Centre International pour la Fiscalité et le Développement
International Centre for Tax and Development
ICTD

Working Paper 82

Can ICTs Increase Tax? Experimental Evidence from Ethiopia

Giulia Mascagni, Andualem T. Mengistu, and Firew B. Woldeyes

July 2018

Bill & Melinda Gates foundation
IDS Institute of Development Studies
UKaid from the British people

ICTD Working Paper 82

Can ICTs Increase Tax? Experimental Evidence from Ethiopia

Giulia Mascagni, Andualem T. Mengistu, and Firew B. Woldeyes

July 2018

Can ICTs Increase Tax? Experimental Evidence from Ethiopia
Giulia Mascagni, Andualem T. Mengistu, and Firew B. Woldeyes
ICTD Working Paper 82
First published by the Institute of Development Studies in July 2018

ISBN: 978-1-78118-456-1

Available from:
The International Centre for Tax and Development at the Institute of Development Studies, Brighton BN1 9RE, UK
Tel: +44 (0) 1273 606261
Email: info@ictd.ac.uk
Web: www.ictd/en/publications

IDS is a charitable company limited by guarantee and registered in England

Charity Registration Number 306371
Charitable Company Number 877338

Can ICTs Increase Tax? Experimental Evidence from Ethiopia

Giulia Mascagni, Andualem T. Mengistu, and Firew B. Woldeyes

Summary

The widespread introduction of information and communication technologies (ICTs) and digitalised data management systems is one of the most important developments among African tax administrations in recent years. However, very little evidence is available on their effectiveness in practice, and how taxpayers respond to these changes. This paper starts filling this gap by reporting three sets of results from Ethiopia.

First, we show that the available data is still not used to its full potential, despite modern ICT systems being in place. Second, we find that a technological innovation, the introduction of electronic sales registration machines (SRMs), had a positive impact on both tax revenue and the accuracy of tax records. However, taxpayers responded to the machines by simultaneously adjusting both reported income and deductible costs, thus reducing the potential revenue gains. Third, we use a letter experiment to show that the main mechanism through which the SRMs increase tax revenue is tax compliance, rather than any change in real business activity.

Keywords: tax compliance; technology; tax administration; taxpayer responses; RCTs.

Giulia Mascagni is Research Director at the International Centre for Tax and Development, and Research Fellow at the Institute of Development Studies.

Andualem T. Mengistu is Director of the Macroeconomics and Trade Centre at the Ethiopian Development Research Institute.

Firew B. Woldeyes is a Research Fellow at the Ethiopian Development Research Institute.

Contents

Summary 3
Acknowledgements 6
Acronyms 6

1 Introduction 7

2 Background and context 10

3 Design, data and empirical methods 12
3.1 Research design 12
3.1.1 Difference-in-difference design 13
3.1.2 Letter experiment design 15
3.2 Data 17
3.2.1 Accuracy of taxpayer reports 20
3.3 Empirical framework 22
3.3.1 Difference-in-difference 22
3.3.2 Letter experiment: implementation and empirical framework 23

4 Results 26
4.1 Diff-in-diff results 26
4.2 RCT results 30

5 Robustness 32

6 Conclusions 33

Appendices 36
Appendix A Text of letters in English 36
Appendix B Summary statistics 38
Appendix C Parallel trends test using monthly data for VAT 39
Appendix D Balance tests (RCT) 40
Appendix E Additional diff-in-diff results 41
Appendix F Additional RCT results 42
Appendix G Additional robustness results 43

References 44

Tables

Table 1 Number of firms in the income tax sample 19
Table 2 Discrepancies between turnover reported for income tax and VAT 22
Table 3 The effect of SRMs on VAT declaration: annual data and positive VAT declaration 27
Table 4 The effect of SRMs on profit tax declaration: positive tax 28
Table 5 The effect of SRMs on accuracy 29
Table 6 The effect of SRMs on the probability of positive VAT declaration:

	annual data	29
Table 7	The effect of SRMs on the probability of declaring positive amounts	30
Table 8	Treatment effects: main results	31
Table 9	Treatment effects on VAT net tax	32
Table 10	Treatment effects: probability of declaring tax > 0	33

Figures

Figure 1	Dynamic response of VAT declarations for machine adoption – annual data	24
Figure 2	Dynamic response of income tax declarations for machine adoption	24

Acknowledgements

This paper was prepared with financial support from the International Centre for Tax and Development (ICTD). We are also grateful to the Ethiopian Revenue and Customs Authority for granting access to the data. Special thanks are due to Kiu G. Molla, who provided excellent support in the preparation of our dataset. We received useful comments from participants at the Ethiopian Tax Research Network's Launch Conference, the ICTD Annual Meeting, an IDS Member Seminar, and a Centre for the Study of African Economies Seminar. All remaining errors are our own.

Acronyms

ERCA	Ethiopian Revenue and Customs Authority
ITT	Intent-to-treat
LTO	Large taxpayers' office
RCT	Randomised control trial
SOE	State-owned enterprise
SRM	Sales registration machine
VAT	Value added tax

1 Introduction

African tax administrations have experienced rapid modernisation in the past two decades. The digitalisation of tax records and widespread adoption of ICTs have been key features of this process, not only for tax administration but also for public sector management more broadly. Most countries today operate automated financial management, customs, and 159 out of 193 UN member states have automated tax systems (World Bank, 2016). One example of this digital revolution in tax administration is e-filing, which allows taxpayers in several low-income countries, including Ethiopia, to file their tax returns online. Some countries, like Rwanda, allow micro taxpayers to file through their mobile phones. Another example is the adoption of electronic sales registration machines (SRMs), which store information on business transactions and communicate this directly to the tax authority. They have already been introduced in several low-income countries, including Ethiopia, Rwanda and Malawi, while others are considering their adoption, like Uganda.

Despite the widespread use of ICTs in tax administration and the radical changes they have brought to the taxpaying process, there is very little evidence on their impact on tax administration and tax compliance in low-income countries. One exception is a recent report on e-filing in Tajikistan, which shows significant decreases in compliance costs and increases in tax payments for some firms (Okunogbe and Pouliquen, 2017). As far as sales registration machines (SRM) are concerned, two recent reports evaluated their impact in Rwanda and Ethiopia. The former shows that SRMs resulted in an increase in tax payments of about 8 percentage points. This average effect, however, hides large differences in terms of sectors and firm size, with smaller firms experiencing larger increases (Eissa and Zeitlin, 2014). In the case of Ethiopia, Ali et al. (2015) looks at the impact of SRM adoption on value added tax (VAT, both reported sales and tax), employment, and firms' net entry. The machines are found to lead to positive and significant increases in VAT and employment, while they have no effect on net entry.

Our paper builds on these results in three ways. First, we provide quasi-experimental evidence on the impact of SRMs on both VAT and profit taxes, as well as checking any impact on the accuracy and consistency of taxpayer self-reports across tax types. Second, for each tax type, we investigate the machines' effect on outcomes on both the income and

expense sides. We find that, while firms increase reported sales, they also make offsetting changes to expenses. Third, using a randomised control trial (RCT), we show that the main mechanism leading from SRM adoption to higher tax revenue is increased tax compliance.

Our analysis is directly related to the recent literature on the key role of information in modern tax administration. Successful revenue administrations in high- and middle-income countries rely heavily on firms as collectors of both information (Kopczuk and Slemrod, 2006) and tax revenue, through withholding of employment income tax (Kleven et al., 2011). Third-party reporting, one example of which is income tax withholding, has been shown to be very effective in curbing opportunities for tax evasion in high- and middle-income countries (Brockmeyer et al., 2016). As a result, the self-employed, for which third-party reported information is not available, comply less than employees (Kleven et al., 2011). Similarly, Pomeranz (2015) shows that evasion is higher in the parts of the VAT chain that are not covered by the paper trail, namely sales to final consumers. However, Carrillo et al. (2017) also show that the effectiveness of third-party reporting can be severely limited in contexts of low state capacity, where it cannot be complemented with more traditional forms of enforcement, such as audits.

Besides its advantages for tax administrators, the availability of clear and verifiable information on taxable transactions can affect compliance through other channels. For example, Naritomi (2016) shows that consumers can act as tax auditors when they are given incentives to require receipts, with positive effects on compliance. Along similar lines, Kleven et al. (2009) argue that employees can contribute to tax enforcement through whistleblowing threats, that are only possible thanks to accurate and accessible business records. The availability of information can also encourage semi-voluntary compliance through peer effects and social norms, as evasion becomes more visible (e.g. not issuing a receipt).

In addition, firms themselves may benefit from the technological innovation induced by the revenue authority, such as the mandatory requirement to use sales registration machines. A recent survey revealed that Ethiopian firms found SRMs particularly useful for getting updated and easily available sales information and to reduce the opportunity of theft by employees (World Bank, 2015). Moreover, the availability of accurate and accessible tax records should reduce corruption and predatory behaviour by tax officials,

as digitalised records make it possible to easily cross-check data for any particular firm (Olken and Pande, 2011). Generally, better business records should make firms' interactions with tax officials easier and less arbitrary, as well as helping to reduce compliance costs (Okunogbe and Pouliquen, 2017).

