ICTD

Working Paper 140

Mandating Digital Tax Tools as a Response to Covid: Evidence from Eswatini

Fabrizio Santoro, Razan Amine and Tanele Magongo

May 2022

BILL & MELINDA GATES foundation
Norad
UKaid from the British people
institute of development studies

ICTD Working Paper 140

Mandating Digital Tax Tools as a Response to Covid: Evidence from Eswatini

Fabrizio Santoro, Razan Amine and Tanele Magongo

May 2022

Mandating Digital Tax Tools as a Response to Covid: Evidence from Eswatini

Fabrizio Santoro, Razan Amine and Tanele Magongo

ICTD Working Paper 140
First published by the Institute of Development Studies in May 2022

ISBN: 978-1-78118-992-4
DOI: 10.19088/ICTD.2022.006

Available from:
The International Centre for Tax and Development at the Institute of Development Studies, Brighton BN1 9RE, UK
Tel: +44 (0) 1273 606261 Fax: +44 (0) 1273 621202
E-mail: info@ictd.ac.uk
Web: ictd.ac/publication
Twitter: twitter.com/ICTDTax
Facebook: facebook.com/ICTDTax
LinkedIn: linkedin.com/company/ICTDTax

IDS is a charitable company limited by guarantee and registered in England
Charity Registration Number 306371
Charitable Company Number 877338

Mandating Digital Tax Tools as a Response to Covid: Evidence from Eswatini

Fabrizio Santoro, Razan Amine and Tanele Magongo

Summary

Many tax authorities changed the mode of interacting with taxpayers from physical to online as a response to the Covid-19 pandemic, to diminish the spread of the virus. Eswatini, the country under study, mandated the use of online tax filing through the e-Tax system for all income tax payers, coupled with a zero-cash-handling policy for tax payment. By means of a difference-in-difference (DID) strategy, reinforced by a propensity score matching (PSM), this paper offers an impact evaluation of the mandate on taxpayer filing and payment behaviour.

We present three sets of results. First, we describe which firms are most likely to register for e-Tax – mostly large firms and those in the primary and tertiary sectors. Second, we show that e-Tax uptake significantly improves filing behaviour, as well as payment behaviour. E-Tax registered taxpayers are less likely to file nil (by 60 per cent), declare more turnover and taxable income, and are 70 per cent more likely to pay conditional on filing. Third, we shed light on the mechanisms behind our main findings, showing that the technology improved accuracy and reduced compliance costs. E-Tax-registered treated taxpayers are more likely to file on time, file for VAT, report more accurately, and, on the payment side, to pay their liabilities in full.

Keywords: tax administration; tax compliance; technology adoption.

Fabrizio Santoro is a Research Fellow at the Institute of Development Studies and at the International Centre for Tax and Development, as well as Research Lead of the DIGITAX programme.

Razan Amine is a Researcher of the DIGITAX programme (ICTD).

Tanele Magongo is the Manager Strategy at the Eswatini Revenue Service. She works on facilitating, monitoring and implementation of ERS' strategic plan and annual organisational plan.

Contents

Acknowledgements

We thank the Eswatini Revenue Service (ERS) for approving the study and providing support throughout its implementation. We are particularly grateful to Edward Groening for helpful comments. We also thank two anonymous reviewers who helped strengthen the study. This paper was supported by the Gates Foundation through the International Centre for Tax and Development DIGITAX programme.

Acronyms

CIT	Corporate income tax
DID	Difference-in-difference
EFD	Electronic fiscal device
EFT	Electronic funds transfer
ERS	Eswatini Revenue Service
GMTI	GovTech Maturity Index
IT	Information technology
ITAS	Integrated technology automated system
LIC	Low-income country
LMIC	Low- and middle-income country
PIT	Person income tax
POS	Points of sale
PSM	Propensity score matching
TIN	Taxpayer identification number
VAT	Value added tax

1 Introduction

The spread of the Covid-19 pandemic has increasingly large fiscal implications, especially in low- and middle-income countries (LMICs) and Africa. There is a significant increase in the expenditure needed to finance health systems and provide social protection to vulnerable groups, and less tax revenue as a result of the global economic slowdown (Arndt et al. 2020; Bachas et al. 2020; Aragie et al. 2021; Mascagni and Lees 2021). The pandemic is also dramatically changing the way that African taxpayers and revenue authorities interact – going from intense physical contact, to a remote online relationship. The increased use of information technology (IT) in tax administration due to the pandemic is taking place in a context where there has been increased adoption of IT solutions over the last decade in Africa (Okunogbe and Santoro 2022; ATAF 2021). In Eswatini – the country under study – this shift to technology took a radical turn in September 2020, due to a mandate to use only online tools for paying taxes.

These developments present LMICs with immediate, and in some ways unprecedented, questions around taxation. First, how to ensure taxpayer compliance through remote interactions and IT solutions? Should governments turn to technology-enabled options, such as e-filing[1] and e-payment[2], as a solution for all taxpayers to remit their taxes without physically interacting with the tax administration? Second, how can these options be implemented smoothly, and be part of a broader journey to digitization – bearing in mind technical considerations, equity considerations and broader political feasibility? Relatedly, how can the tax administration benefit from higher data quality from digital tax filing/payment in its transition towards digitisation?

This paper assesses the impact of the mandate for all income taxpayers in Eswatini to use online filing (e-filing), coupled with a zero-cash-handling policy for tax payments. The e-filing mandate was introduced through legislative reform in September 2020, and applies to the 2020/2021 fiscal year. Manual filing was not accepted in the country from September 2020. All income taxpayers are required to use the online e-Tax platform to process their filing. In conjunction with mandatory e-filing, the tax administration (Eswatini Revenue Service (ERS)) launched a zero-cash-handling policy in April 2021, and it is no longer possible to pay taxes using cash. Instead, digital channels, such as mobile money, credit cards at points of sale (POS) and electronic transfers are accepted.

This paper seeks to address three sets of research questions. First, what is the effect of mandating e-filing on taxpayer filing behaviour – are taxpayers remitting more taxes and filing more accurately? Second, are there any spillover effects on payment behaviour, and the accuracy of tax payments? Third, what are the repercussions in terms of equity of the tax

[1] E-filing refers to the online filing and submission of tax returns, usually on online platforms managed by the revenue authority. As of 2015, 32% of developing countries had introduced e-filing (World Bank 2016). In Uganda, e-filing was introduced in 2009 and by the end of 2011 all returns were filed electronically on the URA web portal (Mayega et al. 2021). In Rwanda, e-filing was launched in 2012 with the aim of assisting clients to file taxes like VAT, PAYE, excise duty and withholding taxes electronically on the Rwanda Revenue Authority (RRA) website directly, without visiting the RRA premises (Ndayisenga and Shukla 2016). In Kenya, e-filing was introduced in 2013 through the adoption of the Integrated Tax Management System (ITMS), initially for VAT collection (Gwaro et al. 2016); this was later replaced with iTax, which enables taxpayers to undertake internet-based registration, filing, paying and status inquiries with real-time monitoring of their accounts (Gwaro et al. 2016). In Nigeria, e-filing was introduced in 2013, but only at the federal level (Mas'ud 2019).

[2] E-payment refers to the payment of taxes through digital financial services, such as: (i) electronic funds transfer (EFT)-based instruments - direct credit and debit transfers that go directly from one account to another; (ii) card-based payment instruments - credit-, debit-, scratch- and charge-card payments that typically involve a physical plastic card, and are initiated, authorised, authenticated, cleared and settled electronically, (iii) electronic money (e-money)-based instruments - online money with payment instructions initiated via the internet, mobile money and prepaid cards - these instruments involve the payer maintaining a pre-funded transaction account with a payment service provider (which could be a bank or another institution).

system, bearing in mind that smaller taxpayers are likely to find it more difficult to comply when using digital solutions?

We rely on data provided by ERS, which granted access to its database through a confidentiality agreement, to address these questions. We make use of detailed information from the taxpayer registry, from which we know which taxpayer registered for e-Tax and when they did so. We also have access to tax filings and payments for 2013-2020 – six years before the mandate and one year immediately after it.

Based on this data, we first map the drivers of e-Tax registration and digital payments uptake. Then, we run a DID analysis on a PSM sample by comparing taxpayers who registered after it became mandatory (treatment group), with those who failed to do so (control group). All early e-Tax registrations are dropped. We are able to adopt this method due to the partial take-up of e-Tax, despite it being mandatory. We measure the impact in terms of filing and payment behaviour – the former is proxied by the amount of tax declared, and the latter by the amount of tax paid. We also consider a range of additional outcomes, such as on-time filing and the accuracy of returns, to investigate further whether e-filing facilitates compliance. We also look at income, expenses and deductions, to understand whether e-filing shapes compliance along these margins. In terms of payments, we finally consider the accuracy of tax payments, as proxied by whether the taxpayer manages to pay their tax liability in full.

We present three sets of findings. First, unsurprisingly, larger firms that are remitting different types of taxes (income tax, PAYE and VAT) are much more likely to register for e-Tax. Also, businesses in primary and tertiary sectors more consistently take up the e-service. Less sophisticated, and less IT-savvy, businesses are much less likely to register for e-Tax. This also applies in 2020-2021, after the mandate, and indicates that registration for e-Tax, while increasing, is not universal even when imposed. A similar pattern is found for the uptake of digital modes of payment – large businesses in the primary sector are more likely to pay taxes using digital means.

Second, in terms of the causal impact of the mandate on tax compliance, the DID analysis indicates that e-Tax uptake significantly improves filing behaviour. Treated taxpayers are almost 60 per cent less likely to file nil, and declare significantly more turnover (+32 per cent) and taxable income (+50 per cent). However, surprisingly, final tax liability remains unchanged. With payment behaviour, we find a strong impact of e-Tax registration on the probability to pay conditional on filing (+70 per cent), as well as on the final tax amount paid (+64 per cent). The divergent impact on tax declared and tax paid is not surprising in a context where filing and payment behaviour are often disconnected, and where revenue authorities often prioritise payment over filing (Santoro and Waiswa forthcoming).

Third, we shed light on the mechanisms behind our main causal estimates. These are:

a) the complementary effect on payment compliance of digital payment means, as compared to cash, is only modest, probably due to e-Tax not directly integrating e-payment options to facilitate payment. We speculate that the zero-cash-handling policy does not have an incremental effect on payment behaviour, which has already been greatly improved by the e-filing mandate;
b) we document a reminder or salience effect that e-Tax is likely to produce, due to the software sending information on deadlines and reminders. Treated taxpayers increase on-time filing (+47 per cent) and the likelihood to file VAT, conditional on having filed for income tax (+135 per cent). The accuracy of reporting – measured by discrepancies between total sales for a given year as reported in income tax and VAT returns – does

not improve. This first finding suggests that e-Tax is helping taxpayers by making the deadline more salient and helping them to meet deadlines;

c) we show that treated taxpayers report more expenses (+67 per cent) and deductions (+42 per cent). This indicates that the system is probably helping taxpayers to track their records – they probably failed to report all expenses with manual filing;
d) treated taxpayers are 53 per cent more likely to pay the full amount of their liability, again suggesting that the main explanatory channel could be better accounting;
e) we show that the largest response to the mandate comes from individual, small taxpayers, and those in urban areas. These findings raise equity concerns, and indicate that the largest improvements in compliance come from those taxpayers who were thought to be less digitally equipped.

