UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2021/150

# The effects of a risk-based approach to tax examinations

## Evidence from Tanzania

Amina Ebrahim,[1] Elineema Kisanga,[2] Ezekiel Swema,[5] Vincent Leyaro,[2] Edwin P. Mhede,[5] Ephraim Mdee,[5] Heikki Palviainen,[3] and Jukka Pirttilä[4]

October 2021

United Nations University World Institute for Development Economics Research
wider.unu.edu

**Abstract:** While technical assistance and increased use of ICT in the area of tax administration have been regarded to hold considerable promise for greater revenue collection, the evidence on how these activities work in the real-world circumstances of developing countries is scant. The paper attempts to fill this gap by evaluating an intervention undertaken jointly by the Finnish and Tanzanian revenue administrations. In this pilot programme, a risk-based method for enhancing firm tax examinations in Tanzania was developed. The results, which are based on a difference-in-differences strategy and administrative data from the Tanzanian Revenue Authority, demonstrate that the intervention increased adjusted taxable income by approximately 15 per cent during the first year of the new approach.

**Key words:** risk-based approach, firm tax audits, Tanzania

**JEL classification:** C93, H26

**Acknowledgements:** We are grateful to colleagues at Finnish Tax Administration and Tanzania Revenue Authority for helpful comments.

**Note:** The order of authors' names appears as per the randomization tool provided by the American Economic Association.

[1] UNU-WIDER; [2] University of Dar es Salaam, Tanzania; [3] Tampere University, Tampere, Finland, corresponding author: heikki.palviainen@tuni.fi; [4] University of Helsinki, VATT Institute for Economic Research, and UNU-WIDER, Helsinki, Finland; [5] Tanzanian Revenue Authority, Dar es Salaam, Tanzania

This study has been prepared within the UNU-WIDER project Building up efficient and fair tax systems – lessons based on administrative tax data, which is part of the Domestic Revenue Mobilization programme. The programme is financed through specific contributions by the Norwegian Agency for Development Cooperation (Norad).



Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-090-0

https://doi.org/10.35188/UNU-WIDER/2021/090-0

Typescript prepared by Siméon Rapin.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

# 1 Introduction

Expanding the tax base and enhancing tax compliance that minimizes tax evasion and increases tax revenue collection has been the core focus and function of many tax administrations, especially in the developing countries. There is also a strong commitment among the donor countries to assist developing economies to reach this objective. Tax and supporting the tax systems are key in the 17th Sustainable Development Goal,[1] and already earlier the Addis Ababa Action Agenda[2] of 2015—via in particular the Addis Tax Initiative[3]—brought development collaboration and technical assistance to centre stage as a means to enhance domestic revenue mobilization.

Tax authorities use various enforcement strategies, including random audits and examinations, but risk-based approaches are common, and with the onset of machine learning, hold a great promise for improving the detection probabilities of non-compliant taxpayers (see Khwaja et al. 2011 for a review of tax practices). However, very little systematic evidence on the impact of these risk-based strategies is available (at least outside of revenue organizations). In our understanding, there are no earlier studies investigating how risk-based compliance interventions work in low-income developing countries like Tanzania.

This paper attempts to fill this gap by examining the revenue impacts of a risk-based tax examination pilot implemented in Tanzania. The pilot was planned jointly by the Tanzanian Revenue Authority (TRA) and the Finnish Tax Administration (VERO), and the Finnish involvement was financed from the development assistance budget of the government of Finland. The intervention is a new method for flagging taxpayers for tax examination on the basis of a data-driven risk assessment. It was designed to improve the existing practice of choosing taxpayers to be examined, which relied more on staff discretion, and was implemented in personal income and corporate income taxation. We use a difference-in-differences design, comparing revenue developments of firms handled by Dar es Salaam tax offices (where the pilot took place) with revenues from corresponding firms in the other five major tax offices in Tanzania, namely Arusha, Dodoma, Mbeya, Morogoro, and Mwanza. The data stem from the administrative information system of the TRA and they are complemented with supplementary material used in the pilot.

Our paper contributes to the literature in four main ways. First, the results of the paper improve our understanding about how risk-based tax examinations work in a low-income county setting, with a lower-than-usual tax administrative capacity. Second, this is to our knowledge one of the first studies to evaluate whether and how technical assistance, financed by donor countries, help receiving countries enhance their tax administrations and revenue collection. Third, while it is a conventional wisdom that tax avoidance and evasion is rampant in developing countries, very little is known about how administrative interventions help lower the extent of these activities. This paper serves as one step towards enhancing our understanding of this matter. Fourth, and more broadly, the paper contributes to the expanding literature which provides causal evidence about enforcement using administrative data from developing countries.

Our context is Tanzania, a low-income sub-Saharan African (SSA) developing country which, as many other SSA countries, strives to raise the tax to GDP ratios. Besides substantial tax reforms

[1] See: https://sdgs.un.org/goals/goal17.

[2] See: https://sustainabledevelopment.un.org/index.php?page=view&type=400&nr=2051&menu=35.

[3] See: https://www.addistaxinitiative.net/.

since independence, studies have shown that tax evasion and avoidance remain rampant in Tanzania and are present in almost all sectors of the economy (Wadhawan 1998). This includes the informal sector where tax evasion was in the range of 35–55 per cent of the total revenue collected by the year 2010 (ESRF 2010). Tax evasion is also rife in the trade sector through under-invoicing custom duties at the port of entry. The reasons for tax evasion in Tanzania include high scheduled tax rates, low salaries, and lack of incentives, for example those offered to customs staff (Epaphra 2015).