A related literature has shown that ICTs can be highly effective in other aspects of public administration. Lewis-Faupel et al. (2016) evaluate the introduction of electronic procurement in India and Indonesia, and find that it significantly improved the quality of public contracts in both countries. Muralidharan et al. (2014) investigate the introduction of smartcards used to pay beneficiaries of employment and pension programmes in India, finding that they led to a faster, more predictable and less corrupt payment process.

Therefore, the introduction of ICTs is expected to bring about several benefits, including and beyond increased tax revenue. However, the extent to which ICTs can realise their full potential in low-income countries, where evasion is widespread and state capacity generally low (Besley and Perrson, 2013), is still unclear. High evasion means that many taxpayers are outside of the tax net altogether, while the tax administration has a strong incentive to focus on those that are more visible and already in the tax net (Gordon and Li, 2009). Moreover, limited capacity may result in the inability to make full use of data, even if it is available in digital format. Anecdotal evidence shows that revenue authorities in low-income countries often do not cross-check their tax records, even if it would be relatively easy to do so (Kangave et al., 2016; Almunia et al., 2017). In this context, ICTs may not yield their expected benefits on tax revenue and compliance, let alone the possible spillover effects along the value chain (Pomeranz, 2015).

Against this background, this paper reports three sets of results that start to fill the research gap in the field of technological innovation in tax administration. First, we provide evidence that the Ethiopian Revenue and Customs Authority (ERCA) is not making full use of the data that is already available to them, which is compatible with the presence of severe capacity constraints. We find several discrepancies in taxpayers' accounts that could easily be detected by simple cross-checks. Second, we provide an evaluation of the impact of SRMs on tax revenue and the accuracy of tax records. Despite having a positive effect on tax revenue, SRMs result in simultaneous adjustments to both firms' reported income and deductible costs, thus substantially reducing the potential revenue gains. We also find

a positive effect on the accuracy of firms' reporting. Third, we use an RCT to test the mechanism through which SRMs increase tax revenue. We show that the revenue increase is mostly due to increases in compliance, rather than real changes in business activity. The RCT also provides an opportunity to indirectly test the effectiveness of SRMs. We show that those who have already adopted them do not respond to an increase in the probability of audit - therefore confirming higher compliance amongst SRM users.

2 Background and context

Ethiopia is one of the most populous and fastest growing countries in Africa. It is also one of the poorest. With a tax-to-GDP ratio of less than 13 per cent, the country experiences a large gap in its capacity to finance development efforts, which is partly filled by foreign aid and borrowing. The low tax ratio is explained by widespread tax evasion, tax incentives for investors and severe constraints in administrative capacity, amongst others. In recent years the government of Ethiopia has been increasing its enforcement efforts, with the aim of tackling evasion. It is against this background that SRMs were introduced with a new law in 2007. SRM adoption was not linked to a specific tax type, although it was expected to help compliance, particularly with the VAT. The VAT was introduced in 2002, and had been slow in producing the expected revenue gains.

Implementation was slow after the 2007 law, partly because of changes in the tax administration itself around the same time. Most notably, ERCA was established in 2008 as a result of the merger of three institutions previously in charge of tax and customs administration.[1] Partly as a result of this, actual implementation of the new SRM law only started in 2008. Each business was legally required to have a machine installed at their premises, through one of the suppliers certified by ERCA. Firms were expected to buy the machine at their own cost.[2] SRM adoption was compulsory for all firms, so they did not have any legal option to avoid it or to voluntarily opt out. ERCA communicated to each

[1] Although ERCA was formally established in 2008, there was much institutional continuity with the previous institutions in charge of tax administration, so this institutional change did not represent a major disruption.

[2] The machine's cost was between 5000 to 14000 birr in 2013.

firm that they should get an SRM, according to the rollout plan discussed below. Once the machine is installed, the firm is meant to bring the relevant receipt to its ERCA branch office, as proof. Firms have generally complied with this process, as not doing so would put them as risk of business closure. In principle, all firms that have an SRM should use it for every transaction, including business-to-business, and without exemptions. Failure to use the machine can result in harsh monetary penalties and even imprisonment, with a serious risk of business closure.[3] As far as our data is concerned, a business appears as an adopter in ERCA's database once it has installed an SRM.

The rollout started in the capital, Addis Ababa, and was planned to take place gradually between 2008 and 2011, with other areas of the country to follow afterwards. The rollout of SRMs was not randomised. VAT-registered businesses in the hospitality sector (i.e. hotels, bars and restaurants) and large taxpayers came first, while others followed largely based on the sector in which they operate. However, in practice the rollout process was quite chaotic and ad hoc. The challenges were mainly related to ERCA's administrative capacity, but also to taxpayer resistance and technological problems related to a weak communication infrastructure. As a result, the process suffered from several delays. As shown in section 3.2, the rollout process continued well beyond 2011, as many of the firms in our dataset were still adopting the SRM in 2012, 2013 and 2014.

The main objective of the introduction of SRMs was to improve the quality of data available at ERCA and, therefore, to increase enforcement amongst registered businesses. All firms in the capital (and, later, all regions) were required to start using SRMs eventually, except micro taxpayers.[4] The key innovation introduced by SRMs was the ability to digitally record data from each sale and transmit it to ERCA in real time. Besides better data, the machines did not change anything else in the process of declaring or paying taxes.[5] Taxpayers are still required to file for VAT monthly, and for income taxes yearly. Declarations for the latter are due within four or two months after the end of the tax year, depending on the type of taxpayer.[6] Most firms follow the Ethiopian fiscal year, which

[3]When there is a technical problem with the machine, the firm is given an official receipt book authorised by ERCA. This can be used for a maximum of five days while the problem is being fixed.

[4]Firms in category A and B (i.e. above 100,000 birr, according to the tax laws of the time) were legally mandated to use the machine, while those in Category C (below 100,000 birr) were not.

[5]For example, taxpayers' returns have not been pre-populated with data from the machines.

[6]Category A taxpayers (turnover above 500,000 birr) are required to file within four months, while

runs from July to June. Therefore the tax declaration for any year ending in June is due between 8 July and 9 November.[7] Usually all declarations in Ethiopia involve some contact with tax officials, either in the process of preparing the declaration, submitting it, or paying the tax due. The frequency of contact with tax officials did not change as a result of SRM introduction. Similarly, there was no major tax reform in the period covered by our data; the relevant tax laws for VAT and income taxes date back to 2002.

Therefore, the introduction of SRMs was largely an increase in enforcement capacity, both because of the increased quantity and quality of data available at ERCA (e.g. to feed into the risk management function), and because of the possible reallocation of resources within ERCA from pure data entry jobs to more advanced tasks, such as audits. SRMs were expected to increase tax revenue, lead to a more effective tax administration and promote better data management.

3 Design, data, and empirical methods

3.1 Research design

Our research design is based on two interrelated components: a quasi-experimental difference-in difference (diff-in-diff) analysis and a letter experiment. With the former, we evaluate primarily the effect of the SRM on tax collection - while also looking in further detail at offsetting responses on the expense side and impact on the accuracy of reporting. Having established that SRMs significantly increase tax revenue, we explore the mechanism through which this effect comes about. Tax compliance is an obvious candidate, as records become more verifiable by the revenue authority and taxpayers may even experience lower compliance costs thanks to the new technology (Okunogbe and Pouliquen, 2017). However, the increases in reported income and tax may also be due to changes in the level of economic activity. On the one hand, this could be the result of improved efficiency generated by bet-

category B taxpayers (turnover between 500,000 and 100,000 birr) are required to file within two months. Source: Proclamation No. 286/2002.

[7]For simplicity, in this paper we refer, for example, to year 2013 to indicate the fiscal year from July 2013 to June 2014.

ter business records and less leakage of firm profits, for example due to theft by employees. In a recent survey, Ethiopian firms highlighted a number of benefits resulting from SRM adoption, such as 'less opportunity for theft' (mentioned by 20 per cent of respondents), 'updated and easily available sales information' (22 per cent), 'easy to comply with tax requirements' (17 per cent) and 'better sales and inventory control' (16 per cent) (World Bank, 2015, p.26). On the other hand, the cost of the machine and increased enforcement pressure might cause a reduction in economic activity, as firms face higher costs than they did before (machine cost, maintenance, more tax payments). While the diff-in-diff analysis cannot shed light into the mechanism leading from SRM introduction to higher revenue, the RCT can. The experiment involves a fairly standard intervention in the literature, consisting of a letter sent to firms by the tax authority (more details in section 3.1.2).[8] These letters increase taxpayers' perceptions of the probability of being detected evading. As such, they should generate increases in tax compliance only for those firms that were not already compliant. Therefore, if the mechanism leading from the SRM to higher tax is compliance, SRM users would not respond to the letter, while non-SRM users would. In this context, the RCT serves both as an indirect test of the impact of the SRM and as a test of the mechanism behind the machines' effectiveness (Ludwig et al., 2011; Pomeranz, 2015). This is particularly relevant in our case, since the policy itself was not randomised.