This paper contributes to at least two strands of literature. First, we add to the thin but growing evidence around the effectiveness of e-services for tax compliance in developing countries. Evidence on e-filing is rather scarce, and only two studies produced causal estimates of its impact:

1. In Tajikistan, Okunogbe and Pouliquen (2022) study the impact of an information and training intervention on e-filing uptake, and how it shapes compliance costs, tax payments and bribes. The authors show that the combination of training and information interventions induced a 34 percentage points (pp) increase in e-filing adoption relative to the control group, but there was no impact from the training intervention alone. E-filing reduced compliance costs, but increased tax payment heterogeneously only for firms who were more likely to evade at baseline (reduced collusion), and decreased tax payments and bribes for firms who were less likely to evade at baseline (reduced coercion).
2. From Africa, a recent study by Jouste et al. (2021) shows a positive, significant and complementary impact of two interventions – a taxpayer registration and education programme, and an e-filing mandate for presumptive tax – on both the number of small business taxpayers that file tax returns, and the tax revenue.

This paper directly adds to recent evidence from Uganda as it focuses on an e-filing mandate – in Tajikistan, e-filing was only offered to taxpayers. Compared to Jouste et al. (2021), this study explicitly isolates the impact of the mandate (which in Uganda was coupled with taxpayer registration and education activities). It also provides a more nuanced picture in terms of behavioural outcomes, as it looks at both filing behaviour and payment behaviour.

Additional, mostly qualitative, evidence on e-filing shows a mixed picture of its potential:

- Research from Nigeria indicates that important constraints exist for take-up of e-filing, such as taxpayer education and age, IT readiness and use of an external auditor, which are particularly relevant for small taxpayers (Efobi et al. 2019; Mas'ud 2019).
- In Zimbabwe, connectivity is a major constraint for uptake, coupled with lack of technical knowledge and training (Obert et al. 2018).
- Positive, cross-country, evidence comes from World Bank Doing Business data, which shows that, five years after their introduction, e-filing and e-payment averages reduced the time it takes businesses to prepare and pay taxes by 25 per cent (World Bank 2015), with exceptionally positive examples from Belarus (from 987 hours to 183 hours), Costa Rica (from 402 hours to 163 hours) and Kenya (from 432 hours to 202 hours). This links to the experimental findings of Okunogbe and Pouliquen (2022) and Kochanova et al. (2020), which indicate that e-filing significantly reduces the likelihood of being inspected by tax officials in the Central Asian region.
- Kochanova et al. (2020) suggest that the implementation of complementary online tools,

such as e-payment or the zero-cash-handling policy in Eswatini, enhances the impact of e-filing.
- Yilmaz and Coolidge (2013) suggest that the impact is greater when adoption is mandated, as in Nepal and South Africa at the time of the study, or as these mandates have been introduced in several LMICs since the pandemic.

This paper attempts to provide more quantitative evidence around some of the aspects above, especially in terms of understanding the complementary effect of the zero-cash-handling policy for tax payments, and measuring the impact on compliance costs and accurate reporting.

More broadly, we add to the policy debate around the potential of technology for tax administration in developing countries (World Bank 2016; Okunogbe and Santoro 2022). Increasing evidence has been produced on the benefits of electronic fiscal devices (EFDs), e-registration and third-party data reporting. Technology can yield large revenue gains, especially in the case of EFDs – as shown in Eissa and Zeitlin (2014), Ali et al. (2015) and Mascagni et al. (2021). However it suffers from similar barriers to uptake, such as poor connectivity, inadequate equipment, limited knowledge and IT sophistication, as those documented for e-filing (Eilu 2018; Mascagni et al. 2021b).[3] It is still not clear how these IT tools strengthen the core functions of tax administrations, and, in particular, how the wealth of new data produced is eventually used by revenue authorities for identification, facilitation and enforcement (Okunogbe and Santoro 2021). The capacity of tax administrations in developing countries to unlock the full potential of this data, in terms of analytical skills, IT software and accessibility to data, is still limited (Mayega et al. 2019; Ligomeka 2019; Okunogbe and Santoro 2022). We attempt to shed light on these mechanisms through the qualitative research in this paper.

In what follows, we start by describing the Eswatini context, its tax system and the tax e-services under study (section 2). Section 3 presents the research design, while the key results are reported in section 4. Additional evidence shedding light on the mechanisms of the main results is discussed in section 5. Section 6 concludes.

2 Institutional context

Tax system. The Kingdom of Eswatini, an LMIC in Southern Africa,[4] is characterised by a tax-to-GDP ratio of 14.7 per cent, in line with the average in sub-Saharan Africa. This is substantially lower than the 34.2 per cent average in OECD countries and about half of that of Southern Africa. Informality is rampant – the informal sector made up roughly 40 per cent of national income on average in 2005-2015 according to Medina and Schneider (2018), higher than the overall regional average of 32 per cent. Appendix Table A1 compares Eswatini and Southern Africa, and shows that Eswatini performs much worse in terms of key fiscal and governance indicators.

[3] For some preliminary evidence on e-registration, see Knebelmann (2019), Kamara et al. (2020) and Okunogbe (2021). See Carrillo et al (2017), Mittal and Mahajan (2017), Brockmeyer and Hernandez (2019), Shah (2020), Chalendard et al. (2020) for more on the effectiveness of third-party data reporting in developing countries.

[4] The country faces major development challenges, and human development indicators are low compared to other LMICs. Based on the international poverty line of US$1.90 a day and the lower-middle-income poverty line of US$3.20 a day, it is estimated that 38% of the Swazi population live in extreme poverty, and a total of 60.4% are poor overall. This is accompanied by an unemployment rate of 23% in 2018. Health issues are difficult to address, with HIV/AIDS and tuberculosis widespread in the country. As of 2018, Eswatini has the twelfth lowest life expectancy in the world, at 58 years. The population growth rate is 1.2%, with a total population of 1.2 million in 2018 (World Bank 2018).

The Eswatini Revenue Service (ERS) collects direct taxes, about 56 per cent of tax revenue in 2020/21, and indirect taxes, about 44 per cent of revenue (ERS 2021). This study focuses on taxpayers registered for corporate income tax (CIT), a 27.5 per cent tax rate imposed on taxable income from corporate business activities, and personal income tax (PIT), taxing individual income through a progressive structure.[5] In the 2020 financial year, CIT represented 14.4 per cent (17.2 per cent in 2019, before the pandemic hit), and PIT 35.8 per cent (35 per cent in 2019) of total tax revenue collected by the ERS. The on-time filing compliance rates recorded for CIT and PIT were 53.2 per cent (55 per cent in 2019) and 30.3 per cent (39 per cent in 2019) respectively (ERS 2021; ERS 2019)**.**

E-Tax, e-filing mandate and zero cash handling. The electronic platform, e-Tax – the focus of this study – is an IT solution enabling e-filing of all the main taxes in the country (income tax, VAT and PAYE).[6] Figures 1a and 1b show snapshots of the platform's website. It operates as a window directly connected with ERS' integrated technology automated system (ITAS). It works mostly online for CIT payers. An offline solution is offered to PIT payers, given the length of their tax return form and the need to report more information. PIT payers can download the form, fill it offline, and then upload it onto e-Tax when connected to the internet. E-Tax operates on computers, tablets and smartphones.

There are many functionalities of e-Tax. Taxpayers can use it to file returns, access their account information, monitor their filing history, log queries and seek assistance (Figure 1a). An e-payment module is not available yet, and taxpayers have to pay their liabilities outside of e-Tax. Nevertheless, the platform shows the payment history and outstanding balance, in addition to the taxpayers' filing behaviour (Figure 1b).

There are many reasons to believe that this solution significantly reduces the compliance costs for taxpayers:

1. Taxpayers can view overdue and upcoming returns and payments, with deadlines clearly indicated on the webpage (Figure 1b).
2. E-Tax helps store information, such as filing and payment history, and provides a more transparent confirmation and tracking of all compliance actions (Figure 1a).
3. Completing a tax return is greatly simplified. When starting a new submission, taxpayers are carefully guided through reporting the key information.[7] While some entries like turnover and tax liabilities are compulsory, some more specific sub-categories, such as type of deduction or allowances, can be skipped as a group, thanks to automatic skipping patterns if taxpayers report they do not have any deductions or allowances to declare. The time it takes to file should be reduced.
4. Taxpayers can seek assistance at any time by submitting a query through the platform, which is then addressed by a dedicated ERS staff member.[8]

[5] PIT has a maximum marginal rate of 33% and exemptions for income below SZL41,000 (US$2,848).

[6] E-Tax initiative is one of many initiatives implemented by ERS to make voluntary compliance easier and more efficient. It comes under the 'Self-Assessment' project that ERS embarked on in the 2020 tax year, through which taxpayers have the responsibility to correctly calculate and determine their own tax liability. The aim is to achieve efficiencies in revenue collection, and promote correct reporting/declaration.

[7] In most cases, such as when selecting the financial year or type of tax they are filing for, taxpayers can select an option from a drop-down list.

[8] Taxpayers can also email or phone the ERS outside of e-Tax, but most taxpayers prefer using the system (interview with ERS staff, 4 November 2021).

Figure 1a E-Tax platform – snapshot 1

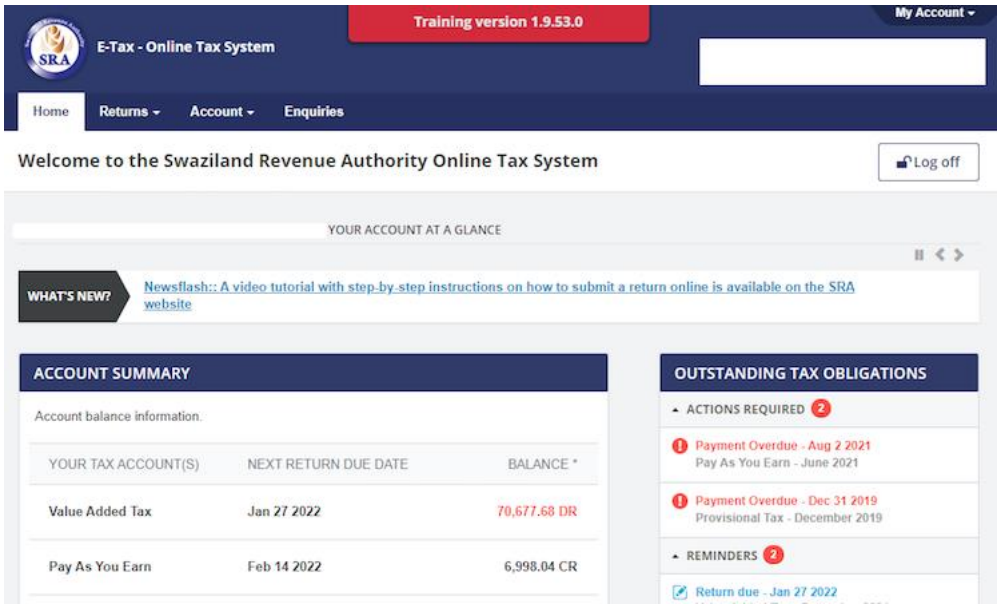
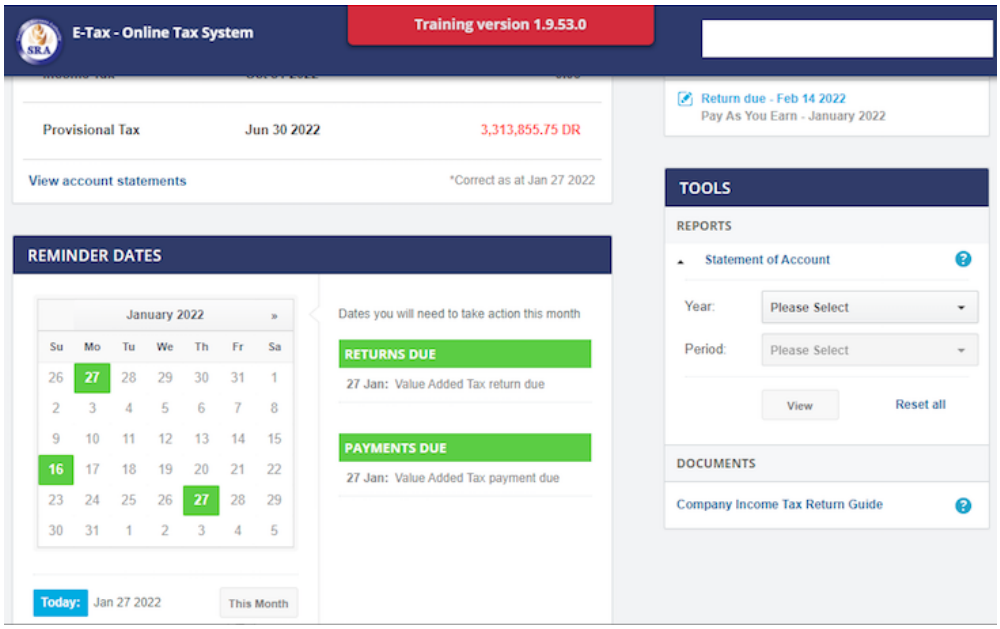


Figure 1b E-Tax platform – snapshot 2



The system was rolled out gradually, starting with large taxpayers, and businesses registered for value added tax (VAT) in March 2014, and adding high net worth individuals in July 2018. E-filing became compulsory for all taxpayer types in September 2020, and the traditional paper-based filing was discontinued. The rationale for this was:

- to reduce physical interaction between taxpayers and tax officials due to Covid-19 social distancing restrictions;
- to reap the benefits e-filing is thought to bring – from improved accuracy and compliance, to reduced corruption and unfairness in the tax system (see section 1).