Over the past two years, the government has introduced tax-awareness campaigns in secondary schools and uses news media such as television and radio to provide tax education in order to minimize oppressive revenue collection methods as a way to address tax evasion and avoidance. Traditional deterrence measures, such as tax audit threats and penalties, remain the main measures to enhance tax compliance in Tanzania as in many other developing countries. Despite the deterrence methods employed, tax revenue collection performance in Tanzania is low. Tax revenue is approximately 12 per cent of GDP in the last decade, which is below the sub-Saharan Africa average of 18 per cent of GDP.

The paper proceeds as follows. A review of the earlier related literature is offered in Section 2. The context of the pilot is explained in Section 3, whereas Section 4 contains a description of the data and the empirical methods that we use. Section 5 presents the results and Section 6 concludes.

## 2 Literature review

### 2.1 Literature on tax evasion[4]

Earlier research on the effects of enforcement policies on tax compliance can be traced back to the work of Allingham and Sandmo (1972), where the income earner's tax compliance decision is regarded to be the result of expected utility maximization. Since the income of the taxpayer is not observed, and the taxpayer is considered a risk-averse agent facing a flat income tax rate, it is assumed that the taxpayer decides what fraction of true income to report facing an exogenous audit (and therefore detection) probability. This basic model creates a trade-off between saving taxes through under-reporting and penalty in case of detection. Tax evasion is expected to decrease relative to the detection probability and the penalty for evasion (Azulai et al. 2014).

A number of empirical studies, both experimental (in the lab and in the field) and non-experimental, have shown that deterrence measures have a positive effect on tax compliance. One such study is a large field experiment by Kleven et al. (2011) done in Denmark. The authors conducted a large-scale random audits programme together with the Danish authority. They found that the evasion rate for third-party reported income is miniscule in comparison to self-reported incomes. The reason is that taxpayers have very limited evasion opportunities when the tax authority receive, often automatically, information on true income earned from third parties. Self-reported income also reacts significantly to having experienced audits in the past.

In addition to information flows and audit risks, it is important to augment the standard model of extrinsic motivation with intrinsic motivations such as morals, social norms, and psychology in order to accurately capture evasion behaviour. While tax moral undoubtedly plays an important

[4] See Slemrod (2019) for an extensive review.

role in curbing evasion (Luttmer and Singhal 2014), our study—and hence this brief survey—focuses on monitoring and inspection conducted by revenue authorities instead.

For developing countries, these remarks imply that since the extent of third-party reporting is not as significant, traditional deterrence tools, e.g. tax examinations and audits, may play a larger role.

### 2.2 Work on the role of information provision

Besides the work on personal income taxation by Kleven et al. (2011), Pomeranz (2015) conducted related work but in the value-added tax (VAT) system. Using a field experiment on firms in Chile, companies were randomly selected to receive audit threat letters; where the focus is to leverage on the potential information trail on the conflicting incentives of sellers against buyers in a VAT chain, i.e. the fact that both buyers and sellers report the same transaction and their incentives to misreport are the opposite. This property breaks down in the retail sector, since the final consumers cannot get a refund for the VAT.

She finds that threats of audit significantly improve compliance among retailers, not covered before by third-party reporting.[5] What these findings posit is the importance of expansion of information trails as top priority, even though a couple of challenges deserve attention.

First, a side-effect of increasing deterrence in some sectors may lead to production shifting towards sectors with less compliance interventions (Carillo et al. 2017). Income shifting may also take place between earnings and self-employment income (Kleven and Waseem 2013). These considerations are arguably more pronounced in a low-income country like Tanzania with more limited administrative capacity.

Second, the impacts may vary by firm size. Kleven et al. (2016) argued that employers and employees may find it easier to collude in smaller firms, enabling evasion of pay-as-you-earn (PAYE) taxes. In larger firms, in contrast, the likelihood of whistleblowing increases. Hence if firm size increases with development, there is a link from economic structures to increased probability of truthful reporting. Evidence supporting this claim has been found by Kumler et al. (2013).

### 2.3 Research on audits and other enforcement mechanisms

Like Kleven et al. (2011), Advani et al. (2017) examine the effects of random audits of individuals in a high-income country (the UK). They find that the audits have long-lasting effects: those audited also report increased incomes in the subsequent years after audits. Risk-based audits have been studied by Løyland et al. (2019), who use a discontinuity approach related to a risk score in Norway. They also find dynamic impacts from audits. The benefit of their approach is an arguably greater external validity, since they evaluate an intervention that tax authorities routinely use in developed countries.

Basri et al. (2020) focus on an additional angle of enforcement, namely the role of dedicated tax offices to certain segments of taxpayers, in their case the impacts of being handled by a medium-sized taxpayer office (MTO) in Indonesia. The introduction of the MTO involved greater scrutiny over the taxpayers handled by the new bureau and resulted in a clear increase in taxable income and taxes paid.

---

[5] This feature of the VAT is also highlighted by the conceptual arguments of Kopczuk and Slemrod (2006) and Keen (2008).

In our understanding, there is no prior hard evidence on the impacts of risk-based tax examinations on revenues in developing countries and very little work in general on how technical assistance works in the area of tax administration, despite the policy focus on support for tax revenue mobilization.

In summary, compliance can be improved by traditional mechanisms (examinations and audit) or by extensive use of information from third parties, and the combination of the two. But the key is implementation, and it is this area where the conditions in low-income countries are the most challenging, also because of the economic structure (the presence of a large number of small informal firms).