3.1.1 Difference-in-difference design

Our analysis starts with a difference-in difference analysis looking at the effect of the SRM on VAT and profit tax (the exact dependent variables we use are described in detail in section 3.2). This initial analysis provides evidence that overall the machines were effective in increasing tax revenue. To explore this result further, we also explore two other sets of outcome variables. First, we unpack the details of firms' responses by looking separately at sales (total income for profit taxes, and VAT sales for VAT) and expenses (total deductible expenses for profit tax, and VAT inputs for VAT). This allows us to identify any offsetting changes that firms may make to their reported expenses, to compensate for higher reported income. Second, we look at the effect of SRMs on the accuracy of firms' reporting, measured by the discrepancy between two measures of turnover that are both reported by the firm,

[8] For a comprehensive review of the literature on tax experiments, see Mascagni (2017).

one for VAT and the other for profit tax purposes (see section 3.2). To make the analysis of VAT and profit tax comparable, all our main results are based on annual data. However, monthly data is available for VAT, as all firms need to declare on a monthly basis for this tax type (see section 3.2.1).

As mentioned in section 2, the introduction of SRMs was not randomised. However, our diff-in-diff estimation would still identify the SRM's effect if the key assumption of pre-treatment parallel trends is satisfied. In other words, treated firms (SRM adopters) should not have systematically different trends pre-adoption compared to un-treated firms (those that have not yet started using the SRM). Note that differences in firm characteristics that are fixed in time, such as location or sector, do not threaten our identification strategy. Given the importance of the parallel trend assumption, we check it in three different ways.

First, we investigated the details of SRM rollout, paying particular attention to any possible link between adoption and time-varying firm-level variables, such as growth and performance. We conducted several consultations with ERCA officials and looked into the details of the rollout plan and its practical implementation. We obtained no indication that rollout could be linked to time-varying firm characteristics. Although in principle the process was organised in clear rounds, in practice the challenges and delays during implementation introduced an element of randomness in actual adoption (although we do not claim random allocation of the SRM).

Second, we directly test the parallel trend assumption using both monthly and annual data for VAT, while for income tax we rely only on annual data. The obvious challenge in the case of annual data is the short time period available before adoption. However, we can still rely on three years pre-adoption for the firms included in our sample (see section 3.2), so that the assumption can be tested also using annual data. Our tests do not reveal any systematic difference in pre-adoption trends (see section 3.3.1), therefore confirming the validity of our identification strategy.

Third, we take particular care in selecting a sample of firms that are relatively homogeneous in terms of location and adoption period. For example, we exclude adopters involved in the initial rounds of the rollout (see section 2), and we only consider firms registered in Addis Ababa. Section 3.2 reports more details on these sample restrictions and on the types of firms included in this study. However, it must be stressed that our

identification strategy would only be threatened if those firms have significantly different trends before adoption, while our model can deal with differences in the level. For example, we can compare firms of different size as long as they have a similar growth trajectory pre-adoption.

3.1.2 Letter experiment design

This component of our analysis relies on an intervention that was carried out as part of a separate study (Shimeles et al., 2017). This RCT involved about 5,400 taxpayers, which were randomly allocated to one of two treatment groups or to a control group.[9] The treatments are letters, with two alternative contents that were meant to potentially stimulate compliance: one based on persuasion and one on deterrence. Treatment allocation was based on stratified randomisation using location and sector as strata.[10] All taxpayers involved in the study are firms, both corporations and non-incorporated businesses that are registered in one of the five main business districts of the capital: Addis Ketema, Bole, Kirkos, Gulele and Nifas-Silk-Lafto.[11] In line with the tax experiments literature, the letters were sent by ERCA, not by the research team, and signed by the manager of the relevant ERCA branch. They were delivered in person by ERCA representatives to treated businesses between June and August 2014, which corresponds with the end of the financial year and the beginning of the declaration period for income taxes.[12] Therefore, the baseline

[9]There are ten sub-cities in Addis Ababa, out of which only five were selected for the original study. Three sub-cities are business areas, while the other two are more residential areas. There are more than 86,000 firms that are registered and keep books of account in the five sub-cities (excluding government enterprises and institutions). For the RCT, about 1,350 taxpayers were selected from each of the three sub-cities in the business areas, and 1350 firms in total from the two sub-cities considered mainly residential. The sampling frame is the universe of taxpayers located in the five sub-cities in Addis Ababa, according to a list provided by the relevant ERCA branches. Taxpayers were grouped into three broad sectors and an equal numbers of taxpayers were selected from each using simple random sampling.

[10]The sector variable distinguishes between wholesalers, other services, and manufacturing and agro-processing. The location variable is based on five districts of Addis Ababa.

[11]The selection of the sub-cities is based on the authors' (Shimeles et al., 2017) discussions with the law enforcement directorate of ERCA. They have covered the three sub-cities known to have a high concentration of business activities and two sub-cities that represent businesses in mainly residential sub-cities.

[12]The ERCA representatives are enumerators who were recruited specifically for this purpose, and were given orientation by officials from ERCA.

declaration is the one for fiscal year 2012, while the post-treatment period is 2013.[13] All these aspects of the RCT design are largely in line with similar field experiments on tax compliance (Mascagni, 2017).

The original study (Shimeles et al., 2017) had two two main objectives. First, the authors wanted to check whether Ethiopian taxpayers are more responsive to coercion or persuasion. To do this, they investigated the effectiveness of two separate treatment letters. The coercion letter informed taxpayers that ERCA is aware that many businesses are not fully compliant, and stressed the existence of audits and penalties to deal with tax evasion. The persuasion letter instead focussed on the civic duty to pay taxes, and their importance in contributing to the country's development, including flagship projects financed entirely through domestic revenue, such as the Renaissance Dam. The original version of both letters is in Amharic – an English translation is reported in the Appendix. The results show that both letters have a large and significant effect on reported tax, ranging from 32 per cent to 38 per cent. However, the coefficients of the two treatments are not statistically different from each other, suggesting that receiving any letter from the revenue authority increases compliance. This finding is in line with the literature from low- and middle-income countries, where receiving any letter from the tax authority is quite uncommon, and might be perceived as a deterrence intervention (Mascagni et al., 2017; Del Carpio, 2014; Ortega and Scartascini, 2016). Second, having found significant treatment effects, the authors explore whether they lasted one year after the intervention was implemented. They find that taxpayers largely regress to their previous compliance behaviour, as the short-term effect almost entirely disappears the following year. Again, this is in line with the very scant literature on the long-term effects of letter experiments (Manoli and Turner, 2014).

Our paper exploits the letter intervention from this RCT as an exogenous variation in the probability of being detected in evading taxes. This variation should increase compliance if taxpayers were not already compliant before the intervention.[14] Note that the

[13]The intervention was implemented in the declaration period for fiscal year 2013 (i.e. 2013/14). As a result, it affects only reporting behaviour, since all economic activity for that fiscal year is complete by the time the treatment is administered.

[14]This is true under the plausible assumption that taxpayers would not pay more than the amount they should pay if they were fully compliant, so the letter is only effective when evasion is present and there is

response elicited by the letter is attributed to increases in compliance (i.e. reporting response), rather than real business activity (i.e. real response), since taxpayers receive the letter once the economic activity for the year has already taken place. In this context, we use this RCT as a mechanism experiment to explain the relationship between introduction of SRM and increased tax. While the diff-in-diff analysis shows that such relationship exists, it cannot identify the mechanism through which it occurs. If the observed increase in tax is due to better compliance, then we expect SRM users not to respond to the letter intervention, since they have already adjusted their compliance as a result of SRM adoption. On the other hand, non-SRM users should respond more than those who have already adopted the SRM, as they are still less compliant. In other words, if the SRM's effect on tax revenue is due to an increase in compliance, those who have adopted it should not respond to successive deterrence interventions - or respond less than non-adopters. By testing the difference in taxpayers' response to the letter between SRM users and non-users, the RCT also provides a useful indirect test for the effectiveness of the SRM (Ludwig et al., 2011; Pomeranz, 2015)).

In the diff-in-diff analysis we look at the SRM's effect on outcomes related both to VAT and income tax. By doing this, we can make the RCT results consistent with the diff-in-diff ones, although the RCT intervention focussed more specifically on income taxes (see letters in Appendix).

3.2 Data

All data used in this study come from administrative records from firms' tax returns. These are firm-level data routinely collected by ERCA in the process of tax collection. The data is available in the same temporal dimension as the relevant declaration: monthly for VAT and yearly for profit taxes. Although we use the monthly VAT data to test the parallel trend assumption and for diff-in-diff robustness, our main results are all based on yearly data. In the case of VAT, the data was annualised simply by adding up the declarations for all months in the year, calculated from July to June to keep consistency with the financial year used for income tax declarations.

---

a margin to increase compliance.

