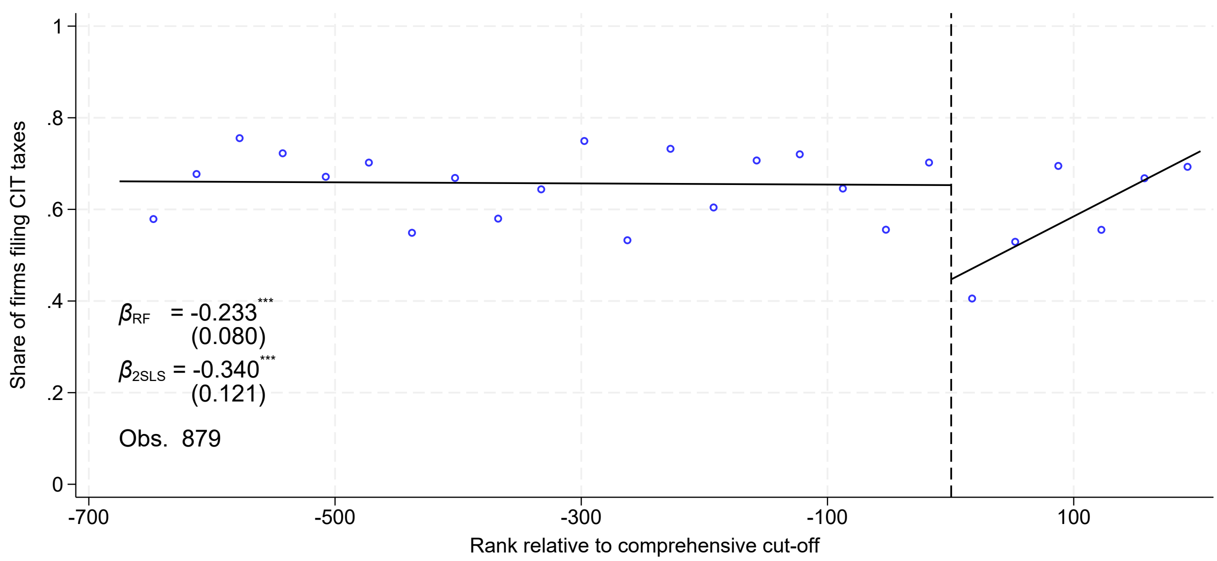
Figure 7

Effect of comprehensive audits on the probability that the firm files CIT returns

(a) Filed CIT return one year post-audit



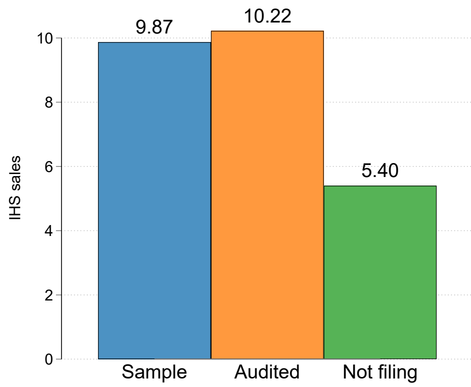
(b) Filed CIT return two years post-audit



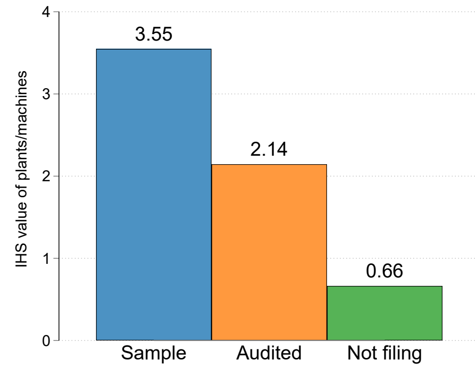
Notes: This figure presents results on the probability that the firm files CIT returns after an audit. In Panel (a), we report the probability that the firm files a CIT return in the year after the audit (July 2022 to June 2023). In Panel (b), we report the probability that the firm files a CIT return in the second year post-audit (July 2023 to June 2024). The outcome variables shown in the figures are residualized. The plots are manually constructed; each point on the graph contains approximately 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 with the `rdrobust` package from Calonico et al. (2014). β_{RF} shows the reduced-form effects—from equation 2—and β_{2SLS} the 2SLS effect—from equations 1 and 3. Heteroskedasticity- robust standard errors are shown in parentheses under the coefficients. "Obs." refers to the number of firms used in the regression. Source: Authors' calculations from URA tax data.

Figure 8
Baseline characteristics of compliers

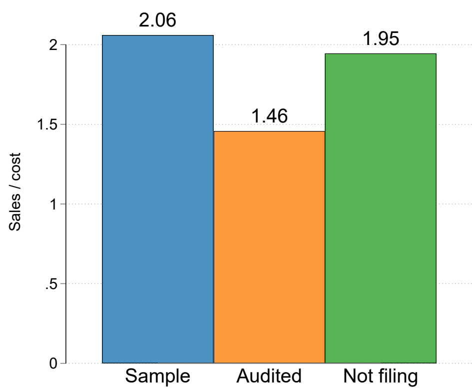
(a) Inverse hyperbolic sine of sales



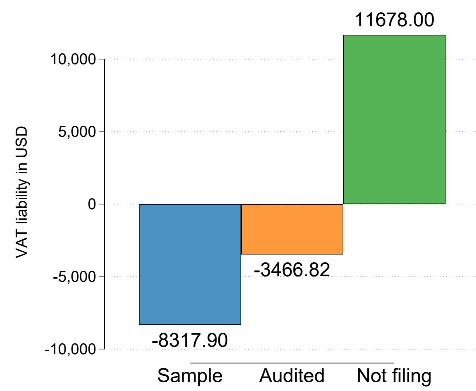
(b) Inverse hyperbolic sine of plants and machines



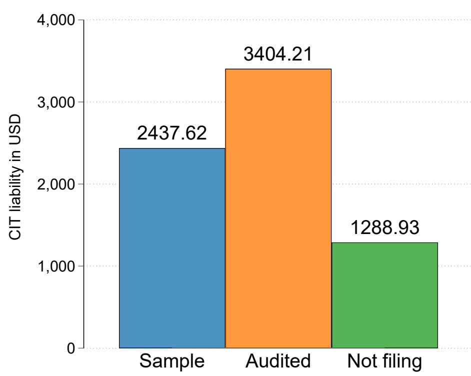
(c) Sales-to-cost ratio



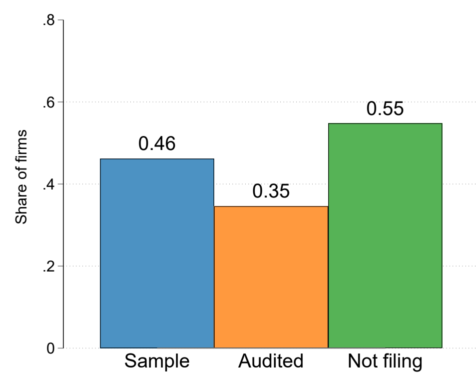
(d) VAT liability in USD



(e) CIT liability in USD



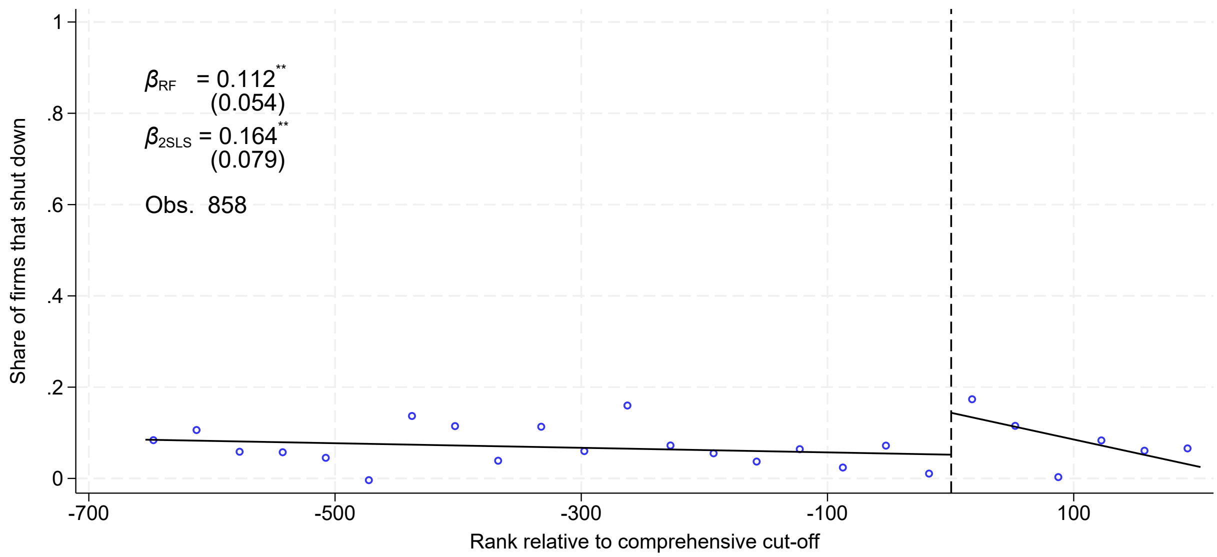
(f) Share of firms in the service sector



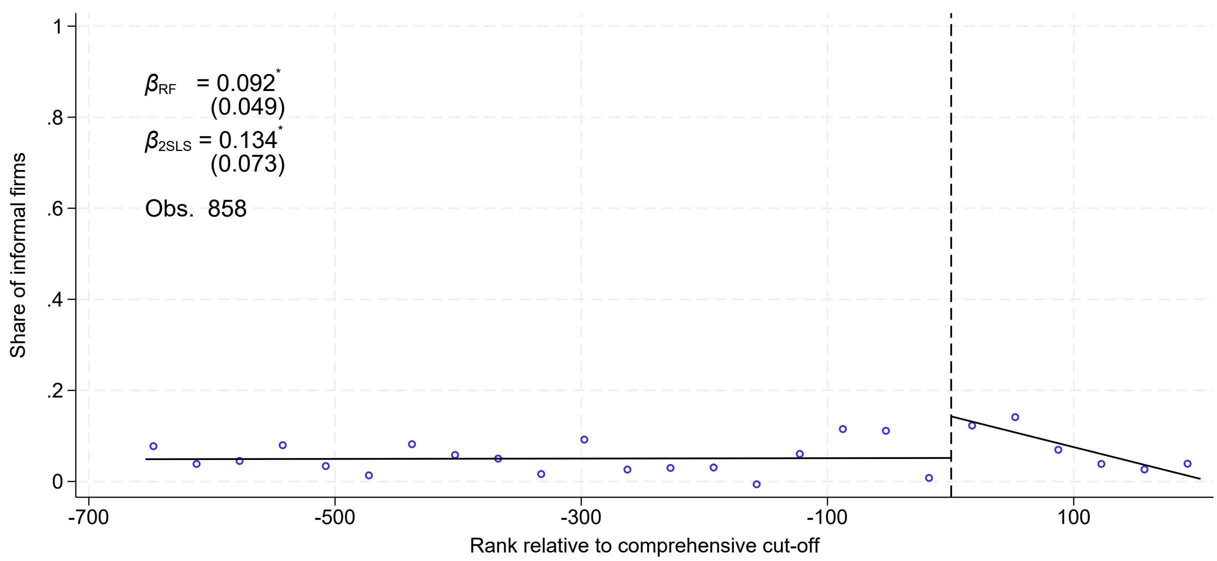
Notes: This figure presents the baseline characteristics of compliers. All variables are based on information submitted in the returns for financial year 2019/20. The first bar ("Sample") presents average baseline characteristics for the sample within the optimal bandwidth of the variable in question. The second bar ("Audited") presents the baseline characteristics for the firms induced into comprehensive audits. The third bar ("Not filing") presents baseline characteristics for firms induced into not filing for two years after the audit. The estimates in bars two and three are derived from equations 4 and 5. In Panel (a) we present the inverse hyperbolic sine transformation of sales, in Panel (b) we present the inverse hyperbolic sine transformation of the value of plants and machinery reported in a firms CIT return. In Panel (c) we show the sales-to-cost ratio computed using the information in a firms CIT return. In Panel (d) we show the total amount of VAT liability submitted, and in Panel (e) we show the total CIT liability submitted. In Panel (f) we show the share of firms in the service sector. Bandwidths used are the optimal bandwidths calculated with the `rdrobust` package from Calonico et al. (2014).