The ERS put a lot of effort into encouraging e-Tax registration and helping taxpayers to navigate the system. Internet kiosks were set up at tax centres country-wide to provide e-filers with assistance and a stable connection.

As we show below, only a minority (41 per cent) of registered taxpayers moved to filing through the e-Tax system after it became compulsory. Qualitative evidence from interviews with ERS suggests that most small taxpayers lack the equipment, an adequate connection, and the knowledge to navigate e-filing. This is in line with evidence from papers like Eilu (2018) and Mascagni et al. (2021b). Manual practices tend to persist – some taxpayers bring their manual returns (or email them) to ERS, and ask staff to enter them onto e-Tax; others seek assistance from intermediaries, or live in rural areas were internet coverage is poor – which is common across the continent (ATAF 2021).[9] ERS is having difficulty contacting large groups of taxpayers to invite them to register on e-Tax. These taxpayers may have submitted incorrect contact information or have changed it without informing the ERS – a pattern observed in other African countries (Nyanga 2021).

More quantitative evidence from administrative data in Figures 2a and 2b shows a sudden jump in e-Tax registrations after the introduction of the mandate in September 2020; this then slowly declines. Most taxpayers had still not registered for e-Tax as of June 2021 – only 41 per cent of registered entities seem to have done so.

Figure 2a E-Tax registration 2013-2021

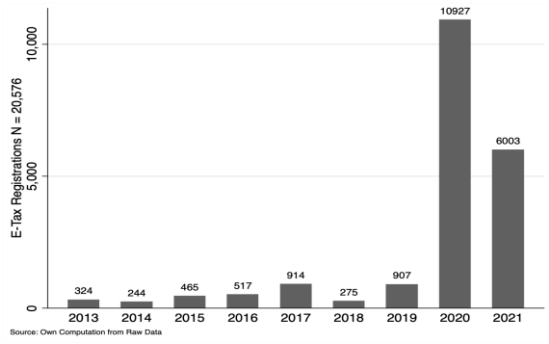
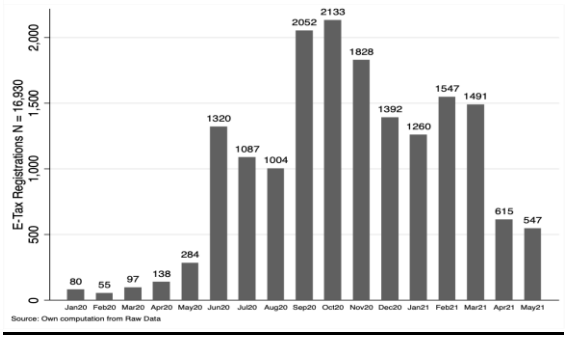


Figure 2b E-Tax registration Jan 2020-May 2021



The ERS made the zero-cash-handling initiative mandatory in April 2021. The only payment options that can now be used are mobile money, point of sale gadgets or electronic bank transfers. In the analysis below, we explore how this initiative at the payment level interacts with the earlier e-filing mandate. A closer look at the administrative data in Figure 3a shows that electronic transfer has been the most popular mode of payment across the nine-year period. It accounted for 91.4 per cent of payments made in 2013, after which there were a substantial number of returns whose payments were being settled in cash. However, taxpayers' dependency on cash has steadily declined over the years, with non-cash modes of payment accounting for almost 93 per cent of payments made in the year preceding the pandemic. Mobile money started to be used for tax payments in 2020, and reached 8 per cent of total payments in 2021.[10] Not surprisingly, digital means of paying taxes, such as electronic transfer and credit card, are used to pay larger amounts of taxes (Figure 3b). Paying by cash is associated with smaller amounts. Interestingly, the tax amounts

[9] In the 2020 ATAF ICT survey, 36% of RAs reported poor internet penetration and power connectivity with a 10-40% monthly power availability (ATAF 2021).

[10] For clarity, we dropped payment by cheque, which accounted for a small share of total payments across the 9-year period (2.6%).

corresponding to mobile money payments follow a similar distribution to that of cash amounts, suggesting higher take-up among smaller taxpayers.

Figure 3a Use of different modes of payments 2013-21

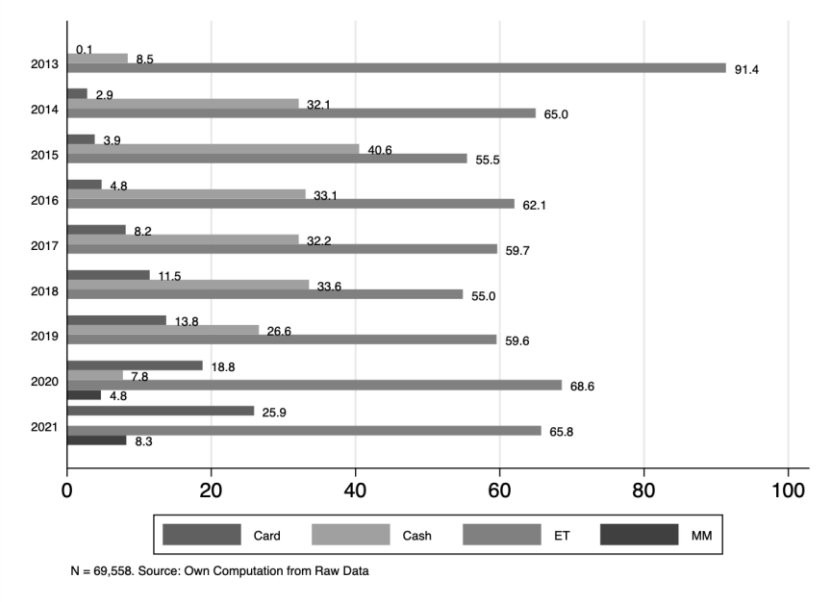
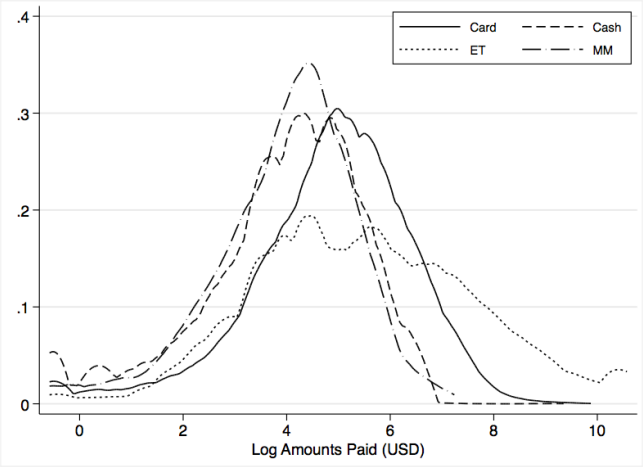


Figure 3b Amount of tax paid by mode of payment 2013-21



The e-filing mandate and zero-cash initiative represent radical IT strategies, and signal a strong drive towards digitisation embracing public institutions – resonating with the broader IT drive in Africa (ATAF 2021). However, according to the latest World Bank GovTech Maturity Index (GMTI), which assigns a 0-1 score and a grouping across four categories (A, B, C, D) as an indicator for the state of a country's public sector digital transformation,[11] Eswatini scores 0.28, and is in group C – this indicates that Eswatini still needs adequate investment to move to a higher level of IT readiness. It also shows that the country performs in line with the LIC average (0.27), but worse than LMICs (0.46).[12]

[11] The GTMI is a composite index based on 48 key indicators on 198 economies based on 4 Indexes: the Core Government Systems Index (CGSI), with 15 indicators; the Public Service Delivery Index (PSDI), with six composite indicators; the Citizen Engagement Index (CEI), with 12 indicators; and the GovTech Enablers Index (GTEI), with 15 indicators. The GTMI is the simple average of the four components measuring the maturity of GovTech focus areas, which are computed as the normalised weighted averages of relevant indicator scores.

[12] According to the index, Eswatini is closer to African countries such as Ethiopia (0.33), Sierra Leone (0.37) and Zimbabwe (0.38), but lags behind with respect to more IT-advanced governments as Rwanda (0.53) and Uganda (0.62).

3 Research design

3.1 Data sources and description

We used different sets of taxpayer-level administrative data from the ERS.[13] First, we had access to the universe of taxpayers registered on the taxpayer registry. This registry is the starting point of our analysis, and was used to identify the study population – more than 60,000 taxpayers.[14] Registrations took place over the period 2007-2021. The registry also contained details of taxpayers and businesses from 26 variables. We used this information as controls in our main specification. This is described in the next section.

The registry indicating who is actually 'treated' by the mandate – the e-Tax registration list – was the second source our data.[15] Crucially, we know the exact date of all e-Tax registrations, which are used in our main specification. As already mentioned in section 2, by the time the data was gathered (May 2021, eight months after the mandate came into force), a minority of 41 per cent taxpayers had registered for e-Tax.

Third, we make use of returns and payments data. We used both CIT (54,450 observations) and PIT (101,667) filings for the period 2013-2021. We focused on the key outcome variables for our analysis on filing behaviour – tax payable, turnover and taxable income – as well as additional entries, such as expenses and deductions for both CIT and PIT payers. A first point to note is that the payment data records fewer payments (71,410 in total) than the sum of CIT (54,450) and PIT returns (101,667). This indicates that a sizeable share of returns are not paid. This also motivates us to explore impact of the mandate on payment behaviour.

We combine all different datasets using anonymised taxpayer identification numbers (TINs). The resulting dataset is a panel data covering all taxpayers in Eswatini, over nine years, from 2013 to 2021.

3.2 Empirical strategy

In this section, we describe the research design and the empirical strategy that we used to estimate impacts. Our main estimation follows a DID strategy,[16] strengthened by a propensity score matching (PSM) approach.[17] Equation (1) below captures the effect of the mandate for all taxpayers in Eswatini to only use e-filing tools for the 2020/2021 financial year. This was implemented in September 2020:

$$Y_{it} = \beta_{0i} + \beta_{1i}(eTax \cdot Post) + \beta_{2i}(eTax) + \beta_{3i}(Post) + \beta_{4i}X_i + \varepsilon_{it} \quad (1)$$

where Y_{it} is a set of outcome variables, e-Tax is a dummy variable that denotes whether the taxpayer registered for e-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $eTax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures our DID coefficient of

[13] All datasets were shared on 20 May 2021.