## 3 Description of the pilot programme

The new method of examination was set up for several reasons including increasing revenue collected, improving the skills of the tax officers, treating taxpayers equally and focusing examination efforts on risky taxpayers and reducing the time spent on examining taxpayers with limited risks. The new method involved the use of an Excel spreadsheet that was developed between TRA and VERO. The spreadsheet raises a flag when information provided by the taxpayer on certain aspects of the tax calculation are found to be outside of reasonable bounds which are pre-determined by TRA and coded into the spreadsheet. These pertain to, for instance, the developments in firms' sales and expenses.[6] Tax officers involved in the process received training and were provided with instructions for using the Excel spreadsheet. In practice, the tax return is filed by the taxpayer and collected by the tax officer. Then, the data handler enters the return information into the TRA information system (called ITAX) and Excel spreadsheet. Further examination is then conducted in the case where the spreadsheet indicates this is advised. It is possible that adjustments are then made which are then recorded again in the ITAX system. It was intended that the pilot would be implemented for both corporate income tax (CIT) payers and value added tax (VAT) payers, however, in practice the pilot was only implemented for corporate income taxpayers. The pilot was implemented in tax regions such as Ilala, Temeke, Kinondoni, and Kariakoo, which are administratively located in Dar es Salaam region. The Large Taxpayer Department (LTD) was also included in the pilot and considered as tax region which is also located in Dar es Salaam region. However, because of more complex selection criteria, in practice the LTD was nearly fully excluded from the pilot programme. We exploit the fact that the pilot was only implemented in the Dar region in our empirical strategy.

## 4 Empirical strategy and data

The empirical strategy follows the standard difference-in-differences (DD) approach:

$$Y_{it} = \alpha + \beta Dar_i + \gamma POST_t + \delta Dar_i \times POST_t + \theta X_{i,t} + \varepsilon_{i,t} \quad (1)$$

where $Y_{it}$ is the outcome variable and $Dar_i$ is the treatment variable (equal to 1 if the tax office is in Dar es Salaam region). Other large regions' tax offices are used in the control group. The adjusted taxable income is used as our main outcome variable, but we also study the number of examinations considered and extra income raised by examinations. This variable contains taxable

[6] Due to security reasons, the exact characteristics of the procedure cannot be disclosed.

income after examinations, audits, and inspections. Because we do not observe the examined companies, this is an intent-to-treat (ITT) estimation. $POST_t$ controls for time-variant conditions for everyone (equals to 1 for the post-intervention period). We observe the tax adjustment decision, therefore the $POST_t$ variable receives the value of 1 after 1 July 2019. The adjustment decisions are recorded until June 2020. Our policy parameter of interest is $\delta$. This parameter shows how much the pilot increased adjusted taxable income in tax offices in the Dar es Salaam region in comparison to selected tax offices in large regions outside of the Dar es Salaam region. We add control variables, $X_{i,t}$, to the basic model to control for any industry or other differences between firms that are unrelated to the treatment. Because companies are not typically examined, audited, or inspected on consecutive years, we use a cross-sectional difference in differences.

We use administrative firm level panel data collected by the Tanzania Revenue Authority (TRA) in this study. The data come from income tax returns for the period 1 July 2015 to 30 June 2020. Each income year contains more than 25, 000 observations. Variables include a set of income and taxable income variables, taxpayer and industry types, and regional location of the firm. Because the pilot was introduced 1 July 2019, we divide our sample into financial years. That is, our first examined year covers 1 July 2016–30 June 2017, and the last examined year covers 1 July 2019–30 June 2020 (the treatment period). From the data set it is observed that number of firms with adjusted taxable income increases from one financial year to another in both treatment and control groups (Table 1).

Table 1: The number of firms with adjustments to taxable income

| Financial year | Control group | Treatment group | Total |
|---|---|---|---|
| 2016/17 | 1,270 | 2,808 | 4,078 |
| 2017/18 | 1,753 | 4,392 | 6,145 |
| 2018/19 | 1,984 | 5,754 | 7,738 |
| 2019/20 | 2,434 | 8,069 | 10,503 |

Source: authors' estimates based on ITAX data.

Table 2 summarizes the number of adjustments that have taken place during the pilot period for the treatment group; it is observed that Ilala has relatively more adjusted income cases (approximately 70 per cent out of the total cases). Such significant higher cases could be explained either by the existence of small- and medium-size business units with relatively low knowledge of recording and reporting financial transaction or by the existence of higher response of business units to adhere to TRA interventions and guidelines as well as the efficacy of the newly introduced system under the pilot exercise. However, it is possible that other factors explain such a significant incidence relative to other tax offices.

Table 2: Number of adjusted taxable income by tax region

| Tax region | Number | Percentage(%) |
|---|---|---|
| Illala | 1,302 | 70.84 |
| Kariakoo | 35 | 1.90 |
| Kinondoni | 379 | 20.62 |
| LTD | 9 | 0.49 |
| Temeke | 113 | 6.15 |
| Total | 1,838 | 100.00 |

Source: authors' elaboration based on Excel spreadsheet data.

Table 3 explores the mean differences between the adjusted income due to tax examinations and the self-declared income. The mean difference is observed to be positive overall, implying that the examinations led to increased taxable incomes and hence increased tax obligations.

Some of the observed reasons for adjustments include unverified expenses and purchases, overstated purchases and understated/unreported sales, disallowed depreciation, unverified loans, understated closing stock, unverified trade payables, unverified shareholder's advances, and unreported non-current assets.