Figure 9
Effect of comprehensive audits on firm closure and informality

(a) Probability that the firm closed



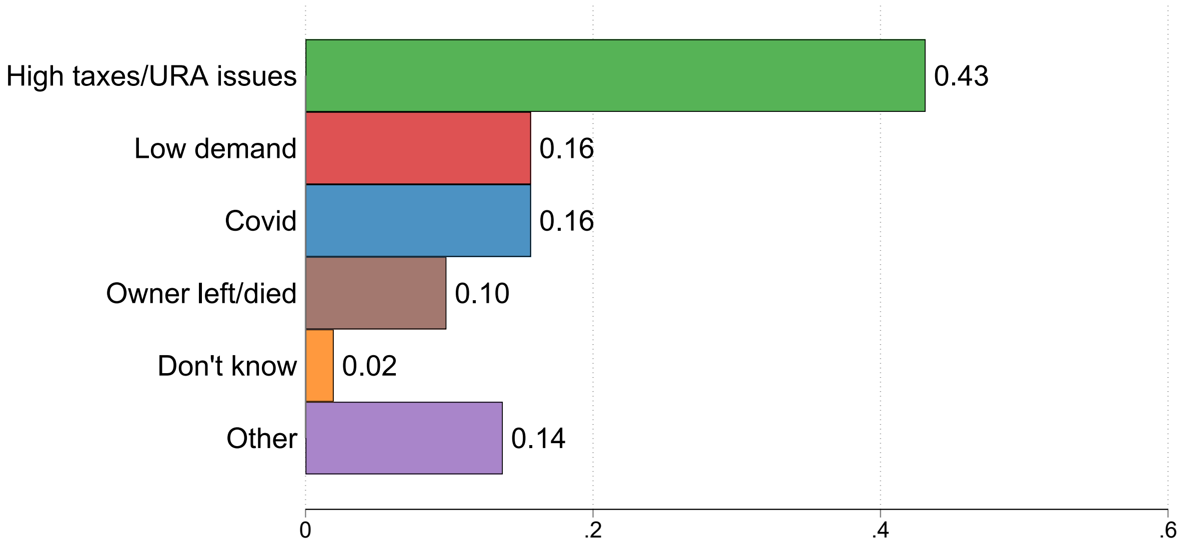
(b) Probability that the firm is informal



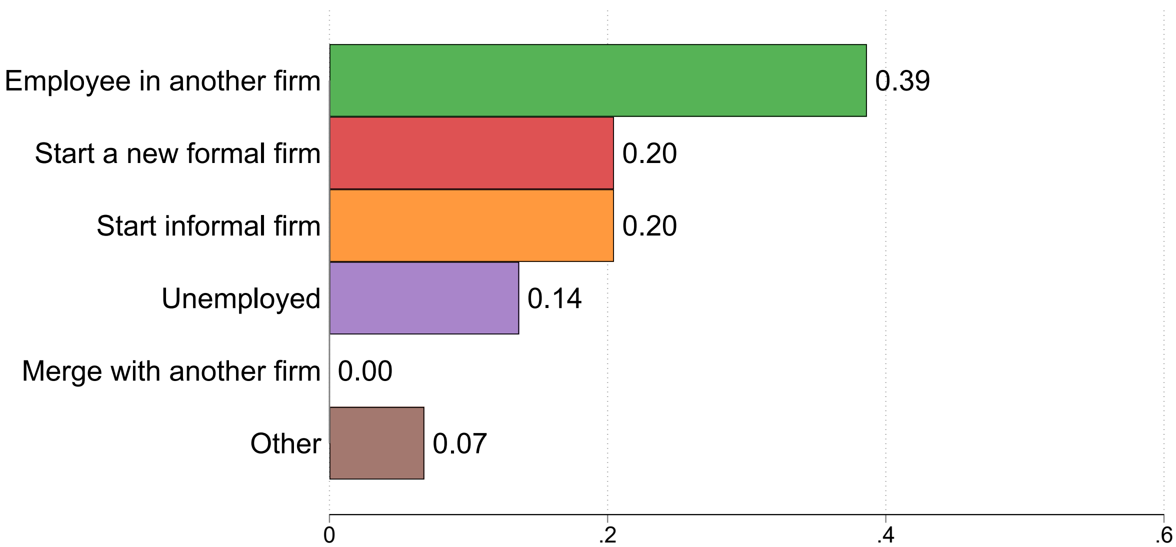
Notes: This figure present results on the probability that the firm had closed (Panel (a)) or was informal (Panel (b)) approximately two years after the audit. "Probability that the firm closed" is an indicator for whether the firm closed or vanished *and* did not file a CIT return for two years after the audit. "Probability that the firm is informal" is an indicator for whether we found the firm, but it did not file a CIT return for two years after the audit. The outcome variables shown in the figures are residualized. The plots are manually constructed; each point on the graph contains approximately 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 with the `rdrobust` package from Calonico et al. (2014) for the outcome variable: whether the firm filed a CIT return in the two years after the audit. This is because we tracked down all the firms within the bandwidth for this variable (our closest proxy of firm exit in the administrative tax data). β_{RF} shows the reduced-form effects—from equation 2—and β_{2SLS} the 2SLS effect—from equations 1 and 3. Heteroskedasticity-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 10
Reasons for firm closure and respondents' subsequent occupation

(a) Why did the firm close?



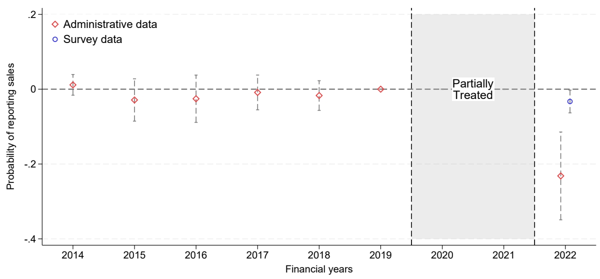
(b) What is the respondent doing now?



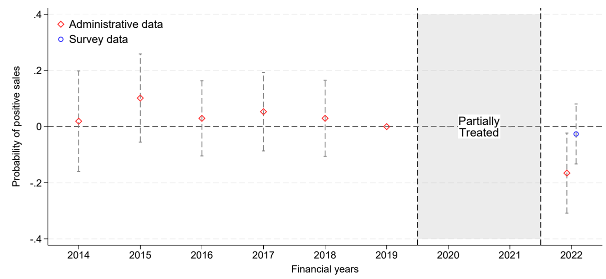
Notes: This figure presents reasons for closure (Panel (a)) and the respondent's subsequent occupation if the firm closed (Panel (b)). In Panel (a), the question was asked of a respondent who could confirm that the firm had closed. The exact phrasing of the question was: "Do you know why it Closed shop?". The question was asked as soon as the respondent had confirmed the firm had closed, and it was an open-ended question. We grouped the responses into the categories shown here. N = 52. In Panel (b), we show the current occupation of the respondent who confirmed that the firm had closed. The exact phrasing of the question was either "Do you know what they are doing now?" or "What are you doing now?", depending on whether we reached someone from the firm. This was a multiple-selection question including the option "other, specify," where the enumerator could type out the answer. The question was asked immediately after the question on why the firm had closed. Occupations are mutually exclusive. N = 44. Source: Authors' survey and calculations.

Figure 11
Effect of comprehensive audits on firms' sales

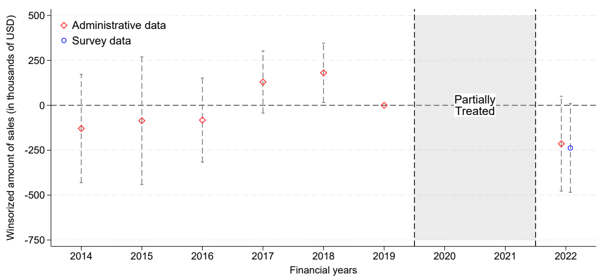
(a) Probability that the firm reports any sales



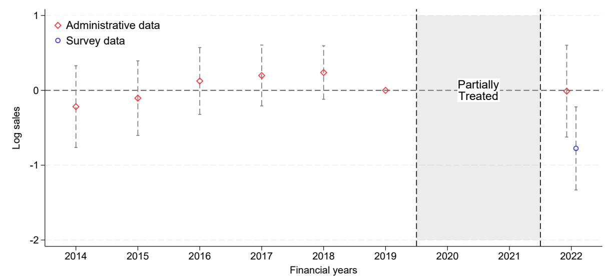
(b) Probability that the firm reports positive sales



(c) Winsorized amount of sales



(d) Logarithmic amount of sales

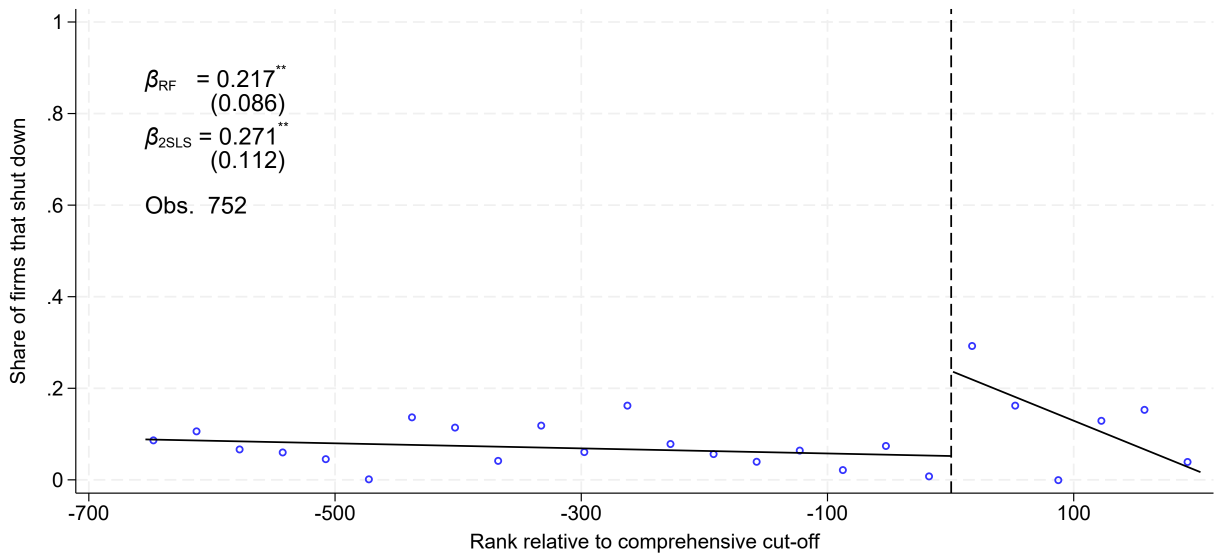


Notes: This figure presents the results of equation 7 from the administrative data and a combination of the administrative and survey data for each of the outcomes shown in Table 5. The blue circles indicate the coefficients estimated from the survey data for the last year (2022), and the red squares represent coefficients estimated from the administrative tax data. We restrict to firms that 1) reported sales in our survey; 2) reported that their last active accounting year was before 2022, so we can infer that their sales in 2022 were zero; or 3) we know closed down, and hence we can also infer that their sales in 2022 were zero. In Panel (a), we present an indicator for whether any sales information was reported (i.e., it was not missing); Panel (b) uses an indicator for whether the firm reported positive sales, conditional on it submitting any sales information. Panel (c) is the amount of sales reported in thousand USD, winsorized at the top and bottom 1 percentiles and with zeroes imputed for firms that did not report sales. Panel (d) is the logarithmic transformation on the amount of sales, effectively conditioning on the firm's reporting sales in 2022. The shaded area indicates the two years of data that we remove when running equation 6. Standard errors are robust to heteroskedasticity and clustered at the firm level.