[14] The taxpayer registry dataset comprises 76,771 observations from a total of 61,638 taxpayers, as taxpayers could register multiple times for the different taxes they remit.

[15] The e-tax registration dataset comprises 31,334 observations corresponding to 20,576 unique registrations - all those taxpayers who registered with e-Tax any time from 2013 (when the service was launched) to May 2021.

[16] The standard DID approach was first introduced by John Snow in 1850, and further explained and critiqued by Bertrand et al. (2004) and Lechner (2011).

[17] PSM was first introduced by Rosenbaum and Rubin (1983).

interest, β_{1i}, which is the effect of being registered for e-Tax and filing after the mandate. To increase statistical precision, we also add time-invariant controls, indicated by X_{it}. The control variables include the firm's sector, location or district, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. $\varepsilon_{\mathrm{it}}$ is an error term.

Our outcome variables originate from both the tax returns and payments dataset. From the returns dataset, we create a first indicator for nil-filing, a dummy variable that indicates whether the firms have filed zero income in the tax returns or not,[18] then the log transformation of reported income, taxable income and final tax liability. From the payments data, the main outcome variables are a dummy variable indicating whether the firm paid taxes at all or not (conditional on having filed), and the log transformation of the tax paid. When the outcome is a binary 0-1 indicator variable, we use a probit model, while for continuous outcomes we recur to a tobit model, more suited for censored – at 0 – variables.[19]

We take particular care in strengthening our DID analysis. Introducing the e-filing mandate in September 2020 with very little notice could be considered a quasi-random shock (section 2). Yet, selection bias into e-filing could severely distort our results. Simple balance tests between those who register early and those who register after the mandate indicate that significant differences exist between the two groups (Appendix Table A2a).[20] Likewise, the same evidence emerges in a more robust regression framework. Figure 4a below reports a coefficient plot from a probit regression of e-Tax registration on a number of different factors, listed on the vertical axis.[21] This shows that a large taxpayer has a strikingly significant likelihood of registering for e-Tax compared to a small one. A similar trend can be found for VAT-registered taxpayers and newly-registered ones. Primary and tertiary sectors are more likely to take up e-filing, compared to the secondary sector. Incorporated taxpayers are less likely to register, mostly due to micro and small entities. Again, taxpayers who did not submit their contact details at the time of registration are much less likely to register for e-Tax.[22] Figure 4b shows similar results when considering digital payments as an outcome.[23] Appendix Tables A5a and A5b show a similar list of predictors of e-Tax registration and digital mode of payment respectively,[24] using a simple machine learning model, a lasso logit model.[25]

---

18 We focus on nil-filing as it is a particularly common and puzzling filing behaviour in Eswatini (Santoro and Mdluli 2019; Santoro et al. 2020), as well as in other African countries (Almunia et al. 2021; Mascagni et al. 2022).

19 For the same reason that motivates us to adopt it here, the tobit model has been used in other studies in tax experiment literature (Slemrod and Weber 2012; Alm and McClellan 2012; Alm et al. 2010).

20 Early adopters are much more likely to be companies than individuals, large in size and operating in the tertiary sector. In addition, those who register early are significantly more likely to submit frequent returns, and more likely to provide an email address or a phone number in their return.

21 Coefficients that cross the vertical 0 line indicate a non-significant correlate, while coefficients that are on the right (left) of the 0 vertical line suggest a positive (negative) correlation with e-Tax registration.

22 This could suggest a lower level of IT sophistication and more hesitancy in disclosing personal contact details (which would make tax evasion more expensive) as two key obstacles to e-Tax registration.

23 Figure 3b shows that the largest predictor of digital payments is being incorporated (company). Being large is also a key determinant. Among the taxpayers across the different sectors of the economy, we find that those from the tertiary and secondary sectors show a lower likelihood of paying tax due electronically. Among other negative determinants, we find having been registered in earlier years and not having a phone number reduces the likelihood of taxpayers filing and paying their tax electronically.

24 The variables PAYE, district Hhohho, and Urban were dropped by the lasso logit model, for negligeable effect on predicting e-Tax registration.

25 The lasso2 coefficients are a result of minimising the residual sum of squares, where the command first runs a full coefficient path for a list of lambda (the tuning parameter chosen by cross-validation), then runs the model selected by EBIC (a type of information criterion). The coefficient of lasso represents the predictive power of each variable, proportional with the magnitude of the coefficient. Post-lasso OLS coefficients are a result of running an OLS regression using the selected predictors.

Figure 4a Correlates with e-Tax registration

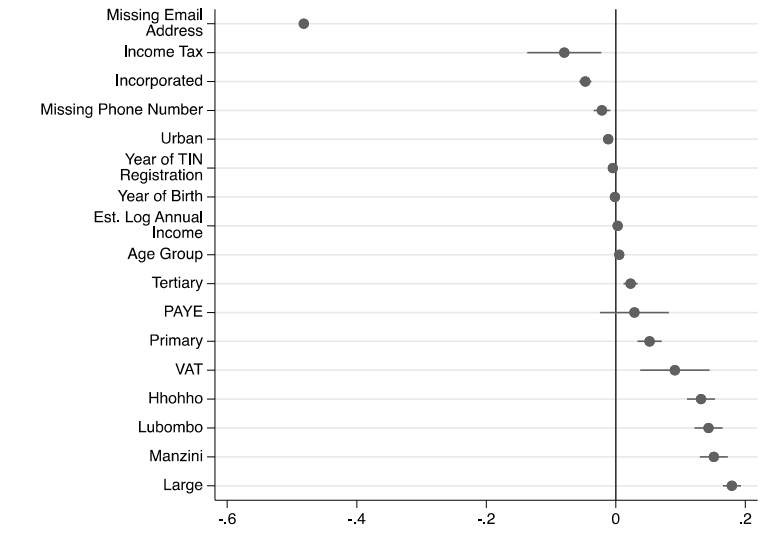
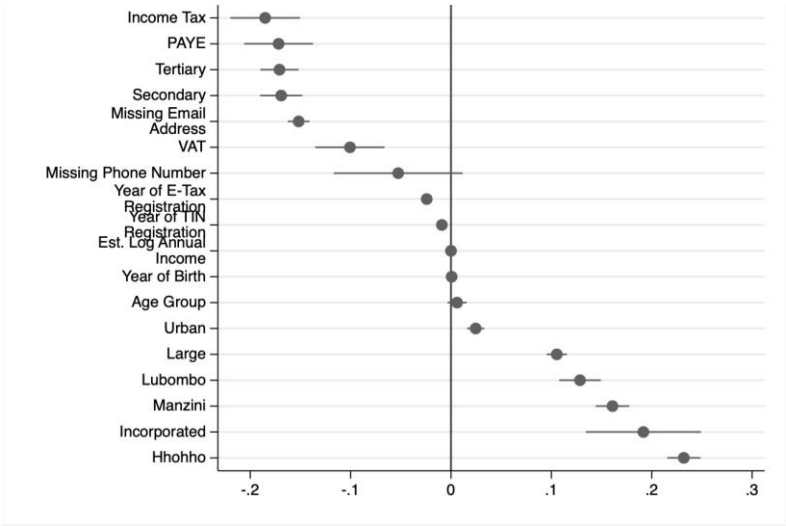


Figure 4b Correlates with digital mode of payment



As a first attempt to control for self-selection, we select a sample of firms that are relatively homogenous. For this, we drop all early registrations – 6,667 taxpayers – who registered voluntarily before the mandate, and hence are not the focus of our study.[26] This means that we consider all those taxpayers who did not register after the mandate as a control group, and compare them to those who registered after the mandate. Early voluntary registrations were likely to introduce bias in our estimates. As shown in Appendix Table A2b, some imbalance still exists between the control and the treatment group as defined above, but the gaps are smaller in magnitude when compared to early registrations (Appendix Table A2a). Below we control for this imbalance further.

Second, we control whether the underlying parallel trend assumption is respected. For this assumption for the DID approach to be sensible, it requires that the treatment and control groups would have similar trends over time in the absence of the treatment. In Figures 4 and 5 below, we indirectly test the parallel trend assumption by showing patterns in filing and payment behaviour over time. Data for up to eight years pre-intervention is available, thus providing a more solid understanding of pre-trends. The e-filing mandate is indicated by the

[26] A similar approach is taken in Mascagni (2021a), studying the impact of the adoption of sales registration machines in Ethiopia.

vertical blue line. Once we drop early adopters, the trends in income, taxable income and tax assessed (Figure 5) seem to be similar before the mandate. Likewise, the trends in payment behaviour tend to be the same (Figure 6). This means that, despite the two groups being different according to background features (Appendix Table A2b), their filing and payment behaviour follows a similar trend.

Figure 5 Parallel trends for variables Log Income, Log Taxable Income, Log Tax Assessed respectively

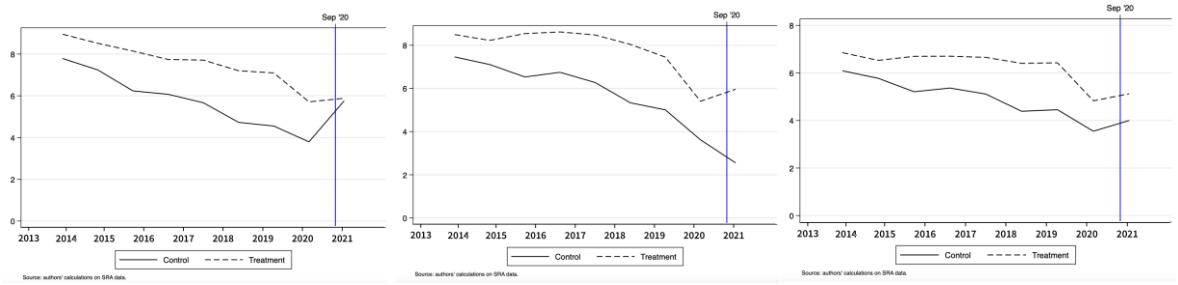
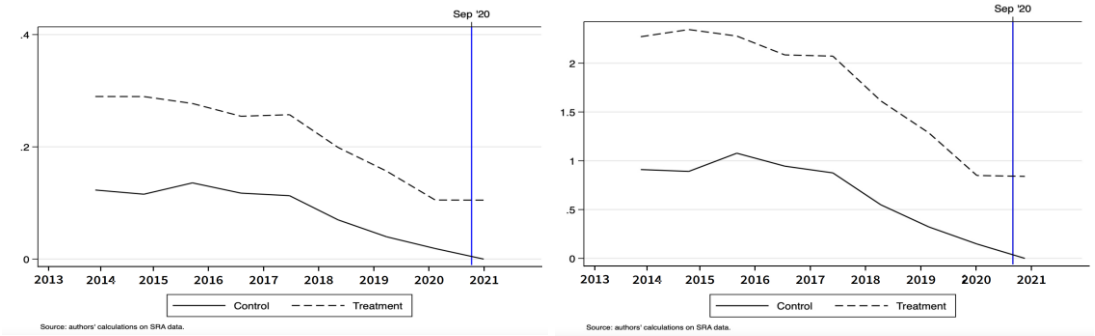


Figure 6 Parallel trends for the variables Paid: Yes No and Log Amount Paid



Despite the fact that the pre-trend in outcomes seems to be parallel, the difference in levels is still a cause of concern. For this reason we adopt a PSM approach as our preferred estimation strategy, after which we run the DID estimation. To better control for the confounding variables likely to explain the partial take-up of the mandate and to arrive at an unbiased treatment effect, we use a PSM estimator to create a better match for the treatment group, based on a vast variety of variables that we gather from ERS datasets. Since the decision to comply with the mandate is not random, PSM mimics randomisation by creating a control group that is very similar to the treatment one based on observables. PSM relies on two main assumptions: *unconfoundedness* – which requires that all confounding factors be included in the set of covariates used to calculate the propensity score – and *overlap* in log odds or common support – which requires that taxpayers with the same score can be either adopters or non-adopters, as if a randomised experiment was carried out. While the first assumption cannot be directly tested, it can be respected by including all the potential confounding variables available in our dataset. The second assumption requires that the propensity score distributions of treated and control units sufficiently overlap, indicating a similar probability to be treated for both groups, in turn mimicking a randomised experiment. Appendix Figure A1 shows that this assumption is satisfied.[27]

PSM can be implemented in several ways – from these we choose a kernel matching algorithm, as it arrives at a more accurate match due to its weighting function. Based on a

[27] Among 13,605 observations, 13,578 observations are retained and 27 observations are out of the common support.

set of observables, a propensity score is created for each observation in our dataset.[28] This score is created by using a range of fixed taxpayer characteristics and pre-mandate variables, likely to determine selection into e-Tax (e.g. turnover), as well as to correlate with our ex-post outcomes.[29] Through the kernel matching algorithm, every observation of the treatment group is then matched with a weighted average of units from the control group.[30] Appendix Figure A2 shows that PSM is successful in reducing imbalance in the different covariates after the match.