Our dataset is a panel covering five years, from 2010 to 2014.[15] Although some data prior to 2010 is available, it does not cover the universe of un-incorporated firms. This is largely due to the extension, in 2010, of the data management system (SIGTAS) to non-incorporated firms, while corporations were already captured in the system. The process of migrating firms to SIGTAS started with VAT-registered firms, and was then completed with the remaining ones. As a result, the data for 2010 is complete for VAT-registered firms, but not for others. Therefore, the VAT analysis uses one more year (2010) than the analysis of income taxes, which starts in 2011. In addition, the sample is restricted to those firms that start using the machine sometime between 2011 and 2014.

We use all years of data for the diff-in-diff (see section 3.3.1), while for the RCT we simply use the baseline (2012) and the post-treatment period (2013).

As far as SRM adoption is concerned, we have data on the exact date of adoption for each firm, which is used to generate a dummy variable for users (0 for non-users). Since the main results are based on yearly data, we consider a firm as an adopter in any given year if they had the machine for at least six months out of twelve. Since the fiscal year starts on 8 July, a firm would need to adopt before January of the following year to be considered a user. This definition is obviously simpler with monthly data, as the SRM dummy simply switches on in the month of adoption. Table 1 reports the sample by year, confirming that the rollout continued throughout our data period. It also shows that the number of observations for VAT taxpayers is lower, which is expected because not all firms included in the dataset are required to register for VAT (see section 2).[16]

Our dataset contains all firms that are registered with ERCA and file tax declarations, except micro taxpayers.[17] We restrict the sample to firms registered in Addis, because both the SRM rollout in the relevant period and the RCT focussed on the capital. We also exclude all taxpayers registered in the large taxpayers' office (LTO) and state-owned enterprises (SOEs), both for the diff-in-diff and the RCT. These taxpayers are generally

[15] These correspond to years 2010/11 and 2014/15 in the Gregorian calendar.

[16] Category B taxpayers (turnover between 100,000 and 500,000 birr) are included in our dataset, but they are not required to register for VAT. The VAT registration threshold is 500,000 birr, therefore affecting only Category A (firms with turnover above 500,000 birr).

[17] Our dataset includes all Category A (turnover above 500,000 birr) and B (turnover between 100,000 and 500,000 birr) taxpayers, while it excludes category C ones (below 100,000 birr).

Table 1: Number of firms in the income tax sample

	Fiscal year	2010	2011	2012	2013	2014
VAT payers	Number of firms	3226	5411	9401	11931	13011
	Number of adopters	0	1670	3171	6471	8102
All firms	Number of firms		15749	23002	29001	29814
	Number of adopters		0	8121	17453	23206

more compliant, and they would be excluded from our sample anyway because they adopted either before or after the relevant period covered here.[18] While the diff-in-diff uses the full population of eligible firms, the RCT is restricted to those who were selected to take part in the experiment (see section 3.1.2 and 3.3.2).

As far as the dependent variables are concerned, we use two sets of variables, both of which are calculated for VAT and for income tax separately. The first set is tax due (for VAT and income taxes), transformed in log. The second set (again for both VAT and income tax) is a dummy variable capturing the probability of filing a return with a positive tax liability. This aspect is particularly relevant in Ethiopia, due to the relatively large proportion of taxpayers with zero tax liability. Some of these are nil-filers, namely taxpayers that declare zero turnover, while others are firms for which expenses are higher than sales (especially for income taxes). Although the phenomenon of nil-filers is found in other African countries, the reasons behind it are still unclear (Mascagni et al., 2017; Almunia et al., 2017). Since in principle the machines can affect both sets of outcomes, we include both of them in all components of our analysis.

In Appendix B we report the key summary statistics for the main variables used in our analysis.

[18] LTO was the first to adopt, while SOEs came towards the end.

3.2.1 Accuracy of taxpayer reports

We measure the accuracy of taxpayers' reports by investigating the discrepancy between turnover as reported for VAT and income tax purposes. In addition to serving as a dependent variable, our measure of accuracy provides some interesting evidence of the extent to which the potential of administrative data in low-income countries remains untapped. This section reports some descriptive statistics on taxpayer reports' accuracy, before describing the empirical framework.

In theory, there is no reason why taxpayers should report a different value of turnover for the two tax types. Turnover as reported on both declarations is used to compute taxes (either VAT or profit tax), so it always has a potential monetary consequence for taxpayers that, in turn, may want to under-report to decrease their tax liability. However, the VAT rate (15 per cent) is applied directly on sales, while the profit tax rate is applied on taxable income, which is turnover minus deductible expenses. In this sense, the monetary gain from under-reporting turnover in the VAT declaration might be greater compared to under-reporting in the income tax declaration, where turnover has a more indirect relationship with taxes, and where the latter can be reduced more safely by inflating deductible expenses. While input costs also reduce tax liability in the case of VAT, through refund claims, the self-enforcement mechanism of VAT means that, in practice, it would be harder to inflate input costs for VAT purposes than for profit tax purposes.

In terms of the measurement of our turnover variable, we are comparing like with like across declarations: total sales on both sides. Importantly, we take total sales for VAT declarations and not VAT sales. Although there are VAT exemptions and zero-rated goods, firms are still required to declare all the relevant amounts in the tax declaration. Although the two measures of turnover should be comparable according to the law, we still expect some mistakes and some level of misinterpretation of the law. To avoid picking up rounding errors and small mistakes, we calculate discrepancies with a margin of error of 200 birr. We consider any small discrepancy below 200 birr to be equal to zero. Still, if firms mistakenly do not report exempted sales in their VAT declaration, and those sales are greater than our margin of error, then we would observe a discrepancy where turnover in the income tax declaration is greater than in the VAT declaration.

Table 2 shows three relevant metrics on accuracy, for firms included in our sample

according to the criteria explained in section 3.2. The first column reports how frequently discrepancies occur, that is the share of total declarations in our sample for which we observe any discrepancy (greater than 200 birr, our margin of error). These numbers are quite large: on average in all years, 60 per cent of firms reported discrepancies in turnover in their declarations. The second column shows that, when a discrepancy occurs, it is more likely to be due to under-reporting in the VAT declaration, compared with the income tax declaration. This case captures about 66 per cent of discrepancies, on average in all years. This observation is consistent with the idea mentioned above – that the monetary gain from under-reporting turnover for VAT declaration purposes is greater than under-reporting in the income tax declaration. However, an alternative explanation would be firms' misunderstanding of their legal obligation to report all sales, even if they are VAT exempt, in their VAT declaration. With the data at hand, we are unable to distinguish between these two alternative explanations. Finally, the third column shows how large the discrepancy is, in cases where one exists at all. This is the absolute value of the ratio between the discrepancy and the larger turnover as reported in the income tax declaration or the VAT declaration.[19] These figures show that discrepancies are not only frequent, but also large – representing on average 48 per cent of the largest turnover declared. While we cannot rule out that these discrepancies are, at least partly, due to mistakes and misinterpretation of the law, their extent and depth suggest that evasion is likely to be part of the explanation. If indeed the amount of the discrepancy could be recovered, so that firms that under-report for VAT would pay the standard 15 per cent rate on the amount of discrepancy, ERCA would have obtained an extra 634 million birr in VAT revenue in 2014 (about USD 23 million).[20] However, there are at least three important caveats to this figure. First, the discrepancy may be entirely due to exempted VAT sales, so that no tax would be due on them. Second, it is likely that firms might increase their input claims when forced to pay more VAT on outputs (as we document in section 4), so that the revenue gains would be eroded by offsetting refund claims. Third, the revenue gain calculation is

[19]Note that, by measuring the ratio using the largest turnover at the denominator, we are implicitly assuming that the correct value is the greater one of the two.

[20]This is a rough calculation, where the amount of discrepancy, for each firm for which VAT turnover is larger than income tax turnover, is multiplied by the VAT rate of 15 per cent. All the revenue gains calculated this way are then added up, for 2014.

based on the assumption that turnover reported in the income tax declaration is correct. However firms may claim it was mistaken and try to revise it downwards. All these caveats mean that the revenue gains may not materialise in practice.

Table 2: Discrepancies between turnover reported for income tax and VAT

	Frequency of discrepancy	VAT < income	Size of discrepancy	N
.				
2010	58.00	71.5	49.4	1871
2011	61.1	70.3	49.6	3304
2012	59.2	67.5	46.2	5564
2013	60.4	63.0	44.6	7206
2014	60.8	65.7	49.1	7908

Finally, it is worth noting that taxpayers would only consciously report discordant figures if they do not think the revenue authority stores the data to cross-check, or if they do not believe that it has the capacity to do so. If the SRMs correct these perceptions, for example by making taxpayers more aware of the data available to ERCA, we would expect to observe a positive impact of SRMs on accuracy.

3.3 Empirical framework

Based on the research design and data described so far, the next two sections go into the details of the empirical framework for both components of our analysis.

3.3.1 Difference-in-difference

In the diff-in-diff component we need to observe firms before SRM adoption, to check for the parallel trend assumption, as well as afterwards. To this aim, treated firms in our data are only those that adopted between 2011 and 2014. By doing this, we exclude early adopters. We obtain a sample of firms that have a minimum of two data points prior to adoption and at least one afterwards (including the adoption year). Control firms are those

who had not yet adopted the machine in the relevant period.