Table 3: Additional average income by tax office (in million)[7]

| Tax region | Mean |
|---|---|
| Illala | -8,750 |
| Kinondoni | 13,654 |
| Temeke | 5,368 |
| Total | 8,065 |

Source: authors' elaboration based on Excel spreadsheet data.[8]

We compare the characteristics of firms in the controlled and uncontrolled firms before and after pilot intervention in Table 4. Firms in the treatment area are larger and more profitable. The largest number of firms are from the service sector and from wholesale and retail.

Table 4: Descriptive statistics before and after the treatment period

| | Before | | After | |
|---|---|---|---|---|
| | Control | Treatment | Control | Treatment |
| *All firms* | | | | |
| Taxable income | | | | |
| Adjusted taxable income, TZS millions | 112.484 | 506.175 | 164.286 | 818.018 |
| | (369.45) | (6,440.84) | (1,239.85) | (8,997.48) |
| Log adjusted income | 3.353 | 3.444 | 3.050 | 3.302 |
| | (1.66) | (2.02) | (1.88) | (2.20) |
| Company type | | | | |
| Limited company | 0.595 | 0.738 | 0.687 | 0.726 |
| | (0.49) | (0.44) | (0.46) | (0.45) |
| Sole proprietor | 0.389 | 0.252 | 0.292 | 0.263 |
| | (0.49) | (0.43) | (0.45) | (0.44) |
| Tax type | | | | |
| Corporate tax | 0.609 | 0.746 | 0.707 | 0.736 |
| | (0.49) | (0.44) | (0.46) | (0.44) |
| Personal income tax | 0.391 | 0.254 | 0.293 | 0.264 |
| | (0.49) | (0.44) | (0.46) | (0.44) |
| Industry type | | | | |
| Agriculture | 0.014 | 0.004 | 0.014 | 0.008 |
| | (0.12) | (0.06) | (0.12) | (0.09) |
| Mining | 0.026 | 0.011 | 0.019 | 0.011 |

[7] At the time of writing (April 2021), US$1 = TZS2,300 (Tanzanian schillings).

[8] The pilot Excel files did not have entries for self-reported income for LTD and Kariakoo (although the ITAX data do have), and that is why those tax offices do not appear in the table.

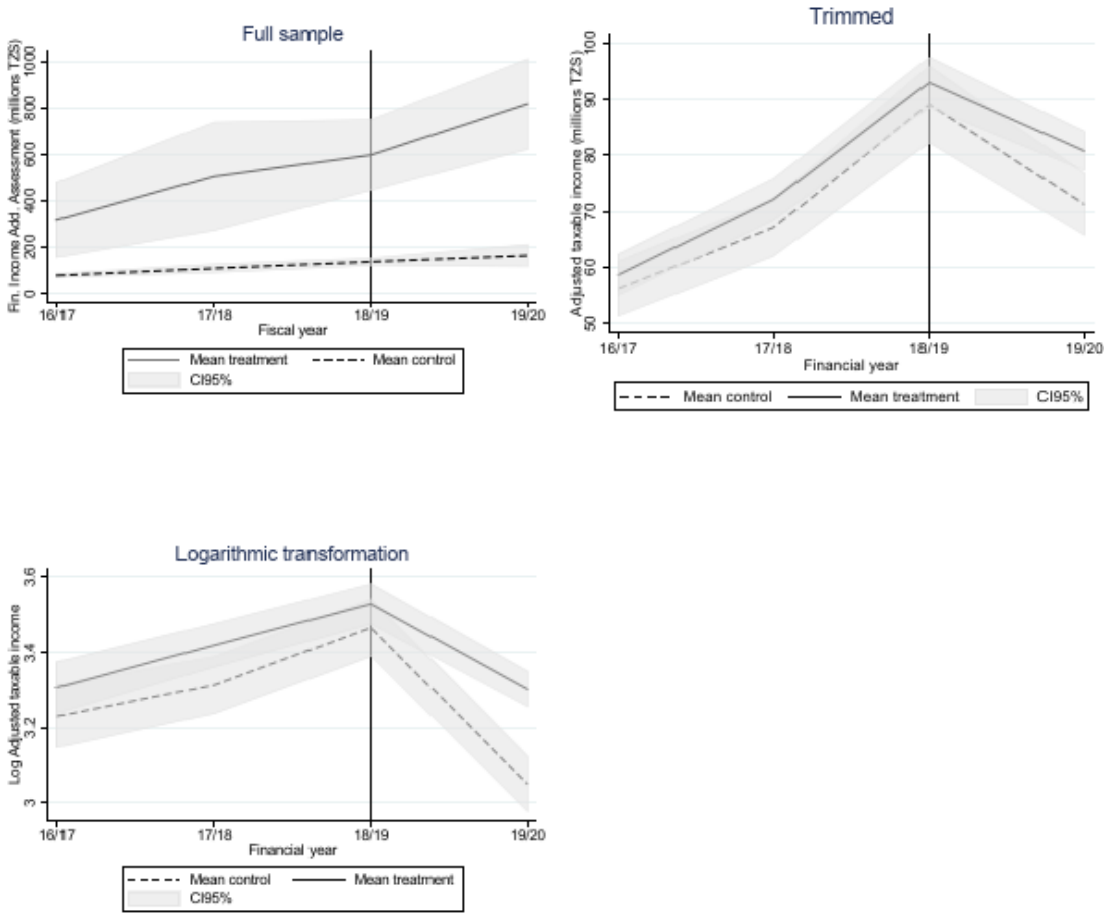
| | | | | |
|---|---|---|---|---|
| | (0.16) | (0.10) | (0.14) | (0.10) |
| Construction | 0.047 | 0.066 | 0.049 | 0.084 |
| | (0.21) | (0.25) | (0.22) | (0.28) |
| Manufacturing | 0.066 | 0.063 | 0.048 | 0.056 |
| | (0.25) | (0.24) | (0.21) | (0.23) |
| Wholesale and retail | 0.375 | 0.307 | 0.276 | 0.309 |
| | (0.48) | (0.46) | (0.45) | (0.46) |
| Information and communication | 0.093 | 0.102 | 0.084 | 0.097 |
| | (0.29) | (0.30) | (0.28) | (0.30) |
| Transportation and storage | 0.012 | 0.033 | 0.022 | 0.037 |
| | (0.11) | (0.18) | (0.15) | (0.19) |
| Finance, insurance, and real estate | 0.060 | 0.110 | 0.048 | 0.091 |
| | (0.24) | (0.31) | (0.21) | (0.29) |
| Services | 0.257 | 0.254 | 0.378 | 0.254 |
| | (0.44) | (0.44) | (0.49) | (0.44) |
| Observations | 5007 | 12954 | 2434 | 8069 |