As a final check on the success of our matching strategy (on top of the sufficient overlap in Appendix Figure A1 and the reduction in imbalance in Appendix Figure A2), we present Figure 7, which shows the distribution of the two groups' log odds,[31] before and after the match. Figure 7 shows that after the match the log odds' distributions almost perfectly converge between the treatment group and the control group, providing robust evidence of the accurate match that has been created for the treatment group.

Figure 7a Matching log odds – filing behaviour

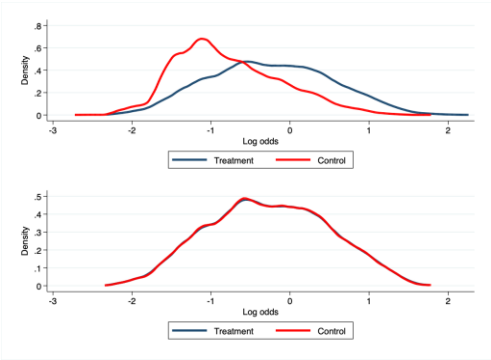
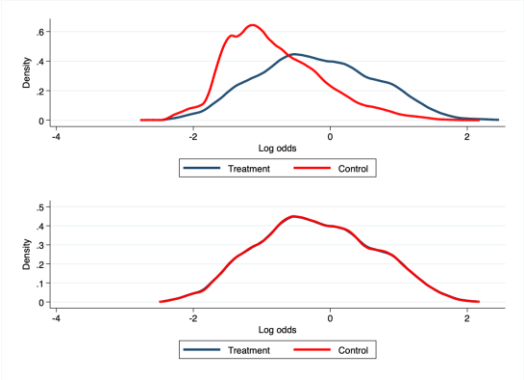


Figure 7b Matching log odds – payment behaviour



4 Results

We first present DID impact estimates on filing behaviour in Table 1. Standard errors are presented in parenthesis. The control group mean is reported at the bottom of the table. Columns 1 and 2 depict the coefficients of the impact of mandatory e-filing on the probability of nil-filing, without and with controls respectively. The first key positive result is the mandate has reduced the probability of nil-filing by a statistically significant 23 percentage points (20pp without controls). The magnitude of this effect is sizeable, about 60 per cent reduction, in a context where almost 40 per cent of the control group mean filed nil (also see Santoro and Mdluli 2019).

28 The propensity score is created by running a logistic model using the treatment (e-Tax registration) as an outcome and the set of cofounders as explanatory variables.

29 The variables that we used for the match are turnover, cost of sales, gross profit, rental income, nil-filers, type of IT, registration year, sector, district, log total assets, log income, log taxable income, log tax payable refund, log land property, status, log total operational expense (available only for CIT taxpayers), and log total deductions (available only for PIT taxpayers). To maximise accuracy, we use three additional variables when creating the match to analyse the payment behaviour: Log Amount Paid, Paid Yes No, which is a dummy for whether the taxpayers made a payment or not, and a categorical variable for payment mode.

30 We selected 0.06 as the bandwidth in the kernel matching algorithm, as common practice in the literature.

31 The log odds ratio is the probability of success divided by the probability of failure.

Table 1 DID estimates of e-Tax policy- filing behaviour

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Nil-filing	**Nil-filing**	**Income**	**Income**	**Taxable income**	**Taxable income**	**Tax amount**	**Tax amount**
E-Tax*Post	-0.20**	-0.23**	1.13	1.90*	3.26***	3.16***	0.40	-0.26
	(0.09)	(0.09)	(1.14)	(1.11)	(0.58)	(0.54)	(0.86)	(0.97)
E-Tax	-0.05***	-0.02***	0.58***	0.26***	0.64***	-0.02	0.48***	-0.09
	(0.00)	(0.01)	(0.06)	(0.08)	(0.05)	(0.07)	(0.05)	(0.06)
Post period	0.29***	0.32***	-2.37**	-3.05***	-4.72***	-4.48***	-1.40	-0.59
	(0.09)	(0.09)	(1.14)	(1.10)	(0.58)	(0.53)	(0.85)	(0.96)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.388	0.388	6.011	6.011	6.267	6.267	5.169	5.169
R-sq.	0.006	0.037	0.005	0.075	0.009	0.094	0.005	0.161
N	61040	61040	61037	61037	57802	57802	60051	60051

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Nil-filing is a 0-1 dummy indicating whether the return is nil. Income, Taxable income and Tax amount are log transformations. Results in col. 1-2 derive from a probit model, while results in col. 3-8 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $eTax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Consistent with the reduction in nil-filing, e-Tax led to an increase in income reported, as shown in columns 3 and 4 – which turns significant when controls are added. The e-filing mandate increases the income reported by 1.9 log units, or about 32 per cent of the control group mean. Likewise, as shown in columns 5-6, the outcome that witnesses the steepest and most significant rise post-intervention is taxable income. This outcome increases by almost 50 per cent of the control group mean, and is highly significantly (column 6). However, there is not sufficient evidence that the final tax assessed has been impacted by e-Tax adoption.

These findings on behaviour show some evidence on the efficiency of the e-Tax system. On the one hand, the mandate reduced zero filing and increased reported taxable income – probably due to the ease of completing an online tax return. This increase in reporting could also be due to a perception of being more closely monitored by the tax agency. On the other hand, the increase in reported taxable income does not carry through to the very final item of the tax return – tax liability. From this analysis, it seems that there is not enough evidence to prove immediate benefits in terms of domestic revenue mobilisation for the tax administration.[32]

Evidence seems to be more positive and less ambiguous for payment behaviour. As shown in Table 2, adoption of e-Tax after the mandate produces a statistically significant increase in both the probability of payment and the amount paid. On the one hand, the mandate increases the probability to pay tax after having filed by 7pp, or a sizeable 70 per cent increase with respect to the control group mean. Remarkably, the average of paying

[32] If tax revenue increased, this in turn raises the question of whether the government will utilise the revenue for development or for corruption.

conditional on filing is very low in the control group, only 10 per cent, and suggests serious tax collection challenges in the country – which e-Tax seems to address.

On the other hand, the mandate boosts the amount paid in tax by 0.5 log units, which translates into a 63 per cent increase over the control group mean. This implies that the mandatory adoption of e-filing has been more effective in improving compliance with payment than filing. The fact that the impact on tax declared and tax actually paid differ so much indicates that these two margins are quite disconnected, as shown in other African contexts (Santoro and Waiswa forthcoming).

Table 2 DID estimates of e-Tax policy on payment behaviour

	(1)	(2)	(3)	(4)
	Paid: Yes No	**Paid: Yes No**	**Amount paid**	**Amount paid**
E-Tax*Post	0.07***	0.07***	0.56***	0.50***
	(0.02)	(0.02)	(0.14)	(0.18)
E-Tax	0.02***	-0.03***	0.14***	-0.28***
	(0.00)	(0.01)	(0.04)	(0.06)
Post period	-0.21***	-0.21***	-1.70***	-1.58***
	(0.02)	(0.02)	(0.13)	(0.17)
Controls	No	Yes	No	Yes
Control Mean	0.100	0.100	0.778	0.778
R-sq.	0.008	0.042	0.008	0.050
N	61065	61065	61065	61065

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Paid Yes-No is a 0-1 dummy indicating whether the taxpayer paid their tax liability. Amount Paid is log transformations. Results in col. 1-2 derive from a probit model, while results in col. 3-4 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for e-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $eTax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

5 Mechanisms

5.1 Role played by mode of payment

The first mechanism to explore refers to the role played by payment mode – digital or non-digital – on the impact of e-Tax on payment behaviour. We test whether any incremental effect is found when merging e-filing and e-payment. This aspect is particularly important given the recent shift towards a zero-cash-handling policy for all tax settlements with ERS (section 2). In order to test for this, we disaggregate Table 2 above by the type of payment adopted. We group modes as mobile money, credit card and EFTs as digital, and cash and cheques (rarely used) as non-digital. As outcomes, we consider both the tax amount paid and an additional variable we create as a proxy for accuracy – a dummy variable that takes a value of 1 if the taxpayer made a full payment (they paid their tax liability in full) and 0 otherwise.[33]

[33] We build this outcome by comparing income tax returns and payment data, and cross-checking the tax liability declared in the return form with the tax actually paid in the payment dataset.

Columns 1 and 2 in Table 3 show that while the amount paid due to e-Tax increases for both payment methods, it is significantly larger for payment through digital means. Likewise, the impact of e-Tax on the probability of fully paying tax liability is more precisely estimated for payers with digital means, while it is only weakly significant for cash payers. Put together, this evidence indicates that digital means do not have a very impressive incremental effect on payment behaviour, in addition to what e-Tax already brings. Yet, they seem to induce slightly higher compliance, or at least the corresponding impact is more precisely estimated. The zero-cash-handling policy only came into force recently (April 2021), and we should not expect a tangible impact too early. From in-depth interviews with ERS, we understand that an explanation for this modest impact of digital payments is that e-Tax is not yet fully configured to offer digital payment solutions – the platform does not contain an e-payment module, and taxpayers can only pay taxes digitally outside it. Mobile money payments are done on a cellphone, while electronic payments are performed through online banking platforms.[34]

Table 3 DID estimates of e-Tax policy on payment behaviour by payment mode

	(1) Amount paid: non-digital	(2) Amount paid: digital	(3) Full payment: non-digital	(4) Full payment: digital
E-Tax*Post	1.14***	1.46***	0.35*	0.18***
	(0.36)	(0.15)	(0.20)	(0.03)
E-Tax	-0.04	-0.17**	-0.01	-0.06***
	(0.08)	(0.07)	(0.02)	(0.02)
Post period	0.00	-2.10***	0.00	-0.15***
	(.)	(0.11)	(.)	(0.03)
Controls	Yes	Yes	Yes	Yes
Control Mean	6.993	8.681	0.333	0.264
R-sq.	0.039	0.113	0.304	0.325
N	4889	5574	4887	5561

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Amount Paid is a log transformation. Full payment is a 0-1 dummy indicating whether the taxpayer paid their tax liability in full. Results in col. 1-2 derive from a tobit model, while results in col. 3-4 come from a probit regression model. We group modes as mobile money, credit card and EFTs as 'digital', while cash and cheques are categorised as non-digital. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

5.2 Ease of filing and reporting

A second mechanism explores the capacity of e-Tax to ease the compliance burden on taxpayers. Thanks to the detailed data available, we are able to investigate:

- whether e-Tax improves on-time registration. This would show whether the IT solution helps taxpayers to respect deadlines, an indicator that could be used as a proxy for ease of complying in the context of Eswatini. As shown in Santoro (2021), taxpayers in Eswatini often do not know when filing deadlines are and end up filing late, triggering severe penalties (Santoro et al. 2020).
- increases in the amount of expenses and deductions reported. This would tell us whether taxpayers are able to claim and deduct all expenses they need, again as a proxy of

[34] The ERS is currently implementing an e-payment module within e-Tax to allow digital payment of taxes; this is expected to go live in 2024.

improved ease to fill a return form. Recent research has shown that less knowledgeable taxpayers tend to leave money on the table by failing to claim their expenses and deductions (Benzarti 2020).