Our estimation follows a rather standard diff-in-diff methodology, with staggered treatment. The standard equation is as follows, where γ_i are firm fixed effects and φ_t is a time fixed effect. *SRM* is a dummy that switches on in the year of adoption (see section 3.2). The dependent variable in our case is either tax due or the probability of declaring positive, as described in section 3.2. In the latter case, we report results from the linear probability model, as it is easier to estimate a fixed effects model using this methodology.

$$Y_{it} = \alpha + \beta SRM_{it} + \varphi_t + \gamma_i + \epsilon_{it} \tag{1}$$

This equation can identify the SRM treatment's effect on tax due after adoption, provided that the parallel trend assumption is satisfied. To test this assumption, we estimate the following equation, where we include years both before (*Pre* variables) and after adoption (*Post* variables), the latter including the year of adoption. As in the previous equation, γ_i are firm fixed effects and φ_t is a time fixed effect, while the dependent variable is tax due. SRM_{it+k} is a dummy equal to 1 if the firm is k periods before(after) adoption in period t.

$$Y_{it} = \alpha + \sum_{k=-3}^{2} \beta_k SRM_{it+k} + \varphi_t + \gamma_i + \epsilon_{it} \tag{2}$$

If there is no parallel trend, we expect the *Pre-adoption* dummies not to be statistically significant. On the other hand, the significance of the *Post-adoption* dummies provides a first indication of the effectiveness of the machine. Figures 1 and 2 plot the coefficients from this regression, both for net VAT and profit tax. They confirm that there is no statistically significant difference between users and non-users before adoption. Moreover, we report in Appendix C the same test based on monthly data for VAT. The same result holds.

3.3.2 Letter experiment: implementation and empirical framework

The empirical framework for the RCT is rather straightforward, as the letter intervention was randomised. However, it is worth mentioning at the outset that the RCT was not designed to look at heterogeneous effect of letters between adopters and non-adopters. The underlying assumption of using the RCT to compare these two groups is that they

Figure 1: Dynamic response of VAT declarations for machine adoption - annual data

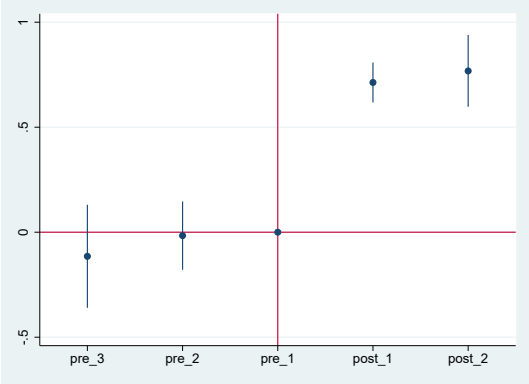
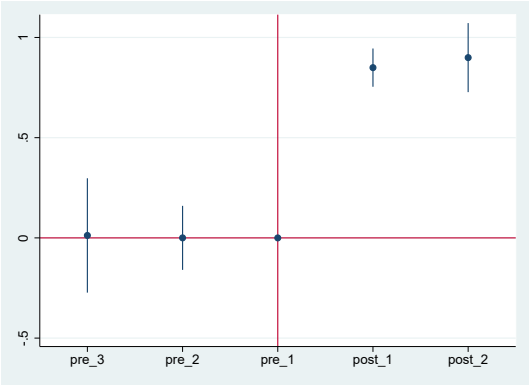


Figure 2: Dynamic response of income tax declarations for machine adoption



would have reacted the same way to the letter intervention, in absence of the machine. In other words, there should be no pre-existing difference in compliance before adoption. We cannot directly test this assumption, so this remains a caveat to our results. However, the details on SRM rollout described in section 2 give us some confidence that the extent of tax compliance was not a relevant criterion to decide on the timing of adoption, especially not in the later stages of the process (i.e. 2012 and 2013, the relevant years for the RCT).

Moreover, there were at least two challenges in the RCT's implementation. The first one is that the randomisation was done on a list of taxpayers, provided by the ERCA main office, which included only information on sectors and sub-cities for all taxpayers registered in Addis Ababa. Although this still made it possible to randomise based on the sector and

location strata, at the design stage there was no information on tax declarations prior to the treatment. Therefore, about 25 per cent of the taxpayers included in the treatment plan (either in the treatments or in the control groups) failed to declare at all for 2013. This proportion is larger for VAT declarations (42 per cent), which could be expected since the RCT focussed particularly on income taxes. As a result, we are left with a sample of 4,051 taxpayers for income tax and 3,151 for VAT. Since some of the taxpayers included in the RCT reported zero tax (see section 3.2), in the regression focussing on those with positive tax we have an even smaller sample. Importantly, there is no significant difference in means between taxpayers included in the RCT sample and those in the broader population, for any of the following variables: VAT sales, VAT liability, turnover reported for income tax, and income tax liability.

The second challenge is related to compliance with the treatment, which, as with many other policy experiments, could not be mandated. Amongst those who declared at all for income tax, about 75 per cent of those assigned to a treatment group actually received the letter. Although we have information on who received the treatment and who did not, we prefer the intent-to-treat (ITT) estimates, as they take into account a level of non-compliance that is inherent in all policy interventions. All our results are ITT effects.

Since this component explores particularly the comparison between two sub-groups (SRM users and non-users) of an already limited sample, our main results focus on the effect of receiving any letter. In other words, we pool the coercion and persuasion treatments together. This choice is supported by the findings of the main study, where the two treatments are not found to have statistically different effects. In fact, the authors recognise that they cannot rule out the possibility that even the persuasion treatment was perceived with an element of deterrence (Shimeles et al., 2017). To make sure, we repeated the regression with disaggregated treatments, and the two coefficients on the letters are indeed largely comparable in magnitude and statistically the same. Although our main results use the pooled treatments, we also explore the effect of disaggregated treatments - finding again that there is no significant difference between the two.

As mentioned in section 3.1.2, the stratified randomisation was successful and there were no statistically significant differences in any of the dependent variables at baseline (2012). We perform balance tests on the whole sample of available taxpayers, using only

those with positive tax, and separately by SRM users and non-users. In all cases, both tax and turnover were balanced at baseline, both for VAT and income tax (see Appendix table A3).

We estimate the following equation:

$$Tax_i = \alpha + \beta Treatment + \gamma X_i + u_i \tag{3}$$

X_i is a vector of firm-specific control variables, including firms' sector, organisational form (corporation or unincorporated business) and location (tax centres). The effect of the letter is captured by the coefficient β. As in the diff-in-diff analysis, the dependent variable is either tax due, in which case we use OLS for estimation, or the probability of reporting positive tax, in which case we use a probit model (reporting marginal effects). In both cases, the dependent variable is tax in the post-treatment period (2013).

4 Results

4.1 Diff-in-diff results

Table 3 reports results for the diff-in-diff estimation for VAT. It includes only taxpayers with positive net VAT, and reports the effect of the SRM on the log of three VAT variables. The first one is VAT sales, which are shown to increase by 87 per cent as a result of adoption of the SRM. However, as shown in the second column, reported inputs also increase and proportionally more than sales – about 150 per cent. The net result on VAT is still positive, with a 70 per cent increase in net VAT, but less than proportional compared to VAT sales. All these effects are statistically significant and economically large. Some degree of cost adjustment is expected when reported sales increase. That is due to firms' desire to appear small, therefore underreporting both sales and costs (Carrillo et al., 2017). However, in our case, costs increase almost twice as much as sales. Besides eroding the potential revenue benefits of the SRM, this large increase in inputs, compared to sales, is hard to reconcile with a real economic response. It may be that taxpayers make large offsetting adjustments to costs because they know this is a harder margin for ERCA to verify. This would be consistent with findings in the literature, where taxpayers are especially found to inflate

less verifiable expenses (Mascagni and Mengistu, 2016; Slemrod et al., 2015; Carrillo et al., 2017).

We find a similar result for income taxes, as reported in Table 4. While SRM adoption increases reported turnover by about 27 per cent, taxpayers also respond by almost doubling the value of deductible expenses (93 per cent), thus managing to increase tax liability by only 18 per cent. Although the taxpayers' response is qualitatively similar for VAT and income tax, it is smaller in magnitude for the latter. This may be due to the self-enforcing nature of the VAT (Pomeranz, 2015) or to the (mis-)perception that SRMs were mostly introduced for VAT enforcement purposes. However, once we estimate the income tax effects only on VAT-registered taxpayers, the effect sizes between VAT and income tax are much more aligned (see Table A4 in Appendix). This suggests that any difference in magnitude is due more to the sample than to the tax type. In other words, VAT payers respond more to the introduction of SRMs. One possible explanation for the larger response is that, for VAT refund in Ethiopia, tax authorities check the financial statements of both transacting parties, which creates a paper trail – increasing the probability of detection. The SRM, by making the process of checking transactions much easier, increases the detection probability.