Note: before reflects the time period 1 June 2016 to 30 June 2019 and after the period 1 July 2019 to 20 June 2020. Standard errors are shown in parentheses.

Source: authors' estimates based on ITAX data.

A DD approach assumes that differences between the treatment and control groups are stable over time. We investigate this in Figure 1 showing the common trends. The figure shows the average adjusted taxable income in millions of Tanzanian schillings (TZS) for the treatment and control groups over the period of study. There appear to be some challenges in the assumption of parallel trends in the pre-reform period, as it is evident the treatment group has companies with a very large adjusted taxable income. We therefore suspect the problem is related to outliers of large firms in the full sample. To address this issue, we trim our sample. We drop observations above 95$^{th}$ percentile from the outcome variable for each financial year and show the effect on the parallel trends in the graph on the right in Figure 1. As a sensitivity analysis, we also estimate the results using other percentiles. The parallel trends for other samples are shown in Figure A1. Table 4 shows that we lose 1,318 observations from the treatment group and 153 observations from the control group. After the trimming, the parallel trend assumption appears to hold well. We also estimate results for the full sample using a logarithmic transformation for ease of interpretation. (below in Figure 1).

Figure 1 Mean adjusted taxable income by group




Source: authors' estimates based on ITAX data.

The trimmed graph in Figure 1 shows that there is a noticeable change in differences between the treatment and control groups after the pilot was introduced. This is what we test in the next section.

## 5 Results and discussion

We now proceed to test the DD regression for our outcomes of interest. We consider the following outcomes: adjusted taxable income for the trimmed sample, logged taxable income, extra income defined as the difference between the self-declared income and the final income, and log of the extra income.[9] Table 5 shows the DD results. We estimate the results with and without control variables. As control variables we use tax type (PIT, corporate tax), taxpayer category (limited company, sole proprietor, etc.), and industries. The first model (column 1) estimates the level (absolute) effect in adjusted taxable income using the trimmed sample described in the previous section to ensure parallel pre-reform trends and exclude any large outliers. Our estimations are average effects, excluding high earning companies that can affect the results significantly. We

[9] Having discovered that the parallel trends assumption does not hold for the extra income variable, we use a trimmed sample (with the observations above the 95th percentile by year dropped) when the outcome is either extra income or the log of it.

report results for the basic DD, DD with controls, and DD with controls using clustered standard errors with clustering at the firm level. The results are estimated with a full sample for the logged adjusted taxable income (2) and below 95$^{th}$ percentile for other outcome variables. We show the results below 95$^{th}$ percentile because the point estimates differ statistically significantly from zero for the extra income variables on the pre-treatment period if samples above 95$^{th}$ percentile are used. Tables A1 and A2 show the results for other percentiles (below 92.5$^{th}$ and 97.5$^{th}$ percentiles).

The first model (1) shows the absolute effects on adjusted taxable income. The effect 7.3 million (TSZ) is not statistically significant at a significance level of 0.05. Table A1 shows that the effect is 13.5 million (TSZ) if a sub-sample with adjusted taxable income below 97.5$^{th}$ is used. This effect is statistically significant. Although the parallel trend assumption appears to hold less, these results imply that the main effects come from higher earning companies.

The second model shows the results from the log-transformed model. Adding the control variables decreases the treatment effect from 0.162 to 0.127 if the log transformation is used. The coefficient 0.127 means that the pilot increased adjusted taxable by approximately 13.5 per cent ($\exp(0.127)\approx 1.135$). The effect is not statistically significant at a 5 per cent significance level, but clustering the standard errors at the firm level increases the standard errors. The effect is significant at the 5 per cent significance level if robust standard errors are used.

The third model shows results from the adjusted indicator, defined to be 1 if the taxable income was adjusted and 0 if not. According to columns (3)–(5), the effect arises from having greater incomes from firms with income adjustments rather than a larger number of adjustments in the treatment area after the pilot. This is in line with the rationale of the pilot, which was to increase the efficiency of examinations rather than being able to conduct more examinations. The percentage increase for the extra income variable is somewhat larger than for taxable income itself. The difference in the magnitude can be explained by the fact that the extra income has a lower base to which the increase is compared.