Table 4 shows the impact of the e-Tax mandate on these outcomes. First, column 1 and 2 show that the mandate increased on-time filing by 22 pp, which rises to 29 pp when accounting for controls. These estimates are statistically significant, and quite sizeable when compared to a control group where only about 60 per cent of taxpayers file on-time. This implies a remarkable increase of 48 per cent in on-time filing.

Table 4 DID estimates of e-Tax policy on on-time filing, expenses and deductions

	(1)	(2)	(3)	(4)	(5)	(6)
	On-time filing	**On-time filing**	**Total operational expense (CIT)**	**Total operational expense (CIT)**	**Total deductions (PIT)**	**Total deductions (PIT)**
E-Tax*Post	0.22***	0.29***	4.30***	3.14**	1.07***	1.03***
	(0.06)	(0.07)	(1.08)	(1.38)	(0.30)	(0.33)
E-Tax	0.07***	0.04***	0.67***	0.08	0.23***	-0.43***
	(0.00)	(0.01)	(0.11)	(0.13)	(0.04)	(0.06)
Post period	-0.30***	-0.36***	-5.93***	-4.38***	-0.94***	-0.91***
	(0.06)	(0.07)	(1.07)	(1.38)	(0.29)	(0.32)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.609	0.609	4.659	4.659	2.422	2.422
R-sq.	0.006	0.044	0.009	0.062	0.002	0.088
N	61040	61040	18889	18889	42147	42147

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. On-time filing is a 0-1 dummy indicating whether the taxpayer filed their tax return by the deadline. Expenses and deductions are log transformations. Results in col. 1-2 derive from a probit model, while results in col. 3-6 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $eTax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Columns 3 and 4 show an increase of a statistically significant 4.3 log units (3.14 without controls) in the operational expenses reported by treated taxpayers. Likewise, the mandate increased the total deductions by 1.03 log units (1.07 without controls). It could also be possible that taxpayers might be claiming more expenses to reduce their tax liability – we do not see this as a rational response, given the greater scrutiny implied by sharing information with the tax agency.

The conclusion is that the e-Tax mandate has a positive and statistically significant impact on all three extra indicators of ease in filing. This evidence reinforces our argument that the IT solution has an impact on taxpayers' compliance behaviour, particularly through the 'facilitation' mechanism. For instance, the platform clearly communicates all coming deadlines, increasing their salience and nudging taxpayers to respect them (Figure 1b). The system also guides the taxpayers in reporting, providing extra information and facilitating the whole tax filing experience. Despite insufficient evidence to prove that the e-Tax system helps to generate immediate extra gains in tax due (Table 1), it has a particular impact on helping taxpayers to comply.

5.3 Accuracy in filing and paying

After showing that e-Tax has an impact on facilitating taxpayer compliance, we now turn to test whether it has any impact on the accuracy of amounts declared. We do this with an exercise cross-matching return items across different items.

With respect to filing accuracy, we compare filing behaviour with income tax and VAT – the latter aggregated at the year level, to be comparable with income tax. We first build the variable Accuracy of Filing, which is a categorical variable for whether taxpayers filed their VAT tax, conditional on having filed their income taxes. Second, in much the same vein as Mascagni and Mengistu (2021), we use the variable Accuracy of Reporting as a categorical variable based on the difference between the turnover amounts reported in the two tax returns. In theory, these amounts should be the same over the same year of reference. We cross-check this, and the indicator Accuracy of Reporting indicates whether the two amounts are the same.[35]

Table 5 below shows the results. Columns 1 and 2 document the positive and statistically significant casual impact of the e-Tax mandate on the accuracy of filing. Treated taxpayers are 0.81pp (0.72 without controls) more likely to file their VAT returns conditional on having filed their income tax – a remarkably large increase of about 137 per cent over the control group. Columns 3 and 4 show the impact of the mandate on the accuracy of the reporting. The results show no conclusive effect on this outcome, whereas only a minority in the control group (28 per cent) report the same turnover amounts across returns.

Table 5 DID estimates of e-Tax policy on accuracy in filing and reporting

	(1) Accurate filing	(2) Accurate filing	(3) Accurate reporting	(4) Accurate reporting
E-Tax*Post	0.72***	0.81***	0.04	0.11
	(0.02)	(0.04)	(0.16)	(0.17)
E-Tax	0.14***	0.15***	-0.08***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
Post period	-0.69***	-0.77***	0.01	-0.06
	(0.01)	(0.03)	(0.16)	(0.17)
Controls	No	Yes	No	Yes
Control Mean	0.591	0.591	0.279	0.279
R-sq.	0.033	0.040	0.012	0.028
N	7264	7264	7264	7264

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Accurate filing is a 0-1 dummy indicating whether the taxpayer filed their VAT return conditional on having filed for income tax as well. Accurate reporting is a 0-1 dummy indicating whether the difference between the turnover reported in income tax and VAT return is nil or within a very small margin. Results derive from a probit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

This evidence may suggest that, again, e-Tax is more effective with basic filing actions, such as actually submitting a tax declaration, rather than reducing discrepancies in returns. This could be due to the increase in salience and awareness on deadlines and filing obligations associated with e-Tax. Also, in-depth interviews indicate that some fields in the VAT return are not mandatory, such as those on zero-rated or exempted supplies – taxpayers would then only declare their VAT-taxable sales. In this sense, the platform could be improved to

[35] We allow for a negligible margin of error of about US$5.

reduce inaccuracies, with the possibility of pre-filling the turnover entry line in the income tax form with information extracted from VAT returns. We return to this point in the Conclusions.

Moving to accuracy with payment, Table 6 investigates the impact of the e-Tax mandate on the probability of making a full payment, an outcome variable already described above. Results show that the e-Tax mandate increased the probability of full payment by a statistically significant 39pp when accounting for controls. Compared to the control group mean, this estimate translates into a large 50 per cent increase. This finding resonates well with the positive impact on tax paid in Table 2, which is not conditional on what the taxpayer declared as tax liability in their return. E-Tax not only raises the tax paid, but also improves the accuracy of these payments, meaning that taxpayers now tend to pay their liability in full. Qualitative evidence from in-depth interviews indicates that this positive impact is mainly ascribed to the increase in transparency and awareness of payments due brought by e-Tax. The software, despite not having a built-in e-payment module (see section 2), clearly shows the payments made so far, as well as the outstanding unpaid liability on which interest is charged and clearly visible to the taxpayers – nudging them to fully pay their tax due.

Table 6 DID estimates of e-Tax policy on full payment

	(1) **Full payment**	**(2)** **Full payment**
E-Tax*Post	0.39***	0.16***
	(0.03)	(0.03)
E-Tax	-0.03***	-0.03**
	(0.01)	(0.01)
Post period	-0.36***	-0.14***
	(0.01)	(0.02)
Controls	No	Yes
Control Mean	0.300	0.300
R-sq.	0.001	0.312
N	10448	10448

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Full payment is a 0-1 dummy indicating whether the taxpayer paid its tax liability in full. Results derive from a probit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

This finding is particularly important in a developing country context, where behaviour in filing and payment compliance is often disconnected, as shown in Santoro et al. (forthcoming) and Santoro and Waiswa (forthcoming). This disconnection tends to take place for a number of reasons. First, the human factor – tax officials, on the one hand, often arbitrarily extract payments from taxpayers without ex-ante verification of tax liability; taxpayers, on the other, pay only a portion of their liability just to please the revenue authority and avoid larger penalties (Santoro et al. forthcoming). Second, inaccuracies in the data itself – payment data is often entered separately, without an unambiguous link with tax returns data, due to missing interfaces in the tax management system (Santoro and Waiswa forthcoming). In the case of Eswatini, these discrepancies reduce thanks to e-Tax. Clear benefits for the tax administration could arise in terms of better measurement of performance and budget

reconciliation – two issues that still persist in some African tax administrations (Santoro et al. forthcoming).

5.4 Heterogeneous impact

Motivated by understanding the impact of e-Tax on different taxpayer categories, we conduct further heterogeneity analysis on our outcomes by various dimensions of heterogeneity. Appendix Tables A3a-g show the results of the heterogeneity analysis on filing behaviour,[36] by six dimensions: taxpayer type (CIT vs PIT), taxpayer size (small vs medium/large, location (urban vs rural),[37] experience with the tax system as proxied by the registration year, business sector, and taxpayer age.

A number of conclusions can be drawn from this analysis. First, individual taxpayers seem to be driving the impact on filing behaviour, as they are less likely than companies to submit zero taxes, more likely to report higher income by 2.18 log units of income, more likely to have higher log taxable income, and more likely to file their taxes on time (Appendix Table A3a). Interestingly, nil-filing from companies is unaffected, in a context where previous research showed how this is often the preferred choice of companies (Santoro et al. 2020). These differences could be explained by the fact that companies might be better equipped to put in place tax minimisation schemes, and are probably less concerned of being on the tax agency's radar – as shown in Santoro et al. (2020). Further research could better explore the difference in practice and motivation between individuals and companies.

Second, and related to the finding above, we find that the mandate was more effective on small taxpayers – Appendix Table A3b shows that this group drives the impact across all seven outcomes of filing behaviour. Medium and large taxpayers are unaffected, compared to small taxpayers. Again, this could be linked to the fact that this group (which is more likely to be composed of CIT payers) is not particularly concerned about the enforcement and monitoring capacity of ERS, which e-Tax is increasing. As a separate explanation, this category could just be better able to navigate the tax system and the facilitation effect of e-Tax appears to be muted.

Lastly, no clear patterns emerge when we compare urban and rural taxpayers, or when we compare more or less experienced entities. No significant differences between sectors are found either – if anything, the secondary sector seems to be less responsive. In terms of age,[38] older taxpayers seem to report higher income and taxable income, and are more likely to pay taxes on time (Appendix Tables A3g.vii).[39]

The heterogeneity analysis of payment behaviour partially aligns with that of filing behaviour. We focus on three outcomes – probability to pay, log tax amount paid and probability to perform a full payment against the tax liability. For the sake of brevity, in Tables A4a-c we consider three taxpayer dimensions only – taxpayer type (CIT vs. PIT), taxpayer size (small vs. medium/large, location (urban vs. rural). Interestingly, while CIT payers were largely unresponsive in terms of filing, they are more likely than PIT payers to pay their taxes and increase the amount paid. PIT payers are significantly more likely to make a full payment,

[36] Again, we use the 7 outcomes for filing behaviour: Nil-filers, Log Income, Log Taxable Income, Log Tax Assessed, Status of Filing, Log Total Operational Expense (CIT), and Log Total Deductions (PIT).

[37] We also perform the analysis across the four districts in the country, which we omit for brevity and can be shared on request.

[38] Variable age is a categorical variable, grouping the taxpayers' age into younger than 30, between 30 and 50 inclusive, and taxpayers older than 50, in line with common practice in age grouping.

[39] Older taxpayers might have more income to disclose. They may be more risk-averse, and less likely to delay their filing correctly.

more than doubling the extremely low control group average of 13 per cent (Appendix Table A4a).