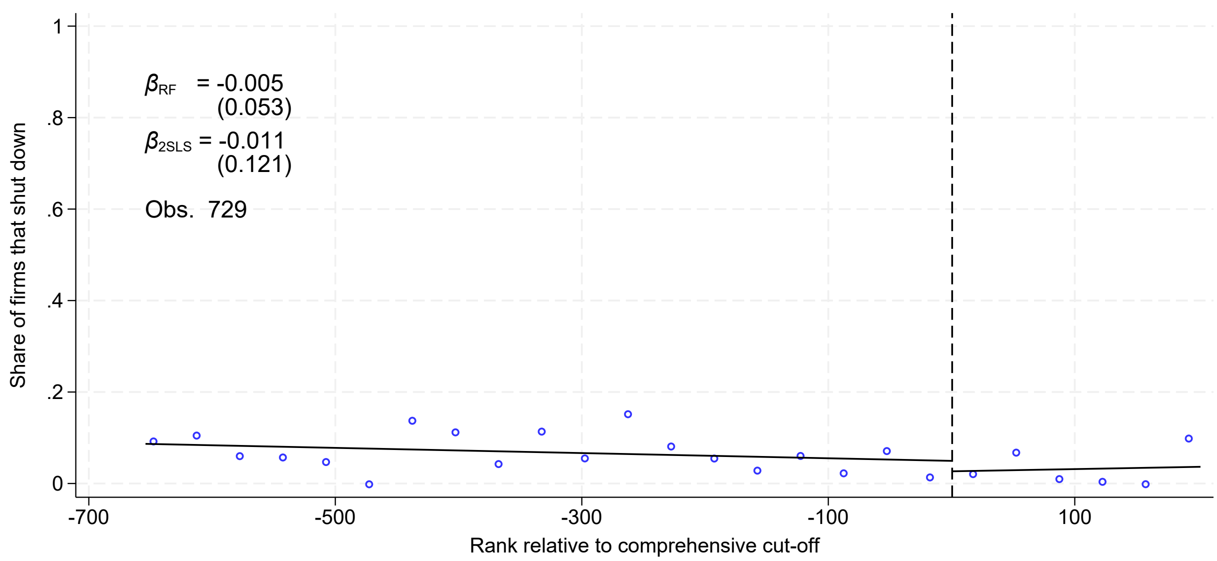
Figure 12

Effect of comprehensive audits on firm shutdown separately by correction share

(a) Correction share above median



(b) Correction share below median

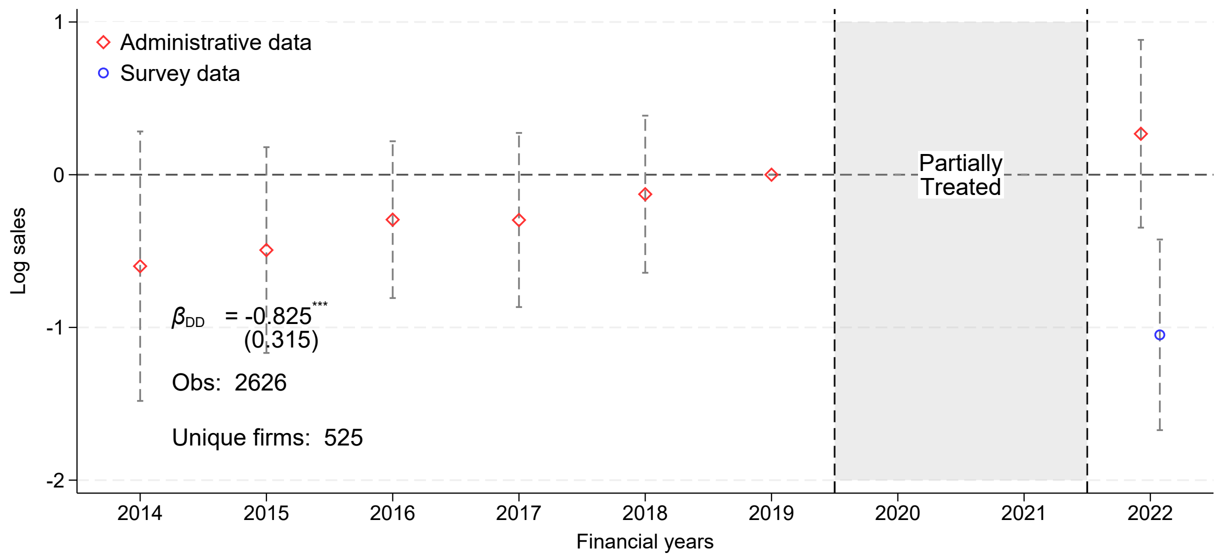


Notes: This figure present results on the probability of a firm's being closed approximately two years after an audit for firms with above-median (Panel (a)) and below-median (Panel (b)) tax corrections as a share of baseline sales. If a firm reported zero sales at baseline, we include it in Panel (a). The outcome variable is an indicator for whether the firm closed or vanished *and* did not file a CIT return for two years after the audit. The outcome variables shown in the figures are residualized. The plots are manually constructed; each point on the graph contains approximately 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 with the `rdrobust` package from Calonico et al. (2014) for the outcome variable of whether a firm filed a CIT return in the two years after the audit. This is because we tracked down all the firms within the bandwidth for that variable (our closest proxy of firm exit in the administrative tax data). β_{RF} shows the reduced-form effects—from equation 2—and β_{2SLS} shows the 2SLS effect—from equations 1 and 3. Heteroskedasticity-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.

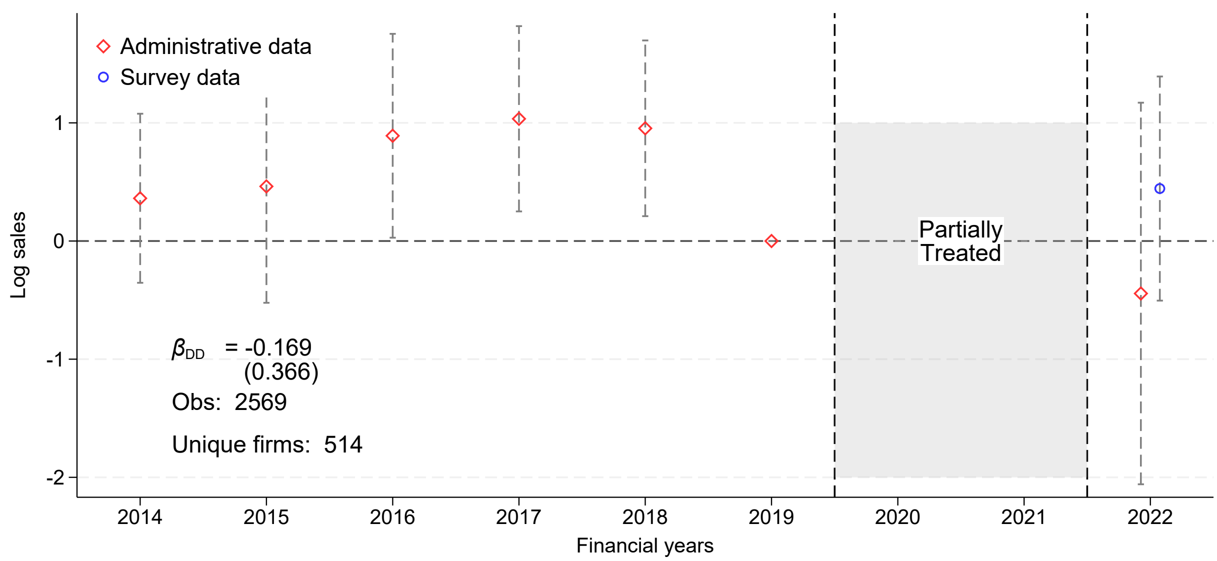
Figure 13

Effect of comprehensive audits on sales separately by belief about audit probability

(a) Belief of audit probability above median



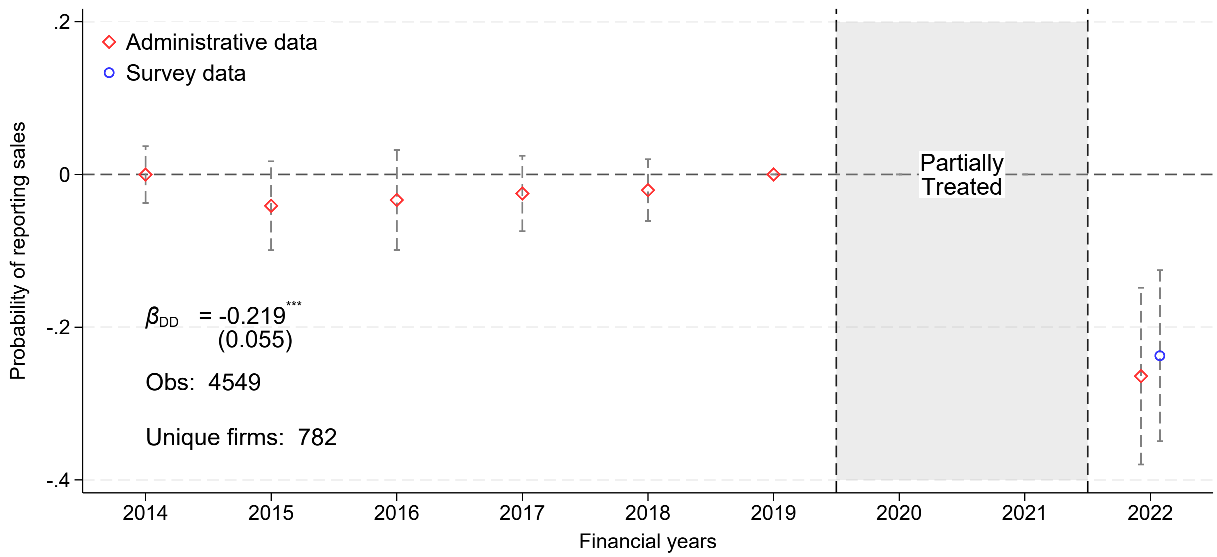
(b) Belief of audit probability below median



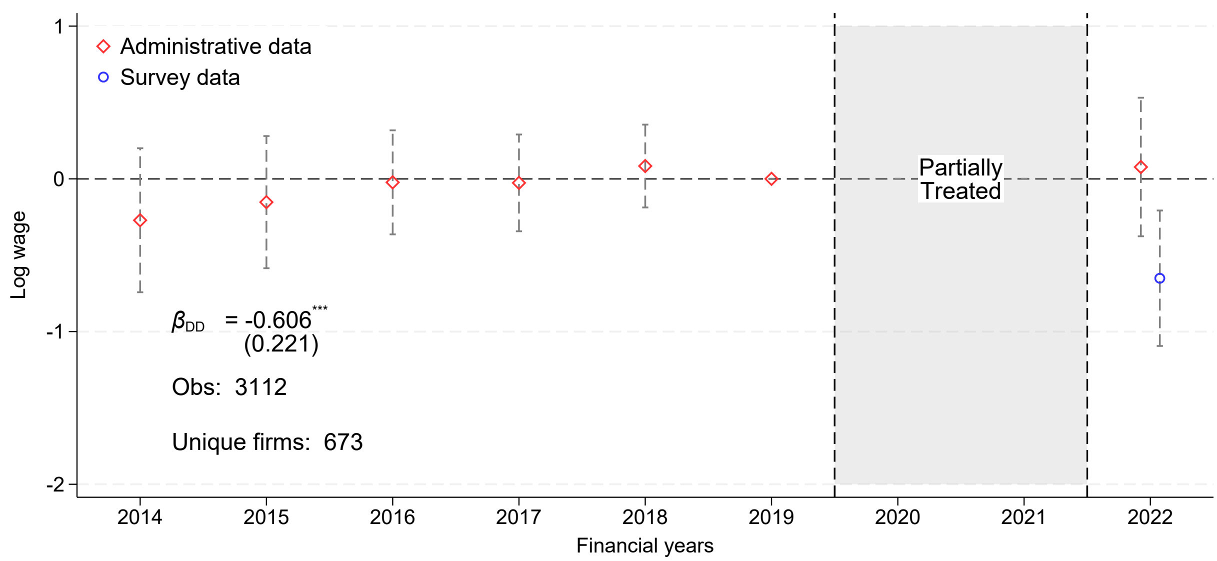
Notes: This figures presents the results estimated from equation 7 with the administrative data and a combination of the administrative and survey data for the logarithmic transformation of sales, separately by firms' perception of audit rates. The blue circles indicate the coefficients estimated from survey data for the last year (2022), and the red squares represent coefficients estimated from the administrative tax data. We restrict to firms that 1) reported sales in our survey; 2) reported that their last active accounting year was before 2022, so we can infer that their sales in 2022 were zero; or 3) we know closed down, and hence we can also infer that their sales in 2022 were zero. We also restrict to firms that answered the following question: "Finally, imagine firms similar to yours in size and sector. In your opinion, what percentage of these firms do you think will be audited in the upcoming financial year?" In Panel (a), we remove firms above the discontinuity that reported a below-median percentage, and in Panel (b), we remove firms that reported an above-median percentage. The shaded area indicates the two years of data that we remove when running equation 6. β_{DD} shows the DD coefficient from equation 6. "Unique firms" is the number of unique firms used in the regression. Standard errors are robust to heteroskedasticity and clustered at the firm level.