Table 5: Estimated treatment effects

| | (1) Adj taxable inc | (2) Log taxable inc | (3) Adjusted | (4) Extra inc | (5) Log extra inc |
|---|---|---|---|---|---|
| Basic DD | 4.078 | 0.162*** | 0.0181* | 3.299 | 0.157*** |
| | (4.050) | (0.0540) | (0.0105) | (2.183) | (0.0492) |
| + controls | 7.314* | 0.127** | 0.000162 | 5.092** | 0.144*** |
| | (4.026) | (0.0543) | (0.0104) | (2.164) | (0.0493) |
| + clustered std. errors | 7.314 | 0.127* | 0.000162 | 5.092** | 0.144** |
| | (4.813) | (0.0709) | (0.0129) | (2.527) | (0.0608) |
| Observations | 27,043 | 28,464 | 28,464 | 27,043 | 26,682 |

Note: controls include tax type (PIT, corporate tax), taxpayer category (limited company, sole proprietor, partnerships, parastatal, unclassified class), and industry type. Signification levels: *** 1%, ** 5% and * 10%. Table shows the results using a full sample for the logged adjusted taxable income (2) and below 95th percentile for the other outcome variables.

Source: authors' estimates based on ITAX data.

As a robustness check, we use simple placebo tests. That is, we treat the periods July 2017–June 2018 and July 2018–June 2019 as placebo treatment periods. We expect the results to be zero for these placebo treatment periods. Our point estimates do not differ statistically significantly from zero on the placebo years for four out of the five outcome variables (see Table 6).[10]

Table 6: Estimated treatment effects

| | (1) Adj taxable inc | (2) Log taxable inc | (3) Adjusted | (4) Extra inc | (5) Log extra inc |
|---|---|---|---|---|---|
| treat_1718 | 3.294 | 0.00870 | 0.0147 | 1.069 | 0.0350 |
| | (4.364) | (0.0733) | (0.0138) | (2.757) | (0.0673) |
| treat_1819 | 2.361 | -0.0255 | -0.0123 | -0.201 | -0.0225 |
| | (5.663) | (0.0769) | (0.0139) | (3.091) | (0.0700) |
| treat_1920 | 10.98** | 0.129 | -0.000859 | 6.156** | 0.155** |
| | (5.261) | (0.0833) | (0.0153) | (2.952) | (0.0731) |
| Observations | 27,043 | 28,464 | 28,464 | 27,043 | 26, 682 |
| R-squared | 0.037 | 0.028 | 0.036 | 0.027 | 0.014 |

Note: controls include tax type (PIT, corporate tax), taxpayer category (limited company, sole proprietor, partnerships, parastatal, unclassified class), and industry type. Signification levels: *** 1%, ** 5% and * 10%. Table shows the results using a full sample for the logged adjusted taxable income (2) and below 95th percentile for the other outcome variables.

Source: authors' estimates based on ITAX data.

Table 7 reports results for three subgroups: CIT payers, PIT payers, and those firms not handled by the Large Taxpayer Department (LTD). The majority of taxpayers are corporations, and hence the positive impact arises from the corporate income side. When the firms handled by the LTD are dropped, the treatment impact declines, and its statistical significance drops.

Table 7: Estimated treatment effects

| | (1) CIT | (2) PIT | (3) No LTD |
|---|---|---|---|
| Basic DD | 0.174** | -0.0333 | 0.0928 |
| | (0.0732) | (0.0713) | (0.0685) |
| Observations | 20,372 | 8,092 | 27,412 |
| R-squared | 0.025 | 0.027 | 0.014 |

Note: dependent variable: Log adjusted income with a full sample. All models with controls, which include taxpayer category (limited company, sole proprietor, partnerships, parastatal, unclassified class) and industry type. Robust standard errors in parentheses. Significance levels: *** 1%, ** 5% and * 10%.

Source: authors' estimates based on ITAX data.

Table 8, in turn, reports the DD results by industries. A large majority of the companies operate in the service sector, and that is where the effect stems from. The treatment impact is not very precisely estimated for the other sectors due to fewer number of firms in these groups. When all the other sectors are combined and services dropped, the treatment impact remains non-significant (not reported).

[10] While log taxable income is not significant in this specification, its magnitude is at the same ballpark, and it is close to being significant.

Table 8: Estimated treatment effects

| | (1) Agric. | (2) Mining | (3) Manuf. | (4) Services | (5) Other |
|---|---|---|---|---|---|
| Basic DD | -0.910 | -0.471 | -0.156 | 0.269*** | -0.00383 |
| | (0.629) | (0.525) | (0.152) | (0.0610) | (0.195) |
| Observations | 220 | 406 | 4,489 | 21,864 | 1,485 |
| R-squared | 0.018 | 0.045 | 0.011 | 0.016 | 0.024 |

Note: dependent variable: Log adjusted income with a full sample . All models with controls, which include tax type (PIT, corporate tax) and taxpayer category (limited company, sole proprietor, partnerships, parastatal, unclassified class). Robust standard errors in parentheses. Significance levels: *** 1%, ** 5% and * 10%.

Source: authors' estimates based on ITAX data.

Our estimations show that the pilot increased adjusted taxable income by approximately 14 per cent for all firms. Because the parallel trend assumption holds for all firms only if a log transformation is used, we trimmed our sample. The null effect on the pre-pilot years gives us confidence that there is no pre-trend we should be concerned about. Adjusted taxable income contains also results from audits and inspections; the average treatment effects are lower than they would be for the treated firms only. We assume that there were no other programmes or reforms related to tax examinations that could intervene on our results. Our results should be interpreted with caution. The parallel trend assumption holds reasonably well, but a longer treatment period would provide more credible results. It is also possible that the COVID-19 pandemic had a differing effect on the treatment and control groups on the treatment period.