Table 3: The effect of SRMs on VAT declaration: annual data and positive VAT declaration

	VAT_sale	VAT_input	Net_VAT
SRM	0.870***	1.493***	0.707***
	[0.027]	[0.069]	[0.032]
Year fixed effect	Yes	Yes	Yes
Firm fixed effects	yes	yes	yes
Number of observations	26938	26938	26938
Number of firms	13566	13566	13566
Adjusted R^2	.2	.1	.12

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5 looks at the SRMs' effect on the accuracy of taxpayer self-reports, which has

Table 4: The effect of SRMs on profit tax declaration: positive tax

	sale	cost	proftax
SRM	0.275***	0.935***	0.177***
	[0.012]	[0.042]	[0.014]
year fixed effect	Yes	Yes	Yes
Firm fixed effects			
Number of observations	77256	77256	77256
Number of firms	30493	30493	30493
Adjusted R^2	.14	.11	.065

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

been shown to be relatively poor in Ethiopia (see section 3.2.1). Both in the case of turnover and costs, the absolute amount of the discrepancy more than halves as a result of SRM adoption (respectively, 55 per cent and 60 per cent).[21] This effect may be related to the availability of more accurate business records, thanks to the SRM, or to an increased perception of detection for misreporting. Since the main role of the machine is to report data directly to ERCA and to facilitate its use for tax enforcement, it is indeed a rational response for taxpayers to improve the accuracy of their own reports. Importantly, these results suggest that discrepancies are related to non-compliance, rather than being due to any legitimate reason.

Finally, we look at the probability of reporting positive amounts of turnover and tax liability, as well as costs, this time including all firms. Table 6 shows a positive effect of the SRM on all these variables for VAT. Table 7 shows increased probability of declaring positive amounts of turnover and tax liability, as well as costs, for income tax purposes for all income tax payers. However, the effects are relatively modest. While the probability of reporting positive VAT sales and tax increases by 14 per cent and 15 per cent respectively, the same figures for income tax are 8 per cent and 5 per cent. But these figures increase

[21] Note that costs for income tax and VAT inputs should, in principle, coincide because we are using total costs from the VAT declaration, not only VAT deductible ones.

Table 5: The effect of SRM on accuracy

	Sale Discrepancy	Cost Discreancy
SRM	-0.524***	-0.578***
	[0.065]	[0.080]
Year fixed effect	yes	yes
Firm fixed effects	yes	yes
Number of observations	42980	42980
Number of firms	18829	18829
Adjusted R^2	.0071	.0054

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

to 13 per cent and 12 per cent when we look at income tax variables for the sub-sample of VAT registered firms (see Table A5 in Appendix). This more modest response is consistent with the persistence of nil-filing observed elsewhere in the literature (Mascagni et al., 2017; Shimeles et al., 2017).

Table 6: The effect of SRMs on the probability of positive VAT declaration: annual data

	VAT sale>0	VAT input>0	VAT declared >0
SRM	0.137***	0.122***	0.154***
	[0.006]	[0.006]	[0.008]
Year fixed effect	yes	yes	yes
Firm fixed effects	yes	yes	yes
Number of observations	42980	42980	42980
Number of firms	18829	18829	18829
Adjusted R^2	.04	.033	.041

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: The effect of SRMs on the probability of declaring positive amounts

	sale>0	cost>0	proftax>0
SRM	0.080***	0.107***	0.046***
	[0.003]	[0.004]	[0.004]
year fixed effect	yes	yes	yes
Firm fixed effects	yes	yes	yes
Number of observations	97566	97566	97566
Number of firms	35469	35469	35469
Adjusted R^2	.022	.079	.0037

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.2 RCT results

Despite the offsetting cost adjustments, the diff-in-diff results still show large increases in tax revenue as a result of SRM introduction. Are these results attributable to compliance, or to real increases in economic activity? The results in Table 8 strongly suggest it is the former. We start from income taxes, as this was the main focus of the original study (Shimeles et al., 2017). Overall, the letter treatment has a positive and significant effect on the income tax liability, which is attributable to an increase in compliance of about 16 per cent compared to the control group. When we look at the sub-group of non-SRM users (column 2), this effect almost doubles, to over 30 per cent, and becomes more significant.[22] Instead, for those who already adopted the machine, the letter has no effect, consistent with the idea that they may already be more compliant. Column 4 confirms the same result using interaction terms between the treatment and the SRM. Consistent with the previous results, it shows that the treatment and the interaction term cancel each other out, thus resulting in no treatment effect for SRM users.

[22] We have also tried to estimate the treatment effects separately for persuasion and deterrence, for non-users. Both treatment effects are significant but not statistically different from each other, as in the original study (Shimeles et al., 2017). These results are available upon request.

Table 8: Treatment effects: main results

	(1)	(2)	(3)	(4)
	All	non-SRM	SRM	All
Treatment	0.160*	0.304**	-0.017	0.258*
	(0.078)	(0.117)	(0.100)	(0.101)
SRM user				0.816***
				(0.105)
Treat*SRM				-0.159*
				(0.076)
Observations	2722	1248	1474	2722

Standard errors in parentheses

Dependent variable: log of income tax due.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Turning to VAT, Table 9 shows that the treatment effects disappear when we consider VAT net tax as an outcome variable. Despite being non-significant, the results in column 2 are qualitatively the same as the ones from Table 8, showing that the interaction between treatment and SRM reduces the treatment effect. The results on sub-groups, while being very imprecisely estimated, are also in line with the findings on income tax – with only non-users having a positive (though non-significant) treatment effect. The weaker results on VAT may be due to two reasons. First, the sample size is substantially smaller, especially for the sub-groups of columns 3 and 4. Second, the letter was specifically targeted at increasing income tax compliance, both in terms of its contents (see Appendix) and its timing (see section 3.1.2). These reasons may largely explain the lack of a treatment effect on VAT. The results of Table 8 instead are largely confirmed when looking at income tax variables for the sub-group of VAT registered taxpayers (see Appendix Table A6).

These RCT results, combined with the increase in revenue highlighted in the diff-in-diff analysis of section 4.1, are highly suggestive that the main mechanism behind the observed increase in tax is an improvement in taxpayer compliance. However, the caveat discussed in

Table 9: Treatment effects on VAT net tax

	(1)	(2)	(3)	(4)
	all VAT TP	all VAT TP	VAT+non-SRM	VAT+SRM
Treatment	0.063	0.131	-0.039	-0.252
	(0.077)	(0.111)	(0.237)	(0.133)
SRM user		0.255*		
		(0.107)		
Treat*SRM		-0.075		
		(0.077)		
Observations	2586	2586	612	1231

Standard errors in parentheses. Dependent variable: log of VAT net tax due.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

section 3.3.2 still holds: we cannot rule out with certainty that adopters and non-adopters would have reacted to the letter in the same way in absence of the machine.

Finally, as in the diff-in-diff, we estimate the treatment effects on the probability of declaring positive. Although the letter generally increases the probability of declaring positive income tax for all taxpayers, this effect is larger for taxpayers who have not yet adopted the machine (see Table 10). This finding essentially mirrors the results of Table 6. Consistent with the results on positive tax, there are no treatment effects on the probability of declaring positive VAT (see Appendix Table A7).

5 Robustness

We check the robustness of our results in two ways. First, we re-estimate all our equations using monthly data for VAT, as more disaggregated data is available in this case. The resulting coefficients are smaller but qualitatively similar: turnover increases, but reported inputs increase by more, thus resulting in only a partial increase in tax (see Appendix

Table 10: Treatment effects: probability of declaring tax > 0

	(1) All	(2) non-SRM	(3) SRM	(4) All
Treatment	0.066***	0.079**	0.055*	0.072***
	(0.016)	(0.024)	(0.022)	(0.020)
SRM user				0.079***
				(0.022)
Treat*SRM				-0.012
				(0.016)
Observations	4051	1867	2171	4051

Standard errors in parentheses. Dependent variable: probability of income tax > 0.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A8).

Second, as far as the RCT is concerned, we check whether our results are robust to taking firm size into account. Appendix Table 8 shows that SRM users have higher tax liability. This is confirmed by t-tests of equal means of turnover and tax, showing that SRM users are significantly larger, although this difference is not large in magnitude. We check whether the RCT results are affected by size, by re-estimating the same equations without the top quartile in terms of turnover and adding controls for size. The results, reported in Appendix Table A9, are qualitatively similar. Although in one specification there seems to be a treatment effect also for SRM taxpayers, this effect is still smaller than for non-users, as expected.

6 Conclusions

This paper presents new evidence on the effect of ICTs on tax compliance, in a literature that so far has left this topic largely unexplored (section 1). Our results contribute to

existing knowledge by showing that, although ICTs are highly effective in increasing tax compliance and revenue in Ethiopia, they are still severely limited by low state capacity and by taxpayers' strategic responses. More specifically, our analysis highlights three sets of results, which can be summarised as follows.