Figure 14
Effect of comprehensive audits on outcomes used in the aggregation exercise

(a) Probability that the firm reports sales



(b) Logarithmic transformation of wage bill



Notes: These figures present the results estimated from equation 7 with the administrative data and a combination of the administrative and survey data for the wage bill. The blue circles indicate the coefficients estimated from survey data for the last year (2022), and the red squares represent coefficients estimated from the administrative tax data. We restrict to firms that 1) reported sales in our survey; 2) reported that their last active accounting year was before 2022, so we can infer that their sales in 2022 were zero; or 3) we know closed down, and hence we can also infer that their sales in 2022 were zero. Panel (a) is the probability of reporting sales, which is our best proxy of exit. Panel (b) is the logarithmic transformation of the wage bill. The shaded area indicates the two years of data that we remove when running equation 6. β_{DD} shows the DD coefficient from equation 6. "Unique firms" is the number of unique firms used in the regression. Standard errors are robust to heteroskedasticity and clustered at the firm level.

Tables

Table 1
Summary statistics for different samples

	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	0.013	4.77	0.277
Share with positive VAT due	0.44	0.45	0.021	0.44	0.783
IHS VAT Due	1.18	1.03	0.250	0.60	0.078
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
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	22.13	13.30	0.000	26.90	0.361
Sectors					
Agriculture	0.05	0.04	0.000	0.03	0.015
Construction	0.14	0.13	0.002	0.13	0.228
Manufacturing	0.05	0.06	0.000	0.11	0.000
Missing sector	0.01	0.01	0.002	0.01	0.579
Retail trade	0.10	0.11	0.512	0.10	0.602
Services	0.52	0.52	0.801	0.46	0.001
Wholesale trade	0.13	0.14	0.022	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. Columns (2) and (4) present averages for the full comprehensive audit sample and the comprehensive audit sample firms that tend to fall within the optimal bandwidth. To derive 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 audit sample is significantly different from the average for the universe of firms. Column (4) repeats the same exercise but compares the average for firms in the optimal bandwidth to the average for firms the full comprehensive audit 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, with two exceptions: “Prob. caught evading” is an indicator for whether any assessment that increased the tax liability of the firm was filed during the financial year. “Number of returns amended” is the total number of VAT amendments filed during the financial year. We count missing values as zeroes. Unless otherwise indicated, variables are conditional on the firm’s filing at least one VAT return. For the CIT data, variables show what is declared in the CIT return for financial year 2019/20, with 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 the firm’s filing a CIT return. For the PAYE data, “Number of employees” is the average number of employees reported conditional on the firm’s filing at least one PAYE return. Sectors are aggregate versions of the sector classification used by the URA. Source: Authors’ calculations from URA tax data.

Table 2
Effect of comprehensive audits on CIT liability one year post-audit

	Probability of	Amount of CIT liability		
	positive CIT	Winsorized	Scaled	IHS
	(1)	(2)	(3)	(4)
RD coefficient	-0.184^{***}	-3.714^{**}	-0.015^{**}	-1.864^{***}
	(0.067)	(1.761)	(0.007)	(0.661)
2SLS coefficient	-0.269^{***}	-5.463^{**}	-0.022^{**}	-2.729^{***}
	(0.100)	(2.670)	(0.010)	(0.997)
N	870	816	497	861
Mean in control	0.33	3.98	0.02	3.01
Bandwidth	667	613	658	658

Notes: This table presents reduced-form and 2SLS estimates for the amount of CIT liability in the year after the audit. In column (1), we present an indicator for the probability of filing a CIT return with a positive CIT liability. In column (2), we present the amount of CIT liability submitted in thousand USD winsorized at the top and bottom 1%. In column (3), we present the amount of CIT liability submitted scaled by baseline sales. In column (4), we present the inverse hyperbolic sine transformation of the CIT liability. We estimate the reduced-form (RF) coefficient by running equation 2. We estimate the 2SLS coefficient by running equations 1 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 3
Effect of comprehensive audits on total VAT liability in the first 5 months post-audit

	Probability of	Amount of VAT liability		
	positive VAT	Winsorized	Scaled	IHS
	(1)	(2)	(3)	(4)
RD coefficient	−0.041	−21.327**	−0.091	−1.371
	(0.062)	(10.159)	(0.073)	(1.526)
2SLS coefficient	−0.060	−30.833**	−0.139	−1.991
	(0.089)	(14.464)	(0.109)	(2.171)
N	895	975	831	931
Mean in control	0.31	-7.29	-0.13	-0.07
Bandwidth	692	772	894	728

Notes: This table presents reduced-form and 2SLS estimates for the amount of VAT liability collected in the first five months of the financial year after the audit. In column (1), we present an indicator for the probability of the firm filing a VAT return with a positive VAT liability. In column (2), we present the total amount of VAT liability submitted across the 5 months in thousand USD, winsorized at the top and bottom 1%. In column (3), we present the amount of VAT liability submitted scaled by baseline sales. In column (4), we present the inverse hyperbolic sine transformation of the VAT liability. We estimate the reduced-form (RF) coefficient by running equation 2. We estimate the 2SLS coefficient by running equations 1 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
Summary statistics on tracking and survey completion rates

	N	Percent
All firms	858	
Closed shop	82	10%
Cannot be found (vanished)	76	9%
Exist	683	80%
Completed survey	301	44%
Refused participation	221	32%
Other	161	24%
NGO	17	2%

Notes: This table presents summary statistics for the tracking exercise and the share of surveys done among the firms found to exist. We define a firm as "Closed shop" if we could confirm with neighbors of the former premises or former employees/owners that the firm closed. "Vanished" is defined as us being unable to reach the firm despite attempting at least three phone calls for each number and a visit to each location on file. "Other" is firms that were found to exist and never officially gave us a refusal but kept delaying or 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 RD for the probability that the firm files a CIT return during the two years after the audit.

Table 5

Effect of comprehensive audits on sales from administrative tax and survey data

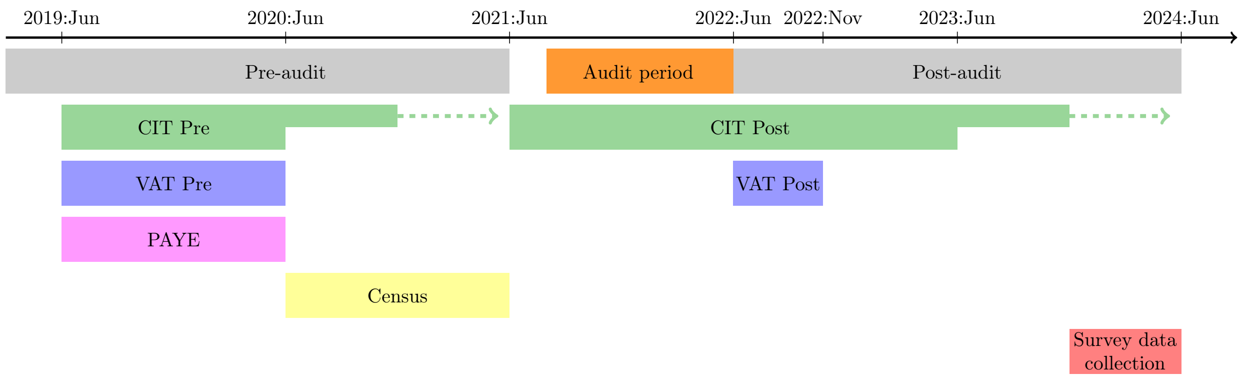
	Panel A: Admin					*Panel B: Admin & survey*				
	Probability		Sales amount			Probability		Sales amount		
	Sales information	Positive sales	Log	Winsorized	Winsorized imputing zeros	Sales information	Positive sales	Log	Winsorized	Winsorized imputing zeros
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DD coefficient	-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 exclusively in the administrative data (Panel A) and in 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 the years 2013–2019 for the pre-period and 2022 for the post-period. We restrict to firms that either 1) reported sales in our survey; 2) reported that their last active accounting year was before 2022, so we can infer that their sales in 2022 were zero; or 3) we know closed down, and hence we can also infer that their sales in 2022 were zero. In columns (1) and (6), the outcome is an indicator variable for whether any sales information was reported (i.e., it is not missing). In columns (2) and (7), the outcome is an indicator variable for whether the firm reported positive sales, conditional on its submitting any sales information. In columns (3) and (8), the outcome variable is the logarithmic transformation of the amount of sales reported. In columns (4) and (9), we show the amount of sales reported, winsorized at the top and bottom 1 percentiles. Columns (5) and (10) use the same outcome variable but impute zeroes for firms not reporting sales. Recall that we include only firms for which we can confidently infer that sales were zero in 2022. Standard errors are robust to heteroskedasticity and clustered at the firm level. Event-study specifications are shown in Figure 11.

A Robustness

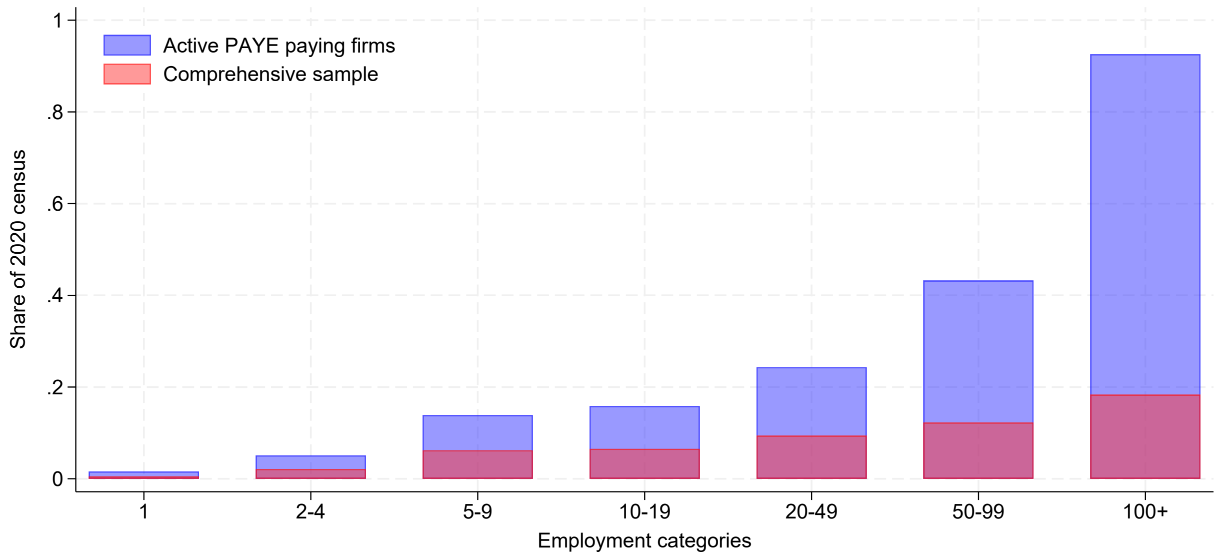
A.1 Figures

Figure A1
Time period covered by each data source



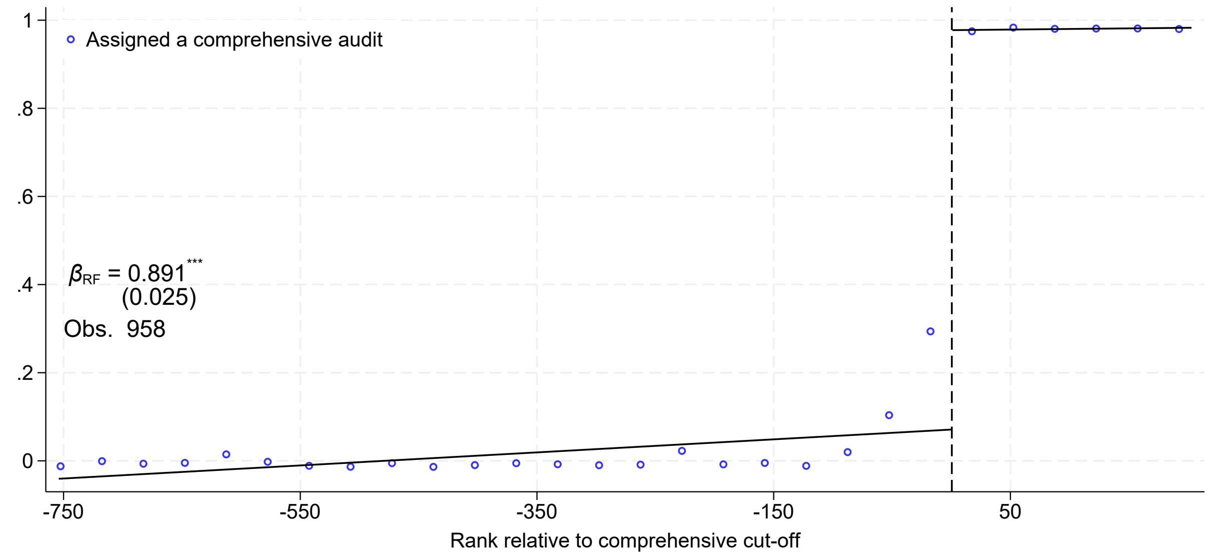
Notes: This figure displays the time period covered for each of the data sources. The CIT data are filed annually and are due 6 months after the close of a firm's accounting year; 79% end their accounting year in June and 6% in December; the year end dates for the remainder are scattered across the calendar year. The dashed green line illustrates the time period when the last set of firms file 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. The survey data were collected between January and July 2024.