Importantly, Appendix Table A4b shows a worrying effect on large and medium taxpayers. This is probably the only negative result of the analysis, and calls for more scrutiny. Large and medium taxpayers are less likely to pay their taxes by 30pp, show a reduction in their amount of tax paid by 3.23 log units, and are less likely to make a full payment by 5pp. Small taxpayers are much more likely to pay their liabilities in full – again a positive result on the accuracy of payment. Yet the negative evidence on medium and large entities parallels the lack of impact on their filing behaviour (Appendix Table A3b), and draws interesting questions on equity. Why is the e-Tax mandate not effective in getting more payment from larger taxpayers, and only effective on the small ones? Are medium and large entities more able to exploit the loopholes in e-Tax, or simply not concerned with the monitoring capacity of ERS? Or are they already the ones more likely to register for e-Tax, so the mandate itself didn't add any significant compliance behaviour? We leave these questions to future research.

6 Conclusions

In conclusion, there is causal evidence on the positive impact of the adoption of mandatory e-filing on tax compliance, both in terms of filing and payment behaviour. Our analysis shows that the mandate reduced the probability of nil-filing and increased reported turnover (Table 1), as well as enhancing the probability of payment and amount of tax paid (Table 2). The latter finding is particularly important – we do not find a significant impact on the tax declared, yet the amount of tax actually paid increases. Among the key explanatory mechanisms, we find only modest evidence that e-payment plays a role (Table 3). More importantly, we show that e-Tax improves on-time filing and pushes taxpayers to claim more expenses and deductions (Table 4). The useful features of the platform, which remind taxpayers of deadlines and help filing accurately, explain this. Furthermore, we document a positive spillover effect on VAT filing (Table 5), again explained by the reminder and informational aspect of the platform. Likewise, e-Tax improves the likelihood of full payment of tax due, thanks to the salience it puts on outstanding tax liabilities. Compared to manual filing, e-Tax seems to have reduced compliance costs by curbing forgetfulness and the difficulty of filing a manual return, and augmented tax payment conditional on filing.

In terms of policy recommendations, this research has proved the positive impact of e-Tax on tax compliance. This is an incentive for the Eswatini Revenue Authority to further promote its adoption. A number of recommendations can be made. First, a key finding is that only a minority (about 40 per cent) of taxpayers managed to register to e-Tax after the mandate. As ERS is currently migrating all remaining taxpayers onto the platform, it could strengthen its facilitation strategy through the internet kiosks (section 2), which could also offer broader taxpayer training and assistance. ERS's recent efforts at sensitising taxpayers through different means (videos on the website, radio, call-ins and 1-1 assistance) could be more directed towards e-Tax adoption. Similar challenges exist in other African countries (ATAF 2021).[40]

Second, the system could be further improved in specific ways. To begin with, we recommend the general strengthening of the system, as it can often fail during peak times,

[40] Where these online platforms exist, there are reports of non-responsiveness by staff, poor issue resolution, task backlog, and some platforms only open between 8am and 5pm while tax activities go on 24/7 (ATAF 2021).

as documented in other African countries (ATAF 2021; Santoro et al. forthcoming).[41] An interview with ERS clarified that taxpayers often leave their tax filing until the last minute – this puts pressure on the network, and means few staff are available to support the spike in queries. These technical difficulties generate frustration and a general sense of unfairness (Santoro et al. forthcoming). Second, to address the only negative result of the study – that on medium and large taxpayers – ERS can introduce a more detailed and strict set of checks and validations for the affected groups in order to reduce their attempts to avoid tax. Third, and separately, an e-payment module could be integrated into the platform so that taxpayers can pay taxes on e-Tax (and not from other platforms). The ERS is planning to do this in the near future.

Further research is needed. It is not clear whether ERS has started making any use of the new data generated by e-Tax and digital payments, which is arguably of higher quality. Anecdotal evidence from the agency reveals that a staff re-organisation is still taking place, and this can lead to issues.[42] In addition, a closer look at how e-Tax is configured to store information is needed, as in-depth interviews unearthed some constraints in data capture.[43] While this paper looks at taxpayer responses to technology, these questions could be addressed through a future study focusing on the consequences on data and IT practices within ERS, taking the internal perspective of the tax administration. Further work is also required on the taxpayer side, through the collection of survey data that could provide extra information on the practical impact of technology on taxpayers. This would help us to understand why taxpayers still select manual filing of tax returns, despite the availability of the e-Tax platform.

41 Taxpayers who responded to the 2020 ATAF ICT survey said the lack of stability in the tax system was the main barrier to using the RA e-system (ATAF 2021).

42 We understand that little capacity is left for responding to requests for amendments. E-Tax currently does not allow amendments, and taxpayers have to submit online queries for these. This means that ERS staff are now very busy with these, having less time to focus on desk audits.

43 For instance, the income tax return seems not to be carefully configured on e-Tax, as there are no validation controls comparing data entered on the balance sheet and actual tax return. Likewise, currently e-Tax cannot pick up a loss brought forward - the taxpayer will incur a larger tax liability as they cannot deduct a previous loss. This reconciliation can be done only manually by ERS staff, implying extra work.

Appendices

Tables

Table A1 Governance and country indicators

	Eswatini	Southern Africa	Year
Tax to GDP ratio[a]	14.7%	22.3%	2018
Tax revenue per capita (USD)[a]	444	949	2015
Informality (% national income)[b]	40.1	32.3	2005-2015
CPI[c]	39	47	2017
Governance indicators[d]			
Control of corruption	-0.44	0.18	2016
Rule of law	-0.32	0.10	2016
Regulatory quality	-0.58	-0.07	2016
Government effectiveness	-0.56	-0.08	2016
Political stability	-0.49	0.19	2016
Voice and accountability	-1.42	0.06	2016
Index of economic freedom[e]	55.9	60.2	2018
Tax burden	74.8	64.9	2018
Government integrity	27	41.4	2018
Judicial effectiveness	35.3	52.6	2018
Business freedom	61.1	63.2	2018
Doing business indicator[f]	59.5	62.3	2018
Starting a business	77.2	79	2018
Registering property	60.8	57	2018
Paying taxes	77.1	76.2	2018
Bank account ownership[g]	29%	42%	2017

Southern Africa: Botswana, Lesotho, Namibia, South Africa and Eswatini.
a Annual Report of the ERS
b Medina and Schneider (2018)
c Transparency International Corruption Perceptions Index. Range: 0-100.
d World Bank (2017). Range: -2.5 (weak) to 2.5 (strong).
e The Heritage Foundation. Range: 0-100.
f World Bank (2018). Range: 0-100.
g Global Findex (2017). Adults (+15 yo) in labor force. Burundi excluded.

Table A2a Balance test table between early adopters and those who register after the mandate

	Early adopters	Adopters after the mandate	Difference
Registration Year	2,013.046	2,012.925	0.121***
	(3.339)	(3.882)	(0.024)
Individual	0.399	0.723	-0.324***
	(0.490)	(0.447)	(0.003)
Non Individual	0.601	0.277	0.324***
	(0.490)	(0.447)	(0.003)
Small	0.929	0.981	-0.052***
	(0.256)	(0.135)	(0.001)
Large	0.070	0.019	0.052***
	(0.256)	(0.135)	(0.001)
Hhohho	0.437	0.422	0.015***
	(0.496)	(0.494)	(0.003)
Lubombo	0.091	0.118	-0.027***
	(0.287)	(0.323)	(0.002)
Manzini	0.394	0.409	-0.015***
	(0.489)	(0.492)	(0.003)
Shiselweni	0.078	0.051	0.027***
	(0.268)	(0.220)	(0.002)
Sector: Primary	0.054	0.040	0.014***
	(0.226)	(0.196)	(0.001)
Sector: Secondary	0.134	0.078	0.056***
	(0.341)	(0.268)	(0.002)
Sector: Tertiary	0.560	0.475	0.085***
	(0.496)	(0.499)	(0.003)
Year of Birth	1,969.526	1,969.431	0.095
	(11.719)	(12.196)	(0.114)
Frequency	6.898	6.443	0.455***
	(1.845)	(2.071)	(0.013)
Missing Email ID	0.037	0.104	-0.068***
	(0.188)	(0.306)	(0.002)
Missing Phone Number	0.001	0.003	-0.002***
	(0.034)	(0.058)	(0.000)
Observations	39,037	57,695	96,732

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Early adopters are taxpayers who registered for E-Tax before the mandate, while adopters after the mandate did so after September 2020.

Table A2b Balance test table between the control group (never adopters) and the treatment group (late adopters)

	Control	Treatment	Difference
Registration Year	2,012.206	2,012.925	0.718***
	(72.469)	(9.236)	(0.304)
Individual	0.747	0.723	-0.024***
	(0.435)	(0.447)	(0.003)
Non Individual	0.253	0.277	0.024***
	(0.435)	(0.447)	(0.003)
Small	0.993	0.981	-0.012***
	(0.083)	(0.135)	(0.001)
Large	0.007	0.019	0.012***
	(0.083)	(0.135)	(0.001)
Hhohho	0.011	0.006	-0.005***
	(0.103)	(0.076)	(0.001)
Lubombo	0.391	0.420	0.029***
	(0.488)	(0.493)	(0.003)
Manzini	0.133	0.117	-0.016***
	(0.340)	(0.322)	(0.002)
Shiselweni	0.397	0.407	0.010***
	(0.489)	(0.491)	(0.003)
Sector: Primary	0.024	0.040	0.016***
	(0.153)	(0.196)	(0.001)
Sector: Secondary	0.076	0.078	0.001
	(0.266)	(0.268)	(0.002)
Sector: Tertiary	0.348	0.475	0.126***
	(0.476)	(0.499)	(0.003)
Year of Birth	1,969.412	1,969.431	0.019
	(0.062)	(0.0607)	(0.087)
Frequency	4.283	6.443	2.160***
	(2.189)	(2.071)	(0.012)
Missing Email ID	0.621	0.104	-0.516***
	(0.485)	(0.306)	(0.002)
Missing Phone Number	0.017	0.003	-0.014***
	(0.130)	(0.058)	(0.001)
Observations	59,376	57,695	117,071

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Control are taxpayers who never registered for E-Tax, while Treatment registered after September 2020.

Table A3a DID estimates of e-Tax policy on filing behaviour by TP type

	(1) Nil-filers: CIT	(2) Nil-filers: PIT	(3) Income: CIT	(4) Income: PIT	(5) Taxable income: CIT	(6) Taxable income: PIT	(7) Tax assessed: CIT	(8) Tax assessed: PIT	(9) On-time: CIT	(10) On-time: PIT
E-Tax*Post	-0.01	-0.33***	0.90	2.18**	2.16**	3.05***	-1.75	1.23	0.19**	0.25***
	(0.17)	(0.08)	(2.18)	(0.86)	(0.86)	(0.80)	(1.45)	(0.84)	(0.09)	(0.07)
E-Tax	0.01	0.00	0.25*	-0.47***	-0.30**	-0.36***	-0.33***	-0.25***	0.07***	0.01
	(0.01)	(0.01)	(0.14)	(0.11)	(0.13)	(0.10)	(0.11)	(0.09)	(0.01)	(0.01)
Post period	0.07	0.42***	-2.06	-3.21***	-3.65***	-4.21***	0.80	-1.96**	-0.27***	-0.31***
	(0.17)	(0.08)	(2.17)	(0.85)	(0.86)	(0.80)	(1.44)	(0.84)	(0.09)	(0.07)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.570	0.326	4.729	6.445	2.714	7.296	1.611	6.395	0.485	0.652
R-sq.	0.049	0.043	0.065	0.055	0.074	0.079	0.034	0.077	0.016	0.052
N	18893	42147	18892	42145	15678	42124	18893	41158	18893	42147

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Nil-filing is a 0-1 dummy indicating whether the return is nil. Income, Taxable income and Tax amount are log transformations. On-time filing is a 0-1 dummy indicating whether the taxpayer filed their tax return by the deadline. Results in col. 1-2 and 9-10 derive from a probit model, while results in col. 3-8 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e-Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3b DID estimates of e-Tax policy on filing behaviour by TP size