First, the potential of existing technology and data systems remains largely untapped in Ethiopia. Most administrative data collected by ERCA is already available in digital format, and should therefore allow for easy cross-checks. However, we show that discrepancies in taxpayers' records are both frequent and large (section 3.2.1). The positive effect of SRMs on accuracy (section 4.1) suggests that these discrepancies are related to non-compliance. Prior to the machine's adoption, taxpayers may think ERCA is unable to cross-check their records and detect mis-reporting. This perception is corrected after the introduction of the machine, when accuracy improves.

Second, despite these limitations, the introduction of SRMs led to large increases in tax revenue for both income taxes and VAT. These effects are both statistically significant and economically large (section 4.1). However, tax revenue does not increase to its full potential, as taxpayers simultaneously adjust both turnover and expenses. Since the latter increases more than turnover, for both income tax and VAT, the revenue gains are lower than they would be without this (over-) adjustment.

Third, our RCT results show that the observed SRM effects are largely due to increases in compliance, rather than real economic activity. Firms that have already adopted the machine generally do not respond to further increases in the probability of being detected, while non-adopters respond significantly (section 4.2). These results are highly suggestive that SRM users are already more compliant due to the machine. The RCT serves both as a mechanism experiment, confirming compliance as the key mechanism, and as an indirect test of the effectiveness of SRMs – notwithstanding the caveats highlighted in sections 3.3.2 and 4.2.

Taken together, our results show that, while ICTs are highly effective, they are no silver bullet in contexts of limited state capacity. Realising their full potential requires substantial investments to boost the revenue administration's capacity to use the available data and ICT systems. Taxpayers still respond to increases in enforcement by shifting evasion to less enforced margins, like deductible expenses. Tackling tax evasion therefore

requires a combination of technological innovation and measures to boost capacity and traditional enforcement.

Appendices

A Text of letters in English

Encouragement Letter

To . . .

Subject: Income Tax Declaration

The Growth and Transformation Plan sets out a vision of becoming a middle-income country, eradicating poverty and ensuring fair and equitable benefits to society. Achieving this vision requires significant financial resources. The private sector engaged in agriculture, manufacturing, foreign trade and infrastructure is highly expected to contribute to this objective. Citizens, investors, development partners and the government need to work in tandem.

The Ethiopian Revenue and Customs Authority believes that you, as a taxpayer, have a big role to play in realising these national goals. Therefore, ERCA cordially asks you to pay your taxes with pride, and support the growth and transformation plan.

ERCA would like to take this opportunity to inform you that it has the mandate to reward outstanding taxpayers as per the income tax proclamation No 286/2002, article 85. Moreover, it has a system of express tax refund and, for those engaged in import-export business, it provides fast and efficient document inspection and examination of declaration. ERCA, therefore, invites you to take this opportunity to play an exemplary role in paying your taxes.

Regards,
Branch Manager

Deterrence Letter

To . . .

Notice: Income Tax Declaration

Taxpayers play a vital role in enabling the effective implementation of extensive development goals put forward by the government. It is therefore the duty of taxpayers to declare their income accurately and in a timely basis. We believe that you are paying your taxes according to the laws of the country.

The Ethiopian Revenue and Customs Authority has carried out an assessment on the tax declaration of taxpayers. The study has revealed that some taxpayers engage in concealing their activities to avoid paying the true tax amount. The Authority has been informing taxpayers of the presence of such discrepancies and encouraging them to make improvements.

As you are well aware, the 2013/14 fiscal year will end in one month's time and the Ethiopian Revenue and Customs Authority has finalised preparations to play its role in the tax declaration process in an effective manner. As per proclamation no. 286/2002, article 38, the authority can audit taxpayers as needed. With the knowledge that you could be one of the taxpayers the Authority will audit, we advise you to declare your taxes truthfully and on time.

It is the duty of taxpayer to pay the true tax and on time as a citizen. If a taxpayer was found to understate his/her tax liability, she/he will be subjected to a penalty of 10 to 50 percent of the understated tax amount as per proclamation no. 286/2002, article 87. Those who have submitted false documents will face criminal charges, which will result in imprisonment of up to 15 years and 100,000 birr monetary penalty. Given these severe consequences, the Authority warns taxpayers to declare their taxes truthfully to avoid penalties and being held accountable for false declarations.

Regards,
Branch Manager

B Summary statistics

Table A1: Summary statistics of VAT reports for the selected sample

	2010			2011			2012			2013			2014			Total		
	mean	SD	Obs.	mean	SD	Obs.	mean	SD	Obs.	mean	SD	Obs.	mean	SD	Obs.	mean	SD	Obs.
TotalsaleVAT	132.89	764	3226	142.08	829	5411	131.27	746	9401	141.74	785	11931	183.36	1147	13011	151.43	906	42980
VATablesale	123.93	706	3226	134.34	794	5411	123.89	714	9401	133.59	738	11931	170.37	1070	13011	141.97	852	42980
VATonsale	18.59	106	3226	20.15	119	5411	18.58	107	9401	20.04	111	11931	25.56	160	13011	21.30	128	42980
TotalinputVAT	109.14	566	3226	110.03	632	5411	100.22	548	9401	114.78	783	11931	153.78	1293	13011	122.38	904	42980
VATableinput	97.49	462	3226	100.93	579	5411	94.12	530	9401	108.42	767	11931	146.89	1272	13011	115.18	879	42980
NetVAT	6.71	67	3226	7.07	64	5411	6.70	59	9401	6.74	49	11931	7.62	56	13011	7.04	57	42980

All mean and standard deviation are expressed in 10000 birr.

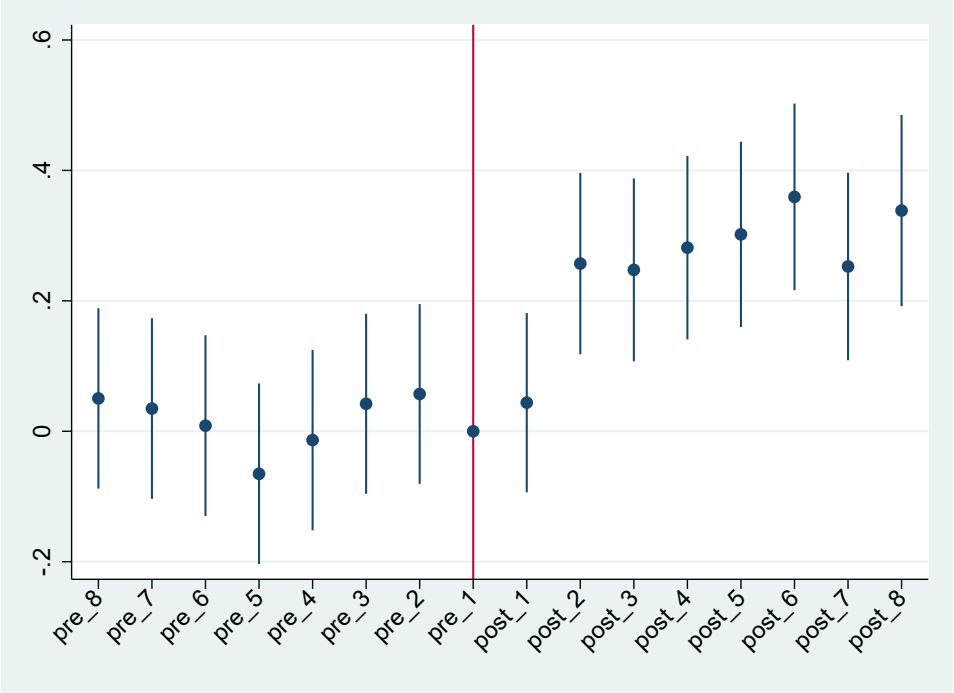
Table A2: Summary statistics of income tax reports for the selected sample

	2011			2012			2013			2014			Total		
	mean	SD	Obs.	mean	SD	Obs.	mean	SD	Obs.	mean	SD	Obs.	mean	SD	Obs.
AnnualTurnOver	31.12	315	15749	43.93	294	23002	59.03	387	29001	91.96	755	29814	61.03	506	97566
Totalcost	27.97	320	15749	42.01	319	23002	56.66	389	29001	105.31	2989	29814	63.44	1678	97566
Profittax	1.09	10	15749	1.27	8	23002	1.77	20	29001	3.04	98	29814	1.93	56	97566

All mean and standard deviation are expressed in 10000 birr.