Figure A2
Share of tax filing firms in each employment category



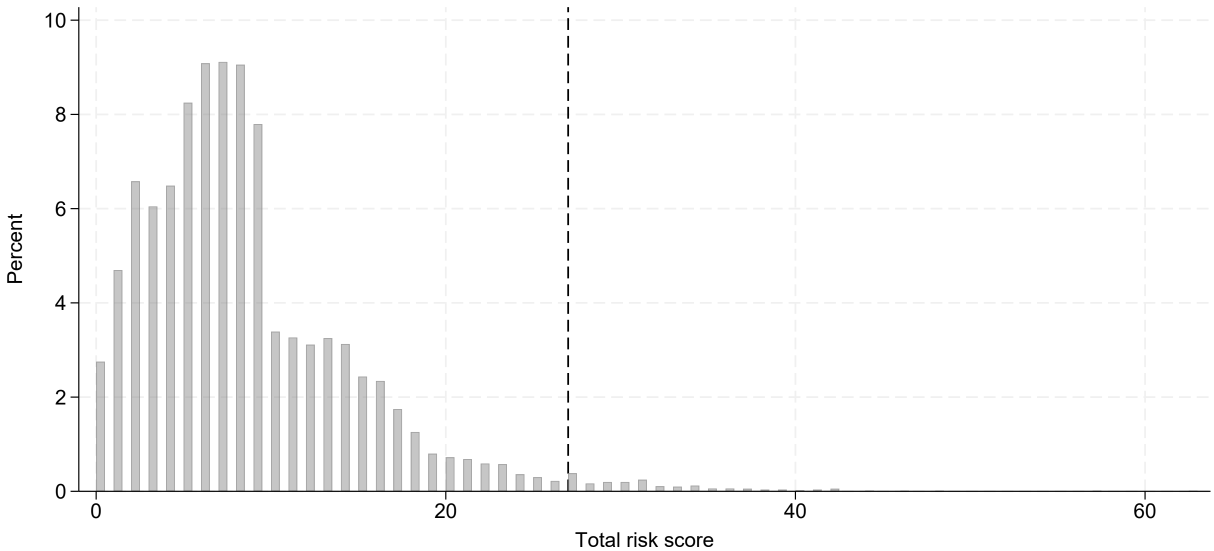
Notes: This figure presents the share of firms filing taxes and the share of firms in the comprehensive audit sample in each employment category. The employment categories are defined by the census, and we construct equivalent categories using the information on number of employees submitted in the PAYE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

Figure A3
Effect of a firm's relative rank on the probability of its being assigned a comprehensive audit



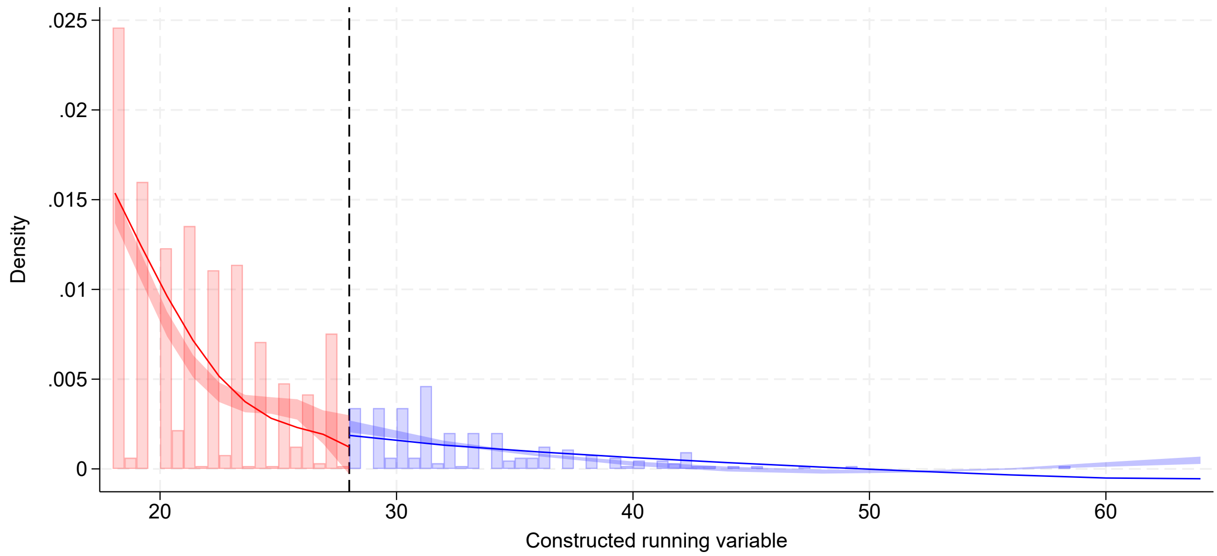
Notes: This figure present results on the probability that a comprehensive audit was assigned. The outcome variable shown in the figure is residualized. The plots are manually constructed; each point on the graph contains approximately 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated with the `rdrobust` package from Calonico et al. (2014). β_{RF} shows the reduced-form effects—from equation 2. Heteroskedasticity-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
Density distribution of total risk score



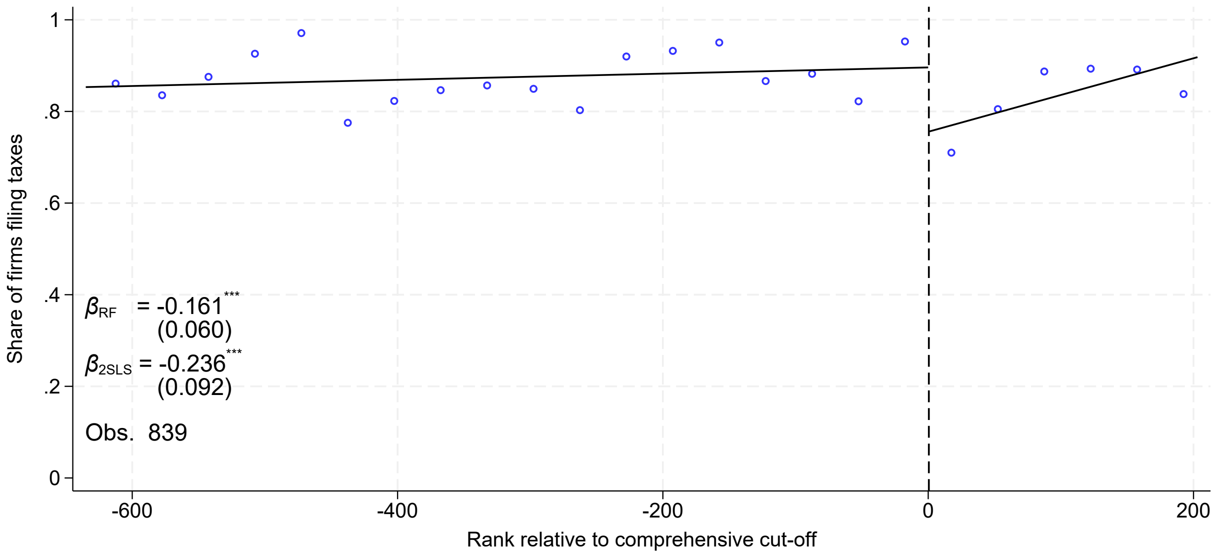
Notes: This figure presents the percent of firms at each discrete risk score in the comprehensive audit sample. Source: URA administrative tax data. Authors' calculations.

Figure A5
Density plot for alternative running variable



Notes: This figure presents the density plot for the alternative running variable discussed in section 5. The figure is produced with the `rddensity` package developed by Cattaneo et al. (2018); Cattaneo et al. (2020). Because the risk score distribution is left skewed, as shown in Figure A4, we restrict the `rddensity` command to run within the optimal bandwidth for one of our outcomes of interests: filing of a CIT return in the two years following the audit. Source: URA administrative tax data. Authors' calculations.

Figure A6
Effect of comprehensive audit on the probability of a firm's filing a CIT or VAT return one year post-audit

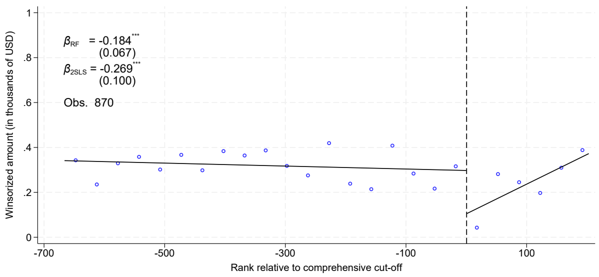


Notes: This figure present results on the probability that the firm files a CIT or VAT return in the year after the audit. The outcome variable shown in the figure is residualized. The plots are manually constructed; each point on the graph contains approximately 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated with the `rdrobust` package from Calonico et al. (2014). β_{RF} shows the reduced-form effects – from equation 2 – and β_{2SLS} the 2SLS effect – from equations 1 and 3. Heteroskedasticity-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.

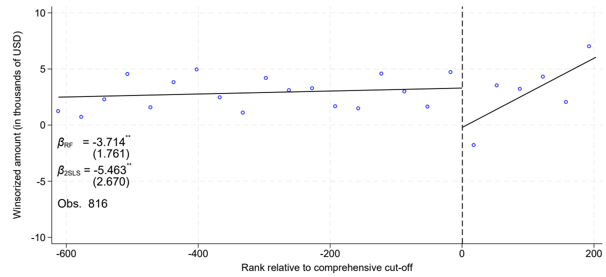
Figure A7

Effect of comprehensive audits on CIT liability one year post-audit

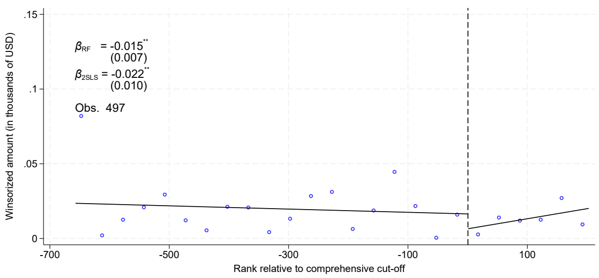
(a) Probability that the firm reports positive CIT liability



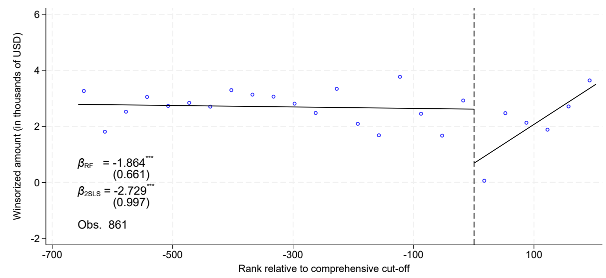
(b) Winsorized amount of CIT liability



(c) Scaled CIT liability



(d) IHS amount of CIT liability

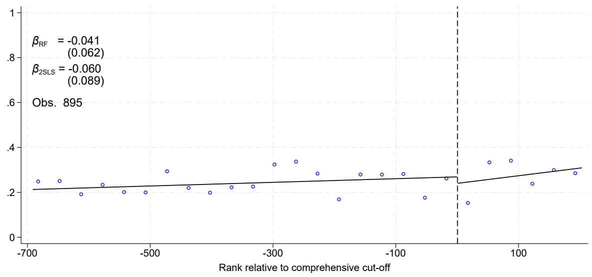


Notes: This figure presents the RD graphs for each of the outcomes shown in Table 2. In Panel (a), we present an indicator for the probability that the firm files a CIT return with a positive CIT liability. In Panel (b), we present the amount of CIT liability submitted in thousand USD winsorized at the top and bottom 1%. In Panel (c), we present the amount of CIT liability submitted scaled by baseline sales. In Panel (d), we present the inverse hyperbolic sine transformation of the CIT liability. The outcome variables shown in the figures are residualized. The plots are manually constructed; each point on the graph contains approximately 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 with the `rdrobust` package from Calonico et al. (2014). β_{RF} shows the reduced-form effects—from equation 2—and β_{2SLS} the 2SLS effect—from equations 1 and 3. Heteroskedasticity- 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.