Finally, we discuss some cost-benefit properties of the project. The Finnish development assistance component costed approximately €60,000. One year before the reform, the additional income raised from the treatment groups was €575 million. When this is multiplied by an estimated treatment effect (0.13) and the corporate income tax rate (0.3), the estimated increase in the tax revenue amounts to €18 million. In other words, the cost of the intervention for the donor per additional revenue is tiny. The Tanzanian Revenue Authority costs are not known to the research team, but TRA did not hire additional staff for the work. It is therefore clear that the benefits of the projects in terms of additional tax revenue clearly dominate any additional intervention costs.

## 6 Conclusions

This paper examined the revenue impacts of a new risk-based method of tax examinations (desk audits) in Tanzania. The method was jointly developed with the Finnish tax administration and represents a capacity development programme designed to bring greater efficiency in tax processes. Our paper therefore also serves as an evaluation of a technical assistance project in the tax area. The results, based on a difference-in-differences estimation strategy and administrative data from the Tanzania Revenue Authority information system, indicate that the pilot we study increased revenues by approximately 15 per cent. The impacts were concentrated in the service sector and among corporate income taxpayers.

In a context where there was limited consistency in the selection of taxpayer for examination, the pilot process has provided additional structure. The pilot decentralized processes standardized the examination process and provided detailed instructions on performing examination. This has translated into reduced examination process times using a systematic approach and clearly defined selection criteria. However, there are still possible efficiency gains. Because of more complex selection criteria, the large taxpayers' department (LTD) was nearly fully excluded from the pilot

programme. While there are some indications of positive revenue gains from the pilot study, these results should be treated with caution due to the short post-treatment period.

## References


Advani, A., W. Elming, and J. Shaw (2017). 'The Dynamic Effects of Tax Audits'. IFS Working Paper W17/24. London: Institute for Fiscal Studies.

Azulai, M., O. Bandiera, F. Blum, H. Kleven, E. LaFerrara, G. Padro, and C. Tejada (2014). 'State Effectiveness, Growth and Development'. IGC Evidence Paper. London: International Growth Center. Available at: https://www.theigc.org/wp-content/uploads/2014/09/IGCEvidencePaperState.pdf (accessed 8 February 2021).

Allingham, M.G., and A. Sandmo (1972). 'Income Tax Evasion: A Theoretical Analysis'. *Journal of Public Economics,* 1(3-4): 323–38. https://doi.org/10.1016/0047-2727(72)90010-2

Basri, M, M. Felix, R. Hanna, and B. Olken (2020). 'Tax Administration vs. Tax Rates: Evidence from Corporate Taxation in Indonesia'. Mimeo. Cambridge, MIT: Massachusetts Institute of Technology.

Carrillo, P., D. Pomeranz, and M. Singhal (2017). 'Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement'. *American Economic Journal: Applied Economics,* 9(2):144–64. https://doi.org/10.1257/app.20140495

ESRF (2010). 'Informal Sector Taxation in Tanzania'. TAKNET Policy Brief Series Number 012-2010. Dar es Salaam: Economic and Social Research Foundation. Available at: http://esrf.or.tz/docs/INFORMATIONSECTORTAXATIONINTANZANIA_12.PDF (accessed 8 February 2021).

Epaphra, M. (2015). 'Tax Rates and Tax Evasion: Evidence from Missing Imports in Tanzania'. *International Journal of Economics and Finance*, 7(2): 122–37. https://doi.org/10.5539/ijef.v7n2p122

Keen, M. (2008). 'VAT, Tariffs, and Withholding: Border Taxes and Informality in Developing Countries'. *Journal of Public Economics,* 92(10-11):1892–906. https://doi.org/10.1016/j.jpubeco.2008.05.006

Kleven, H.J., M.B. Knudsen, C.T. Kreiner, S. Pedersen, and E. Saez (2011). 'Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark'. *Econometrica,* 79(3): 651–92. https://doi.org/10.3982/ECTA9113

Kleven, H.J., and M. Waseem (2013). 'Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan'. *The Quarterly Journal of Economics,* 128(2):669–723. https://doi.org/10.1093/qje/qjt004

Kleven, H.J., C.T Kreiner, and E. Saez (2016). 'Why Can Modern Governments Tax so Much? An Agency Model of firms as Fiscal Intermediaries'. *Economica,* 83(330): 219–46. https://doi.org/10.1111/ecca.12182

Kopczuk, W., and J. Slemrod (2006). 'Putting Firms into Optimal Tax Theory'. *American Economic Review,* 96(2):130–34. https://doi.org/10.1257/000282806777212585

Kumler, T., E. Verhoogen, and J.A. Frías (2013). 'Enlisting Employees in Improving Payroll-Tax Compliance: Evidence from Mexico'. NBER Working Paper 19385. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w19385

Khwaja, M.S., R. Awasthi, and J. Loeprick (eds) (2011). *Risk-Based Tax Audits. Approaches and Country Experiences*. Washington, DC: World Bank. https://doi.org/10.1596/978-0-8213-8754-2

Løyland, K., O. Raaum, G. Torsvik, and A. Øvrum (2019). 'Compliance Effects of Risk-Based Tax Audits'. CESifo Working Paper 7616. Available at SSRN: https://ssrn.com/abstract=3384307 (accessed 23 September 2021).