C Parallel trends test using monthly data for VAT

Figure 3: Dynamic response of VAT declarations for machine adoption - monthly data



D Balance tests (RCT)

Table A3: Letter experiment: baseline summary statistics and balance of randomisation

	All			SRM			Non-SRM		
	Control (mean)	Diff	Obs.	Control (mean)	Diff	Obs.	Control (mean)	Diff	Obs.
Log profit tax	6.596	-0.139	3820	6.726	-0.265	2233	6.426	0.069	1587
		(0.16)			(0.22)			(0.23)	
Log profit tax (positive taxpayers)	9.464	-0.089	2682	9.960	-0.067	1540	8.862	-0.012	1142
		(0.09)			(0.12)			(0.13)	
Log of turnover	10.826	-0.4213	3820	11.43	-0.472	2233	10.038	-0.247	1587
		(0.17)			(0.22)			(0.27)	
Log of turnover (positive taxpayers)	13.16	0.0346	2682	13.663	0.0804	1540	12.55	0.0778	1142
		(0.09)			(0.11)			(0.13)	
Log net VAT	9.3625	0.1522	2841	9.3822	0.0502	2018	9.3211	0.4369	823
		(0.18)			(0.21)			(0.35)	
Log net VAT (positive taxpayers)	11.23	0.0483	2350	11.225	0.0636	1687	11.241	0.0013	663
		(0.09)			(0.10)			(0.17)	
Log sales VAT	11.545	0.1149	2841	11.585	0.0085	2018	11.46	0.4235	823
		(0.21)			(0.25)			(0.40)	
Log sales VAT (positive taxpayers)	13.7	-0.0049	2350	13.708	0.0293	1687	13.682	-0.0957	663
		(0.09)			(0.10)			(0.17)	
Services other than wholesale (%) - profit taxpayers	0.3309	-0.0285	3820	0.3234	-0.0122	2233	0.3406	-0.0537	1587
		(0.02)			(0.02)			(0.03)	
Services other than wholesale (%) - VAT taxpayers	0.3267	0.0429	2841	0.3195	0.0341	2018	0.3416	0.0622	823
		(0.02)			(0.02)			(0.03)	

E Additional diff-in-diff results

Table A4: The effect of SRMs on profit tax declaration: VAT registered an positive declaring firms

	sale	cost	proftax
SRM	0.815^{***}	1.474^{***}	0.746^{***}
	[0.023]	[0.060]	[0.030]
year fixed effect	Yes	Yes	Yes
Firm fixed effects			
Number of observations	28875	28875	28875
Number of firms	13880	13880	13880
Adjusted R^2	.24	.12	.12

Standard errors in brackets

"The results in column(1)-(3) are for all firms including those that declare zero sales."

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A5: The effect of SRMs on the probability of declaring positive amount: VAT taxpayers

	sale>0	cost>0	proftax>0
SRM	0.129^{***}	0.122^{***}	0.117^{***}
	[0.005]	[0.005]	[0.007]
year fixed effect	Yes	Yes	Yes
Firm fixed effects			
Number of observations	42980	42980	42980
Number of firms	18829	18829	18829
Adjusted R^2	.033	.033	.018

Standard errors in brackets

"The results in column(1)-(3) are for all firms including those that declare zero sales."

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

F Additional RCT results

Table A6: Treatment effects: VAT sub-group

	(1)	(2)	(3)	(4)
	all VAT TP	all VAT TP	VAT+non-SRM	VAT+SRM
Treatment	0.154	0.356**	0.374*	0.030
	(0.092)	(0.134)	(0.170)	(0.108)
SRM user		0.205		
		(0.131)		
Treat*SRM		-0.206*		
		(0.092)		
Observations	1843	1843	612	1231

Standard errors in parentheses

Dependent variable: log of income tax due.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7: Treatment effects: probability of declaring VAT net tax > 0

	(1)	(2)	(3)	(4)
	All	non-SRM	SRM	All
Treatment	0.008	0.049	-0.019	0.028
	(0.014)	(0.025)	(0.017)	(0.019)
SRM user				0.088***
				(0.019)
Treat*SRM				-0.026
				(0.014)
Observations	3138	1088	2042	3138

Standard errors in parentheses

Dependent variable: probability of VAT net tax¿0.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

G Additional robustness results

Table A8: SRM effect on positive VAT declaring and the probability of declaring positive VAT: monthly data

	VAT_sale	VAT_input	Net_VAT	VAT sale>0	VAT input>0	VAT declared >0
SRM	0.503***	0.612***	0.433***	0.036***	0.070***	0.016***
	[0.024]	[0.054]	[0.025]	[0.006]	[0.006]	[0.006]
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	yes	yes	yes	yes	yes	yes
Number of observations	71209	71209	71209	137320	137320	137320
Number of firms	3523	3523	3523	3631	3631	3631
Adjusted R^2	.15	.041	.1	.016	.016	.0099

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A9: Treatment effects: results with size

	(1) non-SRM	(2) SRM	(3) All	(4) non-SRM	(5) SRM	(6) All
Treatment	0.259*	-0.111	0.161	0.207**	0.081	0.231***
	(0.106)	(0.098)	(0.093)	(0.079)	(0.071)	(0.067)
SRM user			0.816***			0.243***
			(0.099)			(0.073)
Treat*SRM			-0.106			-0.073
			(0.072)			(0.052)
Observations	972	928	1900	972	1474	2722

Standard errors in parentheses

Dependent variable: log of income tax due.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

Ali, M., A. Shifa, A. Shimeles, and F. Woldeyes (2015). Information technology and fiscal capacity in a developing country: Evidence from Ethiopia. Working paper 31, International Centre for Taxation and Development.

Almunia, M., F. Gerard, J. Hjort, J. Knebelmann, D. Nakyambadde, C. Raisaro, and L. Tian (2017). An analysis of discrepancies in tax declarations submitted under value-added tax in Uganda. Final report S-43312-UGA-1, International Growth Centre.

Besley, T. and T. Perrson (2013). Taxation and development. In A. Auerbach, R. Chetty, M. Feldstein, and E. Saez (Eds.), *Handbook of Public Economics*. Elsevier.

Brockmeyer, A., M. Hernandez, S. Kettle, and S. Smith (2016). Casting a wider tax net: Experimental evidence from Costa Rica. Technical report, mimeo.

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.

Del Carpio, L. (2014). Are the neighbors cheating? Evidence from a social norm experiment on property taxes in Peru. Mimeo, INSEAD.

Eissa, N. and A. Zeitlin (2014). Using mobile technologies to increase VAT compliance in Rwanda. mimeo, McCourt School of Public Policy, Georgetown University.

Gordon, R. and W. Li (2009). Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics 93*(7), 855 – 866.

Kangave, J., S. Nakato, R. Waiswa, and P. L. Zzimbe (2016). Boosting Revenue Collection through Taxing High Net Worth Individuals: The Case of Uganda. *ICTD Working Paper 45*.

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), 651–692.

Kleven, H. J., C. T. Kreiner, and E. Saez (2009, August). Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries. *NBER Working Paper 15218*.

Kopczuk, W. and J. Slemrod (2006, May). Putting firms into optimal tax theory. *The American Economic Review 96*(2), 130–134.

Lewis-Faupel, S., Y. Neggers, B. A. Olken, and R. Pande (2016). Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia. *American Economic Journal: Economic Policy 8*(3), 258–283.

Ludwig, J., J. R. Kling, and S. Mullainathan (2011, Summer). Mechanism Experiments and Policy Evaluations. *Journal of Economic Perspectives 25*(3), 17–38.

Manoli, D. S. and N. Turner (2014). Nudges and learning: Evidence from informational interventions for low-income taxpayers. Working paper 20718, NBER.

Mascagni, G. (2017). From the lab to the field: A review of tax experimemts. *Journal of Economic Surveys*.

Mascagni, G. and A. Mengistu (2016). The corporate tax burden in Ethiopia: Evidence from anonymised tax returns. Working paper 48, International Centre for Taxation and Development.

Mascagni, G., C. Nell, and N. Monkam (2017). One size does not fit all: A field experiment on drivers of compliance and delivery methods in Rwanda. Working paper 58, International Centre for Taxation and Development.

Muralidharan, K., P. Niehaus, and S. Sukhtankar (2014, October). Building State Capacity: Evidence from Biometric Smartcards in India. *NBER Working Paper 19999*.

Naritomi, J. (2016, May). *Consumers as Tax Auditors*. Ph. D. thesis, London School of Economics.

Okunogbe, O. and V. Pouliquen (2017). Technology, Taxation and Corruption: Experimental Evidence from the Introduction of Electronic Tax Filing. *Forthcoming*.

Olken, B. A. and R. Pande (2011, September). Corruption in Developing Countries. *NBER Working Paper 17398*.

Ortega, D. and C. Scartascini (2016). Don't blame the messenger: A field experiment on delivery methods for increasing tax compliance. mimeo, Inter-American Development Bank.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the Value Added Tax. *American Economic Review 105*(8), 2539–69.

Shimeles, A., D. Z. Gurara, and F. Woldeyes (2017). Taxman's dilemma: Coercion or persuasion? evidence from a randomized field experiment in Ethiopia. *American Economic Review 107*(5), 420–24.

Slemrod, J., B. Collins, J. Hoopes, D. Reck, and M. Sebastiani (2015). Does credit-card information reporting improve small-business tax compliance? NBER Working Paper 21412, NBER.

World Bank (2015). Tax Compliance Cost Burden and Tax Perceptions Survey in Ethiopia. Technical report, World Bank Group.

World Bank (2016). World development report 2016: Digital dividends. Technical report, Washington, DC: World Bank.

International Centre for Tax and Development
at the Institute of Development Studies
Brighton BN1 9RE, UK
T: +44 (0)1273 606261
F: +44 (0)1273 621202
E: info@ictd.ac
www.ictd.ac