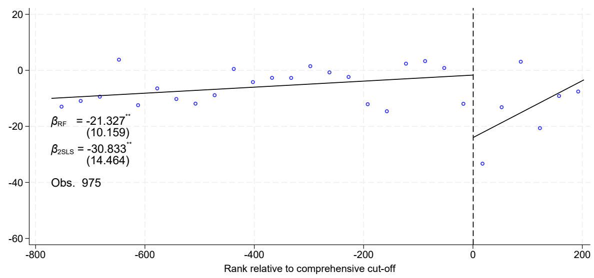
Figure A8

Effect of comprehensive audits on VAT liability in first five months post-audit

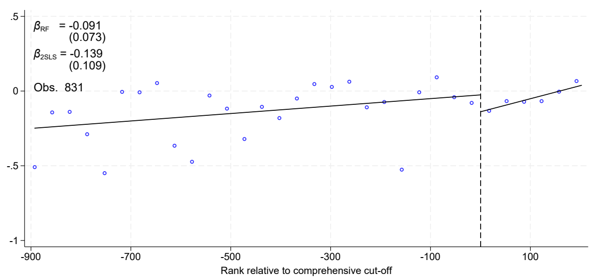
(a) Probability that the firm reports positive VAT liability



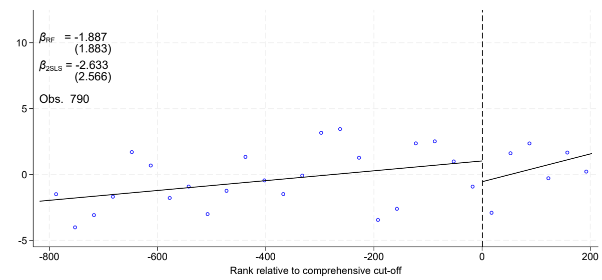
(b) Winsorized amount of VAT liability



(c) Scaled VAT liability



(d) IHS amount of VAT liability

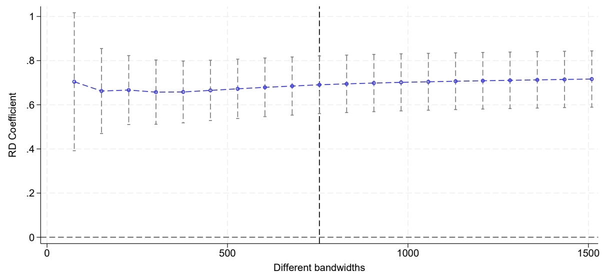


Notes: This figure presents the RD graphs for each of the outcomes shown in Table 3. In Panel (a), we present an indicator for the probability that the firm files a VAT return with a positive VAT liability. In Panel (b), we present the total amount of VAT liability submitted in thousand USD winsorized at the top and bottom 1%. In Panel (c), we present the amount of VAT liability submitted scaled by baseline sales. In Panel (d), we present the inverse hyperbolic sine transformation of the VAT liability. The outcome variables shown in the figures are residualized. The plots are manually constructed; each point on the graph contains approximately 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 with the `rdrobust` package from Calonico et al. (2014). β_{RF} shows the reduced-form effects—from equation 2—and β_{2SLS} the 2SLS effect—from equations 1 and 3. Heteroskedasticity- 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.

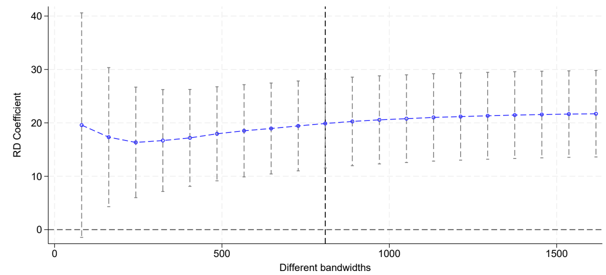
Figure A9

Effect of comprehensive audits on key outcomes under alternative bandwidths

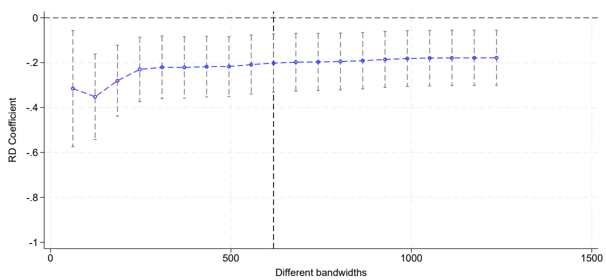
(a) Likelihood that the firm receives a comprehensive audit



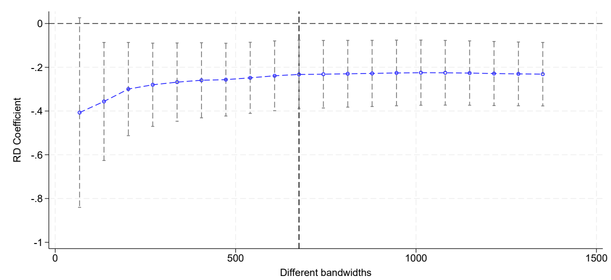
(b) Increase in tax liability from tax corrections



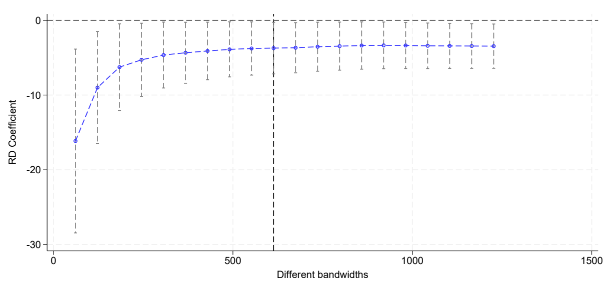
(c) Likelihood that the firm files a CIT return one year post-audit



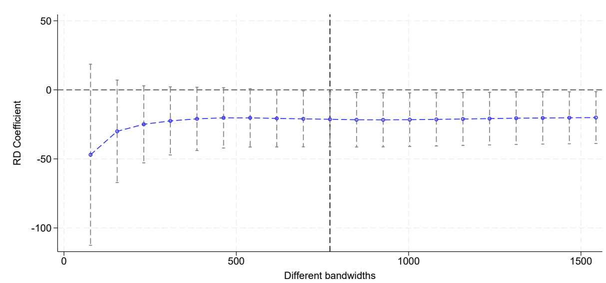
(d) Likelihood that the firm files a CIT return two years post-audit



(e) CIT liability one year post-audit



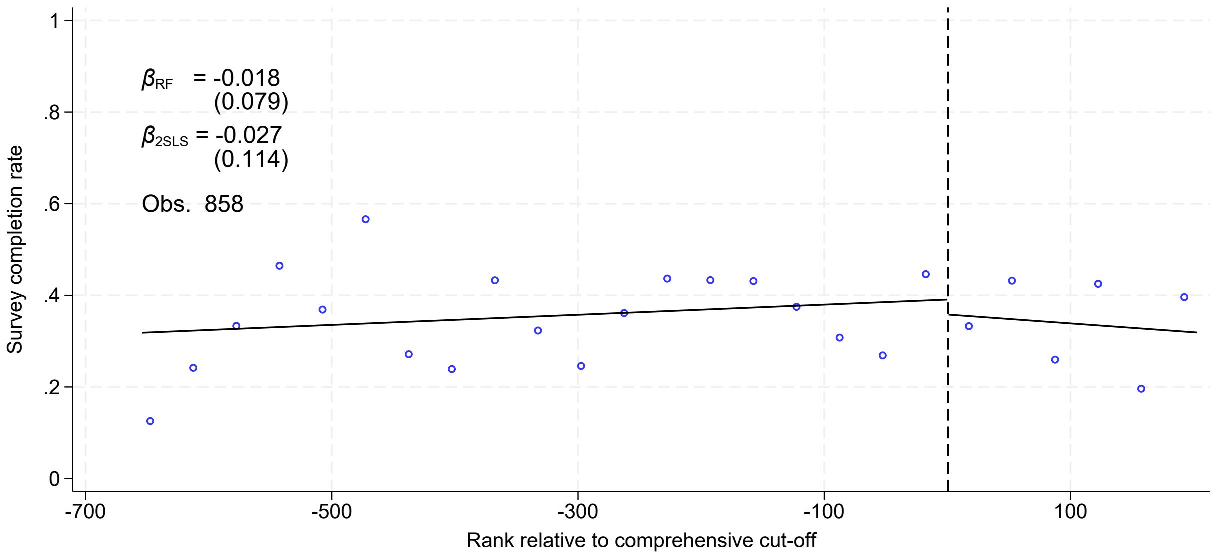
(f) Total VAT liability during first 5 months post-audit



Notes: This figure presents the results from equation 2 for some of our key outcomes of interest under different bandwidths. In Panel (a), the outcome is the probability that the firm receives a comprehensive audit; in Panel (b), the outcome is the increase in tax liability from tax corrections in thousand USD, winsorized at the top and bottom 1%; in Panel (c), the outcome is the probability that the firm files a CIT return in the year after the audit; in Panel (d), the outcome is the probability that the firm files a CIT return two years after the audit; in Panel (e), the outcome is the amount of CIT liability in thousand USD, winsorized at the top and bottom 1%; and in Panel (f), the outcome is the total amount of VAT liability submitted during the first five months of the year after the audit, winsorized at the top and bottom 1%. The vertical dashed line represents the optimal bandwidth selected according to Calonico et al. (2014). The figure shows 95% confidence intervals.

Figure A10

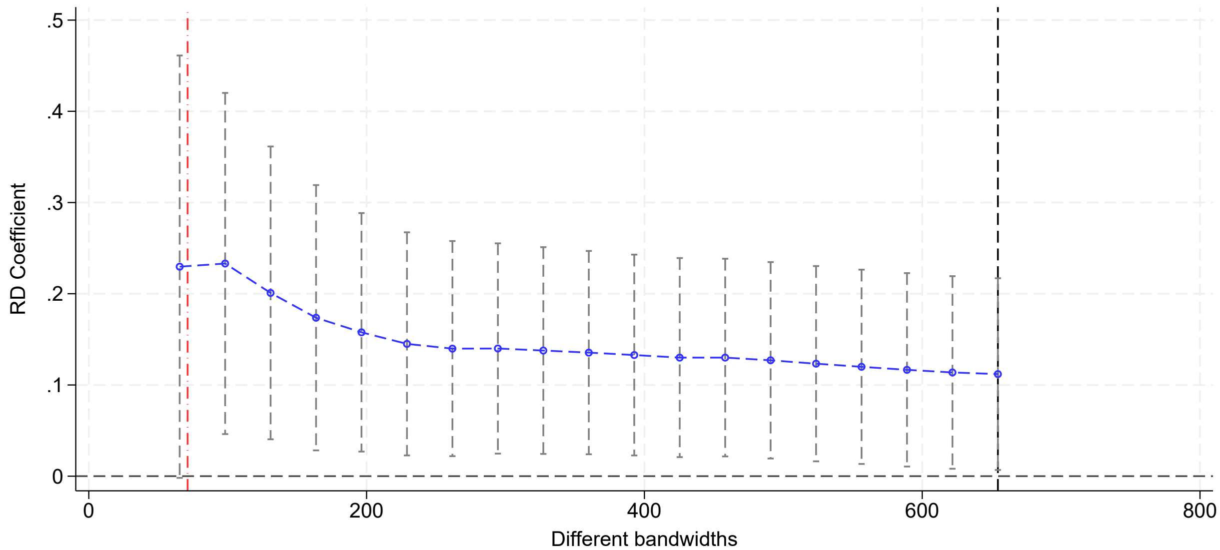
Effect of comprehensive audits on probability that the firm completes the survey