Luttmer, E.F.P., and M. Singhal (2014). 'Tax Morale'. *Journal of Economic Perspectives*, 28(4): 149–68. https://doi.org/10.1257/jep.28.4.149

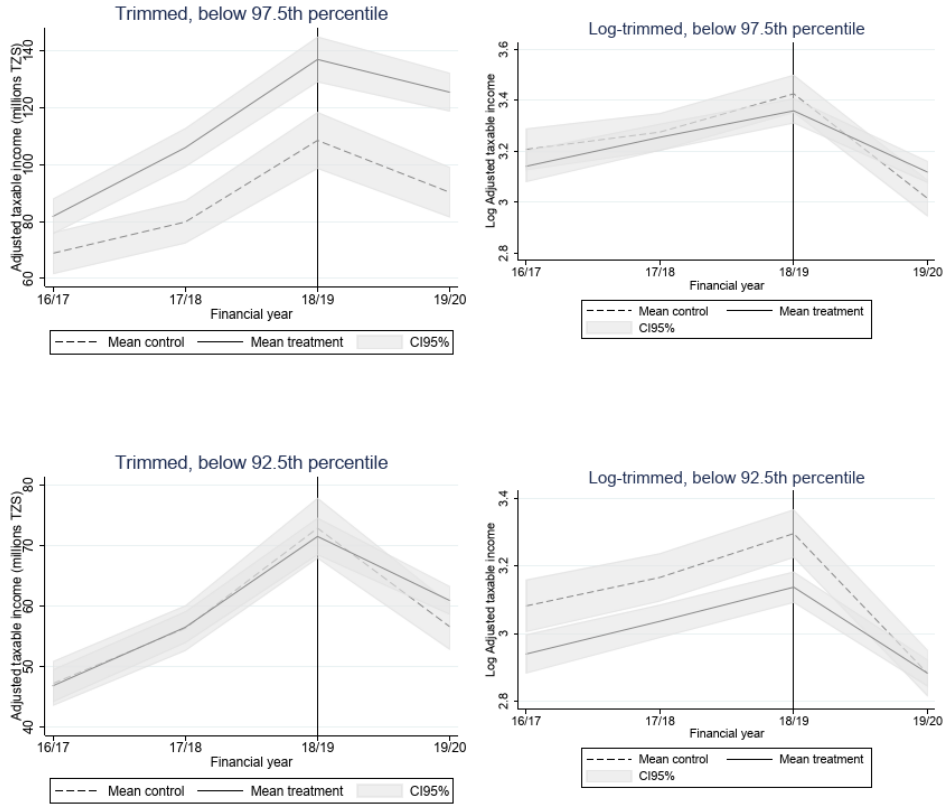
Pomeranz, D. (2015). 'No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax'. *American Economic Review,* 105(8): 2539–69. https://doi.org/10.1257/aer.20130393

Slemrod, J. (2019). 'Tax Compliance and Enforcement'. *Journal of Economic Literature*, 57(4): 904–54. https://doi.org/10.1257/jel.20181437

Wadhawan S.C., and C. Gray (1998). 'Enhancing Tax Transparency in Tax Administration: A Survey'. African Economic Policy Discussion Paper 3. Arlington: EAGER Publications/BHM.

# Appendix A

Figure A1: Parallel trends using adjusted taxable income below 97.5th percentile (above) and below 92.5th percentile (below). Log transformed adjusted taxable income on the right side.




Source: authors' estimates based on ITAX data.

Table A1: Estimated treatment effects, adjusted taxable income below 97.5th percentile

| | (1) Adj taxable inc | (2) Log taxable inc | (3) Adjusted | (4) Extra inc | (5) Log extra inc |
|---|---|---|---|---|---|
| Basic DD | 9.046 | 0.144*** | 0.0167 | 3.054 | 0.155*** |
| | (6.695) | (0.0516) | (0.0106) | (2.384) | (0.0496) |
| + controls | 13.52** | 0.114** | -0.00127 | 4.845** | 0.143*** |
| | (6.650) | (0.0517) | (0.0105) | (2.374) | (0.0497) |
| + clustered std. errors | 13.52 | 0.114* | -0.00127 | 4.845* | 0.143** |
| | (8.467) | (0.0663) | (0.0130) | (2.749) | (0.0612) |
| Observations | 27,755 | 27,755 | 27,755 | 27,063 | 26,709 |

Note: controls include tax type (PIT, corporate tax), taxpayer category (limited company, sole proprietor, partnerships, parastatal, unclassified class) and industry type. Signification levels: *** 1%, ** 5% and * 10%.

Source: authors' calculations based on ITAX data.

Table A2: Estimated treatment effects, adjusted taxable income below 92.5th percentile.

| | (1) Adj taxable inc | (2) Log taxable inc | (3) Adjusted | (4) Extra inc | (5) Log extra inc |
|---|---|---|---|---|---|
| Basic DD | 3.798 | 0.136*** | 0.0151 | 1.363 | 0.145*** |
| | (2.815) | (0.0487) | (0.0110) | (0.962) | (0.0451) |
| + controls | 5.612** | 0.0905* | -0.00418 | 1.628* | 0.111** |
| | (2.802) | (0.0486) | (0.0108) | (0.965) | (0.0450) |
| + clustered std. errors | 5.612* | 0.0905 | -0.00418 | 1.628 | 0.111** |
| | (3.293) | (0.0612) | (0.0134) | (1.097) | (0.0545) |
| Observations | 26,332 | 26,332 | 26,332 | 24,360 | 24,016 |

Note: controls include tax type (PIT, corporate tax), taxpayer category (limited company, sole proprietor, partnerships, parastatal, unclassified class) and industry type. Signification levels: *** 1%, ** 5% and * 10%.

Source: authors' calculations based on ITAX data.