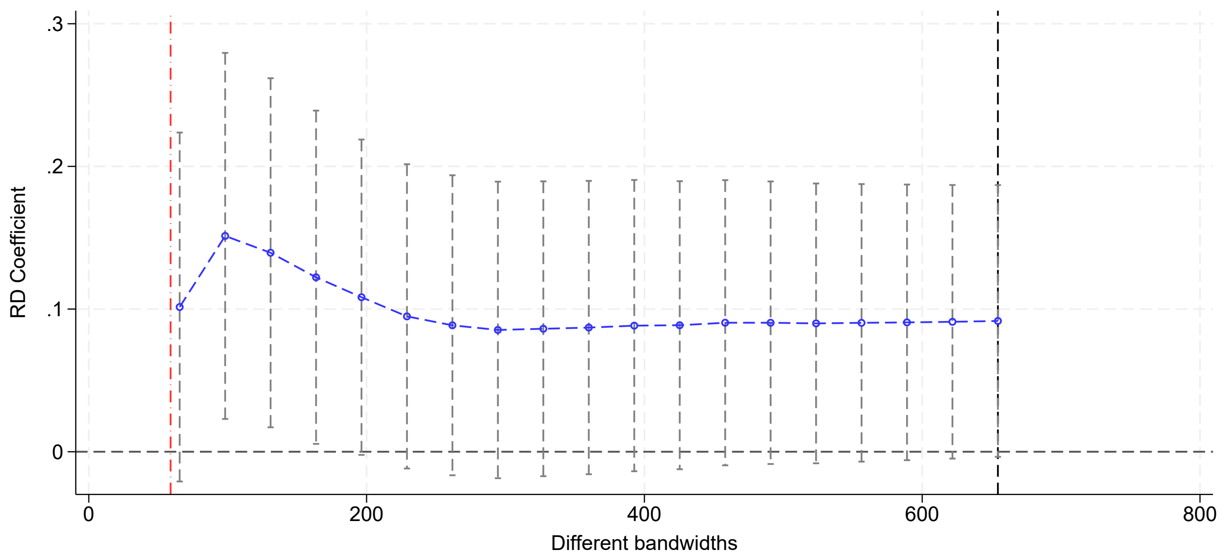
Notes: This figure presents results on the probability that the firm completes a survey. The outcome variable shown in the figure is residualized. The plots are manually constructed; each point on the graph contains approximately 35 firms. The black lines show the linear best fit for the residualized data. The black dashed line indicates the threshold. Bandwidths are the optimal bandwidths calculated with the `rdrobust` package from Calonico et al. (2014). β_{RF} shows the reduced-form effects – from equation 2 – and β_{2SLS} the 2SLS effect – from equations 1 and 3. Heteroskedasticity-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 A11
Effect of comprehensive audits on the probability that the firm closed or is informal under alternative bandwidths

(a) Probability that the firm closed



(b) Probability that the firm is informal



Notes: This figure presents the results from equation 2 for whether a firm closed (Panel (a)) or operates informally (Panel (b)) with different bandwidths. The vertical dashed black line represents the optimal bandwidth selected according to Calonico et al. (2014) for the outcome variable of whether the firm filed a CIT return in the two years after the audit. This is because we tracked down all the firms within the bandwidth for that variable (our closest proxy of firm exit in the administrative tax data). The dot-dashed red line is the optimal bandwidth selected according to Calonico et al. (2014) for the outcome of interest when the package is restricted to the sample we tracked. The figure shows 95% confidence intervals.

A.2 Tables

Table A1
Changes in number of firms with each sample restriction

	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
Restrict to firms assigned to comprehensive and desk audits	13,014	231		12,783	

Notes: This table presents the number of firms that remain after each sample restriction is imposed. In column (1), we present the total number of firms in the sample, whereas in columns (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 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
Baseline covariate balance test around comprehensive audit selection threshold

	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	0.45	0.39	776
Prob of positive VAT Payable	0.54	0.00	0.94	820
IHS VAT Payable	4.64	-0.79	0.59	809
Prob of positive VAT Due	0.43	0.02	0.74	827
IHS VAT Due	0.56	1.03	0.56	815
Number of suppliers	2.49	-1.11	0.20	713
Number of buyers	1.48	0.11	0.69	833
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	24.46	-7.27	0.65	671
Agriculture	0.03	-0.01	0.73	954
Construction	0.13	-0.01	0.82	938
Manufacturing	0.09	0.05	0.42	963
Missing sector	0.01	0.03	0.36	831
Retail trade	0.10	-0.03	0.49	900
Services	0.48	-0.03	0.75	901
Wholesale trade	0.16	0.01	0.86	913
				999
P-value of joint f-test (min. bandwidth)			0.15	324
P-value of joint f-test (med. bandwidth)			0.44	360
P-value of joint f-test (max. bandwidth)			0.53	396
P-value of joint f-test (min. bandwidth) imputing 0			0.55	813
P-value of joint f-test (med. bandwidth) imputing 0			0.60	936
P-value of joint f-test (max. bandwidth) imputing 0			0.64	1023

Notes: In this table, we present results from balance tests for a wide range of covariates measured at baseline. We run equation 2 with the optimal bandwidth selected with the `rdrobust` package from Calonico et al. (2014). In the last 6 rows, we test whether the baseline covariates are jointly different from 0. We run the specification with 3 different bandwidths, the minimum, median and maximum bandwidths selected across all the baseline covariates, and separately for when we impute zeroes 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, with two exceptions. Variables are conditional on the firm's filing at least one VAT return. For the CIT data, variables show what is declared in the CIT return for financial year 2019/20. Variables are conditional on the firm's filing a CIT return. For the PAYE data, "Number of employees" is the average number of employees reported conditional on the firm's filing at least one PAYE return. Sectors are aggregate versions of the sector classification used by the URA. Source: VAT, CIT, and PAYE return data from the URA. Authors' calculations.

Table A3
Effect of comprehensive tax audits on CIT liability for year of the audit

	Probability of	Amount of CIT liability		
	positive CIT	Winsorized	Scaled	IHS
	(1)	(2)	(3)	(4)
RD coefficient	−0.012	−0.375	−0.007	−0.069
	(0.083)	(2.751)	(0.009)	(0.827)
2SLS coefficient	−0.018	−0.553	−0.010	−0.102
	(0.120)	(3.973)	(0.013)	(1.189)
N	859	794	543	832
Mean in control	0.42	4.97	0.02	3.78
Bandwidth	656	591	735	629

Notes: This table presents reduced-form and 2SLS estimates for the amount of CIT liability in the year of the audit (but filed after). In column (1), we present an indicator for the probability that the firm files a CIT return with a positive CIT liability. In column (2), we present the amount of CIT liability submitted in thousand USD winsorized at the top and bottom 1%. In column (3), we present the amount of CIT liability submitted scaled by baseline sales. In column (4), we present the inverse hyperbolic sine transformation of the CIT liability. We estimate the reduced-form (RF) coefficient by running equation 2. We estimate the 2SLS coefficient by running equations 1 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 A4
Effect of comprehensive audits on different margins of CIT returns

	CIT liability	Probability		CIT liability
	Winsorized	Filing a CIT return	Positive CIT liability conditional on filing	Log
	(1)	(2)	(3)	(4)
RD coefficient	-3.528^{**}	-0.232^{***}	-0.186^{*}	-0.153
	(1.671)	(0.079)	(0.107)	(0.507)
N	939	939	616	302
Mean in control	4.11	0.67	0.50	8.40
Bandwidth	736	736	736	736

Notes: This table presents reduced-form results for different margins of changes in firms' CIT filings. In column (1), the outcome is the amount of CIT liability in thousand USD, winsorized at the top and bottom 1%. In column (2) the outcome is the probability that the firm files a CIT return in the second year post-audit. In column (3), the outcome is the probability that the firm files a CIT return with a positive CIT liability, conditional on its filing a CIT return. In column (4), the outcome is the logarithmic transformation of the CIT liability amount submitted. To make the results comparable, we fix the bandwidth to be the optimal bandwidth for the outcome in column (4). The coefficients are estimated with equation 2. 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

First stage and effect on filing of returns with different sets of controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Probability the firm receives a comprehensive audit						
RD coefficient	0.719^{***}	0.686^{***}	0.691^{***}	0.682^{***}	0.673^{***}	0.667^{***}	0.662^{***}
	(0.067)	(0.067)	(0.067)	(0.069)	(0.068)	(0.070)	(0.066)
N	958	958	958	928	933	933	933
Bandwidth	755	755	755	755	760	760	760
	Panel B: Probability the firm files a CIT return one year post-audit						
RD coefficient	-0.192^{***}	-0.210^{***}	-0.202^{***}	-0.191^{***}	-0.161^{**}	-0.157^{**}	-0.165^{***}
	(0.067)	(0.065)	(0.066)	(0.066)	(0.064)	(0.066)	(0.061)
N	821	821	821	795	831	831	831
Bandwidth	618	618	618	618	656	656	656
	Panel C: Probability the firm files a return two years post-audit						
RD coefficient	-0.207^{**}	-0.225^{***}	-0.233^{***}	-0.233^{***}	-0.219^{***}	-0.279^{***}	-0.227^{***}
	(0.081)	(0.080)	(0.080)	(0.082)	(0.080)	(0.078)	(0.076)
N	879	879	879	850	881	881	881
Bandwidth	676	676	676	676	707	707	707
Station FE	No	Yes	Yes	Yes	Yes	Yes	No
Sample FE	No	No	Yes	Yes	Yes	Yes	No
Removing fictitious firms	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 that the firm receives a comprehensive audit (Panel A), the probability that the firm files a CIT return the year after the audit (Panel B), and the probability that the firm files a CIT return in the second year post-audit (Panel C) with different empirical specifications. 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 remove "fictitious" firms, defined as firms labeled suspicious by the URA; in column (5), we add sector fixed effects; and in column (6), we add risk controls. These are 70 separate controls that control for the risk score and expected potential revenue of each of the underlying 35 risk parameters. Finally, in column (7), we run a LASSO specification to select the controls. In this specification, we do not penalize the station fixed effects because—based on conversations with the URA—they are an important part of selection. We fix the bandwidth to the optimal bandwidth selected when we run the specification in column (3). We estimate the coefficients by running equation 2 and varying $\mathbf{X}_f$. Standard errors are robust to heteroskedasticity.

Table A6
Effect of comprehensive audits on changes in tax liabilities with different sets of controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Increase in tax liability from tax corrections						
RD coefficient	22.854***	20.146***	19.881***	19.589***	19.404***	17.823***	19.210***
	(4.401)	(4.253)	(4.250)	(4.295)	(4.299)	(4.095)	(4.149)
N	1012	1012	1012	982	982	982	982
Bandwidth	809	809	809	809	809	809	809
	Panel B: CIT liability one year post-audit						
RD coefficient	−2.457	−3.278*	−3.714**	−3.668**	−3.747**	−4.259**	−3.510**
	(1.698)	(1.727)	(1.761)	(1.826)	(1.802)	(1.869)	(1.627)
N	816	816	816	790	790	790	790
Bandwidth	613	613	613	613	613	613	613
	Panel C: Total VAT liability during 5 months post-audit						
RD coefficient	−22.685**	−22.090**	−21.327**	−23.184**	−22.437**	−14.125	−22.371**
	(10.015)	(10.156)	(10.159)	(10.586)	(10.324)	(9.534)	(9.457)
N	975	975	975	945	945	945	945
Bandwidth	772	772	772	772	772	772	772
Station FE	No	Yes	Yes	Yes	Yes	Yes	No
Sample FE	No	No	Yes	Yes	Yes	Yes	No
Removing fictitious firms	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 increase in tax liability from tax corrections (Panel A), CIT liability in the year after the audit (Panel B), and VAT liability in the first five months post-audit (Panel C) with different empirical specifications. In column (1), we do not use any controls; in column (2), we include only 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 remove "fictitious" firms, defined as firms labeled suspicious by the URA; in column (5), we add sector fixed effects; and in column (6), we add risk controls. These are 70 separate controls that control for the risk score and expected potential revenue of each of the underlying 35 risk parameters. Finally, in column (7), we run a LASSO specification to select the controls. In this specification, we do not penalize the station fixed effects because—based on conversations with the URA—they are an important part of selection. We fix the bandwidth to the optimal bandwidth selected when we run the specification in column (3). We estimate the coefficients by running equation 2 and varying $\mathbf{X}_f$. Standard errors are robust to heteroskedasticity.

Table A7
Alternative threshold: Effect of comprehensive audits for threshold based on initial audits assignment

	Probability			Winsorized amount		
	Receiving comprehensive audit	Filing CIT return 21/22	Filing CIT return 22/23	Tax correction	CIT liability	VAT liability
RD coefficient	0.611^{***}	-0.168^{***}	-0.190^{**}	16.709^{***}	-2.744	-19.867^{**}
	(0.068)	(0.062)	(0.078)	(4.050)	(1.781)	(9.764)
N	849	850	924	958	895	1022
Mean in control	0.01	0.85	0.67	2.77	4.05	-7.27
Bandwidth	641	642	716	750	687	814

Notes: This table presents results from a robustness check where we use the alternative (assigned) threshold. That is, we run equation 2 using the alternative cut-off. 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

Alternative sample: Effect of comprehensive audits for the sample of all firms

	Probability			Winsorized amount		
	Receiving comprehensive audit	Filing CIT return 21/22	Filing CIT return 22/23	Tax correction	CIT liability	VAT liability
RD coefficient	0.639^{***}	-0.162^{***}	-0.192^{***}	20.633^{***}	-3.798^{**}	-13.522
	(0.065)	(0.057)	(0.072)	(5.168)	(1.771)	(9.387)
N	1421	1220	1277	1419	1288	1341
Mean in control	0.01	0.85	0.67	5.33	4.56	-9.27
Bandwidth	1191	990	1047	1189	1058	1111

Notes: This table presents results estimated in the full sample of all firms. The purpose of this check is to determine whether our results hinge on our comparing firms assigned to desk audits with firms assigned to comprehensive audits. That is, we run equation 2 using firms assigned to any enforcement intervention (comprehensive, issue, desk) or to no enforcement intervention. 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 A9

Dimension reduction exercise: Effect of comprehensive audits estimated with predicted outcomes

	Probability			Winsorized amount		
	Receiving comprehensive audit	Filing CIT return 21/22	Filing CIT return 22/23	Tax correction	CIT liability	VAT liability
RD coefficient	-0.002	-0.028	-0.011	-0.327	-0.432	-0.396
	(0.005)	(0.029)	(0.026)	(0.402)	(0.805)	(1.745)
N	958	821	879	1012	816	975
Mean in control	0.01	0.86	0.67	2.97	3.98	-7.29
Bandwidth	755	618	676	809	613	772

Notes: This table presents results from our dimension reduction exercise. We predict each of our main outcomes using all baseline covariates and run an RD on the predicted outcome. The purpose is to see whether differences in baseline covariates can predict the treatment effects we find. That is, we run equation 2 on the predicted outcome. 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 A10
Baseline characteristics of compliers

	All	Compliers			
		Audited		Not filing CIT	
	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	12.02	12.94	0.07	11.38	0.42
IHS Purchases	11.19	12.98	0.00	10.11	0.40
Prob of positive VAT Payable	0.54	0.53	0.84	0.51	0.80
IHS VAT Payable	4.77	4.10	0.67	4.18	0.86
Prob of positive VAT Due	0.44	0.44	1.00	0.58	0.34
IHS VAT Due	0.47	0.52	0.98	5.47	0.25
VAT Due	-8317.90	-3466.82	0.67	11678.00	0.11
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 of filing CIT	0.77	0.80	0.64	1.03	0.00
IHS Turnover	9.87	10.22	0.75	5.40	0.04
IHS Cost of sales	8.48	10.47	0.07	5.03	0.09
Prob of positive pre-tax profits	0.39	0.25	0.10	0.16	0.09
IHS Pre-tax profits	-0.39	-3.89	0.06	-4.20	0.15
Prob of positive tax liability	0.36	0.33	0.69	0.26	0.51
IHS Tax liability	3.21	2.91	0.74	2.14	0.39
Tax liability	2437.62	3404.21	0.59	1288.92	0.42
IHS Value of plant/machines	3.55	2.15	0.20	0.66	0.01
Sales/Cost of sales	2.06	1.46	0.01	1.95	0.89
Agriculture	0.03	0.01	0.47	0.02	0.80
Construction	0.14	0.14	0.98	0.04	0.23
Manufacturing	0.07	0.24	0.03	-0.07	0.00
Missing sector	0.01	0.03	0.45	0.11	0.40
Retail trade	0.10	0.12	0.63	0.13	0.75
Services	0.46	0.35	0.20	0.55	0.63
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 induced into comprehensive audits (the comprehensive audit complier average). In column (4), we present baseline characteristics for firms induced into not filing CIT returns after the audit (the nonfiling compliers mean). Columns (5) and (6) present the p-value for the 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 zeroes only for "VAT due" and "Tax liability." For the VAT data, variables represent monthly averages over the 12 months from July 2019 to June 2020. For the CIT data, variables show what is declared in the CIT return for financial year 2019/20. Sectors are aggregate versions of the sector classification used by the URA. Source: VAT, CIT, PAYE return from the URA. Authors' calculations.

Table A11
Filing rates across tracking categories

	File vat or cit	File cit 22/23
Closed	0.72	0.38
Vanished	0.64	0.38
Exist	0.95	0.72
NGO	0.94	0.72

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

Table A12
Effect of comprehensive audits on firm closure or informality with different sets of controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Probability that the firm closed						
RD coefficient	0.108**	0.115**	0.112**	0.097*	0.098*	0.083*	0.112**
	(0.052)	(0.053)	(0.054)	(0.052)	(0.051)	(0.050)	(0.052)
N	858	858	858	830	858	858	858
Bandwidth	655	655	655	655	655	655	655
	Panel B: Probability that the firm is informal						
RD coefficient	0.083*	0.096**	0.092*	0.098**	0.086*	0.095*	0.092*
	(0.049)	(0.048)	(0.049)	(0.049)	(0.049)	(0.052)	(0.048)
N	858	858	858	830	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 that the firm closed (Panel A) and the probability that the firm is informal (Panel B) with different empirical specifications. In column (1), we do not use any controls; in column (2), we include only 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 remove "fictitious" firms, defined as firms labeled suspicious by the URA; in column (5), we add sector fixed effects; and in column (6), we add risk controls. These are 70 separate controls that control for the risk score and expected potential revenue of each of the underlying 35 risk parameters. Finally, in column (7), we run a LASSO specification to select the controls. In this specification, we do not penalize the station fixed effects because—based on conversations with the URA—they are an important part of selection. We fix the bandwidth to the optimal bandwidth selected when we run the specification in column (3). We estimate the coefficients by running equation 2 and varying $\mathbf{X}_f$. Standard errors are robust to heteroskedasticity.

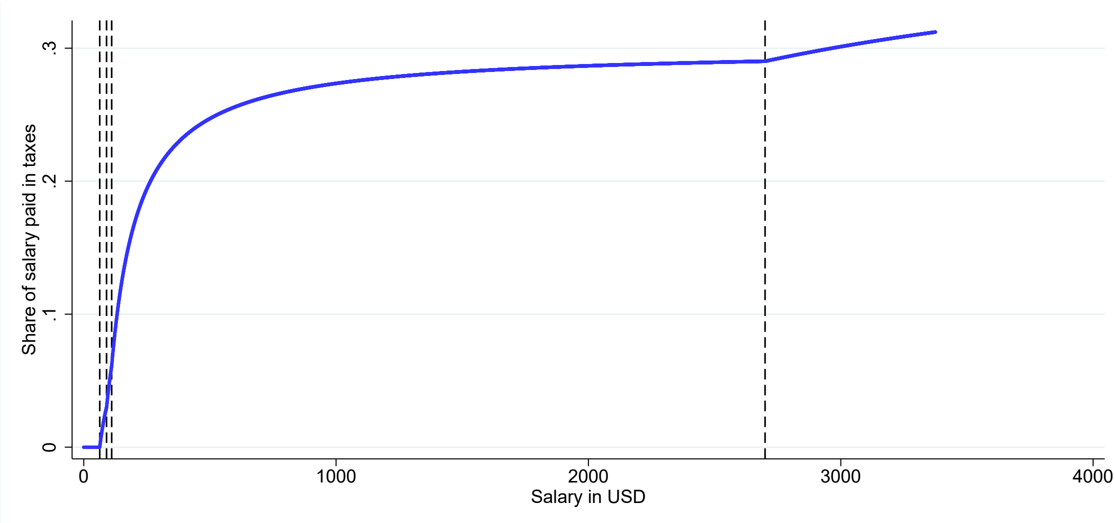
Table A13
Effect of comprehensive audits on sales and different measures of costs

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

Notes: This table presents reduced-form and 2SLS estimates for sales, costs and number of employees reported by firms. In column (1), we present an indicator for whether the firm reported having positive sales. In columns (2) and (3), we present the amount of sales, winsorized at the top and bottom 1st percentile and the logarithmic transformation. In columns (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 the optimal bandwidth for the probability that the firm files a CIT tax return in the two years after the audit. We estimate the reduced-form (RF) coefficient by running equation 2. We estimate the 2SLS coefficient by running equations 1 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: <u>TAX COMPLIANCE ADVISORY FOR THE PERIOD ENDED DECEMBER 31, 2020</u>.

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.**

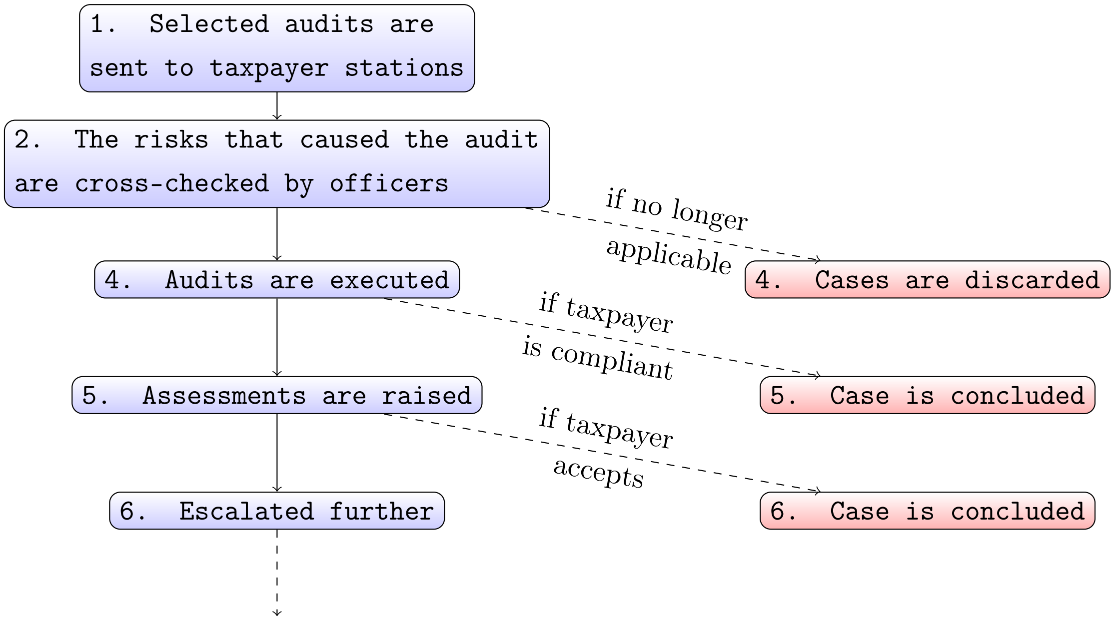
You are therefore advised to amend the company's tax returns to rectify the above issues and pay resultant tax liability.

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
Graphic illustration of URA audit process



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

C Math

In this section, we derive equations 11 and 12 used in the aggregation exercise in section 8.

Assume output y_i is produced with only labor with a constant-returns-to-scale production function $f(l_i)$

$$y_i = A_i f(l_i)$$

where A_i is the productivity of firm i. Profits of the firm are then given by

$$\pi_i = p_i A_i f(l_i) - w l_i$$

For a small change (prices and wages do not adjust), the total derivative of profits with respect to labor is given by

$$d\pi_i = p_i \left[dA_i f(l_i) + A_i \frac{\partial f}{\partial l_i} dl_i \right] - w dl_i$$

Changes in labor are unlikely to affect the productivity of the firm; hence, we set $dA_i = 0$ and obtain

$$\begin{aligned} d\pi_i &= \left[p_i A_i \frac{\partial f}{\partial l_i} - w \right] dl_i \\ &= \left[p_i A_i \frac{\partial f}{\partial l_i} l_i - w l_i \right] \frac{dl_i}{l_i} \\ &= [s_i - c_i] \frac{dl_i}{l_i} \end{aligned}$$

where the second equality follows from multiplying and dividing by l_i and the third equality follows by definition under constant returns to scale. Finally, we aggregate across all firms and consider discrete changes Δ instead infinitesimal changes d. Thus, the first-order approximation is given by

$$\Delta Y_g = \sum_{i \in g} \Delta \pi_i \approx \sum_{i \in g} [s_i - c_i] \times \%\Delta l_i \tag{13}$$

This holds for a constant-returns-to-scale production function with one input. It also holds separately for each input, under multiple inputs. For it to hold for changes in labor if the firm uses multiple inputs, we need to impose the additional assumption that inputs are perfect complements.

The derivation of equation 12 follows from substituting in $\%\Delta l_i = (1 - p_e)\%\Delta l_i + p_e \%\Delta l_i$

$$\Delta Y_g = \sum_{i \in g} \Delta \pi_i \approx \sum_{i \in g} [s_i - c_i] \times [(1 - p_e)\%\Delta l_i - p_e]$$

$$\approx \sum_{i \in g} [s_i - c_i] \times (1 - p_e) \% \Delta l_i - \sum_{i \in g} [s_i - c_i] \, p_e$$

$$\approx \sum_{i \in g | exist} [s_i - c_i] \times \Delta \log l_i - \sum_{i \in g} [s_i - c_i] \, p_e$$