	(1) Nil-filers: small	(2) Nil-filers: medium/ large	(3) Income: small	(4) Income: medium/ large	(5) Taxable income: small	(6) Taxable income: medium/ large	(7) Tax assessed : small	(8) Tax assessed : medium/ large	(9) On-time: small	(10) On-time: medium/l arge	(11) Total expenses (CIT): small	(12) Total expenses (CIT): medium/ large	(13) Total deductions (PIT): small	(14) Total deductions (PIT): medium/ large
E-Tax*Post	-0.23**	-0.03	1.89*	-0.05	3.14***	0.46	-0.28	0.26	0.29***	-0.03	3.13**	0.98	1.02***	-0.16
	(0.09)	(0.03)	(1.11)	(0.64)	(0.54)	(0.38)	(0.97)	(0.37)	(0.07)	(0.07)	(1.38)	(1.28)	(0.33)	(0.68)
E-Tax	-0.02***	-0.00	0.25***	1.14	-0.03	-0.68	-0.10	-0.89	0.04***	-0.06	0.09	-4.66**	-0.41***	-3.61***
	(0.01)	(0.06)	(0.08)	(1.19)	(0.07)	(0.94)	(0.06)	(0.86)	(0.01)	(0.10)	(0.13)	(1.86)	(0.06)	(0.63)
Post period	0.32***	0.00	-3.06***	0.00	-4.49***	0.00	-0.60	0.00	-0.36***	0.00	-4.37***	0.00	-0.89***	0.00
	(0.09)	(.)	(1.10)	(.)	(0.53)	(.)	(0.96)	(.)	(0.07)	(.)	(1.38)	(.)	(0.32)	(.)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.390	0.087	5.970	11.891	6.225	12.041	5.128	10.880	0.609	0.688	4.658	6.387	2.376	7.503
R-sq.	0.035	0.090	0.068	0.062	0.082	0.120	0.147	0.149	0.044	0.044	0.062	0.484	0.052	0.147
N	60294	746	60291	746	57060	742	59306	745	60294	746	18844	45	41446	701

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Nil-filing is a 0-1 dummy indicating whether the return is nil. Income, Taxable income, Tax amount, Expenses and Deductions are log transformations. On-time filing is a 0-1 dummy indicating whether the taxpayer filed their tax return by the deadline. Results in col. 1-2 and 9-10 derive from a probit model, while results in col. 3-8 and 11-14 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3.c DID estimates of e-Tax policy on filing behaviour by urban/rural

	(1) Nil-filers: rural	(2) Nil-filers: urban	(3) Income: rural	(4) Income: urban	(5) Taxable income: rural	(6) Taxable income: urban	(7) Tax assessed : rural	(8) Tax assessed : urban	(9) On-time: rural	(10) On-time: urban	(11) Total expenses (CIT): rural	(12) Total expenses (CIT): urban	(13) Total deductions (PIT): rural	(14) Total deductions (PIT): urban
E-Tax*Post	-0.23	-0.20*	2.39	1.02	3.08***	2.92***	-1.88	0.64	0.30***	0.29***	2.88	2.47	0.44	1.37***
	(0.16)	(0.11)	(1.63)	(1.39)	(0.84)	(0.67)	(1.86)	(1.09)	(0.11)	(0.08)	(2.12)	(1.73)	(0.50)	(0.40)
E-Tax	-0.03**	-0.01*	0.49***	0.15	0.28**	-0.16*	0.15	-0.21***	0.05***	0.03***	0.32	-0.03	-0.10	-0.61***
	(0.01)	(0.01)	(0.15)	(0.10)	(0.13)	(0.09)	(0.11)	(0.08)	(0.01)	(0.01)	(0.27)	(0.14)	(0.10)	(0.08)
Post period	0.34**	0.28**	-3.56**	-2.15	-4.59***	-4.15***	0.73	-1.37	-0.34***	-0.37***	-4.32**	-3.63**	-0.45	-1.16***
	(0.16)	(0.11)	(1.62)	(1.39)	(0.82)	(0.66)	(1.85)	(1.08)	(0.11)	(0.08)	(2.11)	(1.73)	(0.48)	(0.39)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.371	0.395	6.151	5.949	6.351	6.228	5.303	5.110	0.623	0.603	4.916	4.590	2.071	2.602
R-sq.	0.044	0.038	0.081	0.077	0.083	0.103	0.130	0.176	0.047	0.045	0.097	0.063	0.106	0.085
N	17787	43253	17785	43252	17052	40750	17475	42576	17787	43253	4160	14729	13627	28520

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Nil-filing is a 0-1 dummy indicating whether the return is nil. Income, Taxable income, Tax amount, Expenses and Deductions are log transformations. On-time filing is a 0-1 dummy indicating whether the taxpayer filed their tax return by the deadline. Results in col. 1-2 and 9-10 derive from a probit model, while results in col. 3-8 and 11-14 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3d DID estimates of e-Tax policy on filing behaviour by TIN registration year

	(1) Nil-filers: before 2014 TY	(2) Nil-filers: after 2014 TY	(3) Income: before 2014 TY	(4) Income: after 2014 TY	(5) Taxable income: before 2014 TY	(6) Taxable income: after 2014 TY	(7) Tax assessed : before 2014 TY	(8) Tax assessed : after 2014 TY	(9) On-time: before 2014 TY	(10) On-time: after 2014 TY	(11) Total expenses (CIT): before 2014 TY	(12) Total expenses (CIT): after 2014 TY	(13) Total deductions (PIT): before 2014 TY	(14) Total deductions (PIT): after 2014 TY
E-Tax*Post	-0.26***	-0.24*	1.41	2.28	2.67**	3.29***	1.60	-0.87	0.09	0.32***	-0.98	3.64**	1.24***	0.95*
	(0.10)	(0.13)	(1.09)	(1.47)	(1.09)	(0.60)	(1.03)	(1.21)	(0.09)	(0.07)	(3.08)	(1.75)	(0.40)	(0.50)
E-Tax	-0.05***	0.04***	0.69***	-0.52***	0.42***	-0.85***	0.16**	-0.70***	0.07***	-0.00	1.18***	-1.27***	-0.51***	-0.30***
	(0.01)	(0.01)	(0.09)	(0.15)	(0.09)	(0.14)	(0.07)	(0.12)	(0.01)	(0.01)	(0.15)	(0.22)	(0.07)	(0.11)
Post period	0.35***	0.32**	-2.53**	-3.29**	-3.90***	-4.58***	-2.40**	0.00	-0.18**	-0.34***	-0.76	-3.97**	-1.07***	-0.96**
	(0.10)	(0.12)	(1.09)	(1.46)	(1.09)	(0.59)	(1.02)	(1.20)	(0.09)	(0.07)	(3.07)	(1.75)	(0.39)	(0.49)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.379	0.417	6.126	5.624	6.411	5.773	5.295	4.740	0.616	0.586	4.814	4.282	2.381	2.578
R-sq.	0.037	0.061	0.079	0.091	0.092	0.126	0.164	0.182	0.032	0.089	0.059	0.094	0.094	0.077
N	45131	15909	45130	15907	42927	14875	44721	15330	45131	15909	12827	6062	32300	9847

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Nil-filing is a 0-1 dummy indicating whether the return is nil. Income, Taxable income, Tax amount, Expenses and Deductions are log transformations. On-time filing is a 0-1 dummy indicating whether the taxpayer filed their tax return by the deadline. Results in col. 1-2 and 9-10 derive from a probit model, while results in col. 3-8 and 11-14 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3e.i DID estimates of e-Tax policy on turnover by sector

	(1) Primary	(2) Secondary	(3) Tertiary
E-Tax*Post	3.51***	-1.50	1.32
	(1.23)	(2.21)	(1.61)
E-Tax	1.86***	0.39	0.49***
	(0.40)	(0.31)	(0.11)
Post period	-4.54***	0.14	-3.09*
	(1.15)	(2.18)	(1.60)
Controls	Yes	Yes	Yes
Control Mean	5.025	4.626	5.253
R-sq.	0.230	0.093	0.068
N	2373	5547	28601

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3e.ii DID estimates of e-Tax policy on taxable income by sector

	(1) Primary	(2) Secondary	(3) Tertiary
E-Tax*Post	2.22	-0.66	3.03***
	(1.45)	(2.20)	(0.60)
E-Tax	1.66***	-0.11	0.03
	(0.37)	(0.27)	(0.10)
Post period	-4.23***	-0.92	-4.89***
	(1.38)	(2.18)	(0.59)
Controls	Yes	Yes	Yes
Control Mean	4.567	3.574	4.568
R-sq.	0.138	0.074	0.078
N	2099	4954	26257

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3e.iii DID estimates of e-Tax policy on tax assessed by sector

	(1) Primary	(2) Secondary	(3) Tertiary
E-Tax*Post	-0.53	-0.75	-1.13
	(2.21)	(1.59)	(1.34)
E-Tax	0.34	-0.26	-0.05
	(0.30)	(0.23)	(0.09)
Post period	-0.80	-0.40	-0.17
	(2.18)	(1.57)	(1.33)
Controls	Yes	Yes	Yes
Control Mean	3.458	2.527	3.419
R-sq.	0.148	0.094	0.133
N	2326	5471	28203

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3e.iv DID estimates of e-Tax policy on total deductions by sector

	(1) Primary	(2) Secondary	(3) Tertiary
E-Tax*Post	2.12***	0.37	0.18
	(0.49)	(0.50)	(0.37)
E-Tax	0.24	-0.58	-0.23**
	(0.18)	(0.37)	(0.10)
Post period	-2.04***	-0.28	-0.23
	(0.42)	(0.37)	(0.36)
Controls	Yes	Yes	Yes
Control Mean	0.713	1.094	1.133
R-sq.	0.139	0.064	0.044
N	1162	1496	15105

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3e.v DID estimates of e-Tax policy on total operational expenses by sector

	(1) Primary	(2) Secondary	(3) Tertiary
E-Tax*Post	-0.28	-1.51***	4.63***
	(4.04)	(0.42)	(1.49)
E-Tax	1.33**	-0.01	-0.02
	(0.54)	(0.33)	(0.14)
Post period	-1.06	0.00	-5.74***
	(3.99)	(.)	(1.48)
Controls	Yes	Yes	Yes
Control Mean	3.799	4.194	4.865
R-sq.	0.223	0.077	0.053
N	1209	4051	13496

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3e.vi DID estimates of e-Tax policy on nil-filing by sector

	(1) Primary	(2) Secondary	(3) Tertiary
E-Tax*Post	-0.14	0.04	-0.17
	(0.14)	(0.23)	(0.14)
E-Tax	-0.15***	-0.01	-0.02**
	(0.03)	(0.02)	(0.01)
Post period	0.23*	0.08	0.29**
	(0.14)	(0.23)	(0.13)
Controls	Yes	Yes	Yes
Control Mean	0.505	0.552	0.470
R-sq.	0.099	0.051	0.037
N	2373	5547	28603

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a probit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3e.vii DID estimates of e-Tax policy on on-time filing by sector

	(1) Primary	(2) Secondary	(3) Tertiary
E-Tax*Post	0.34***	-0.14	0.36***
	(0.11)	(0.19)	(0.11)
E-Tax	-0.02	0.06**	0.06***
	(0.03)	(0.02)	(0.01)
Post period	-0.39***	0.07	-0.44***
	(0.10)	(0.19)	(0.11)
Controls	Yes	Yes	Yes
Control Mean	0.557	0.503	0.541
R-sq.	0.154	0.055	0.050
N	2373	5547	28603

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a probit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3g.i DID estimates of e-Tax policy on turnover by age group

	(1) Age<30	(2) 30<=Age<=50	(3) Age>50
E-Tax*Post	-0.08	1.75	-4.80*
	(0.94)	(1.33)	(2.78)
E-Tax	-0.66	-0.29**	-0.34**
	(0.83)	(0.15)	(0.17)
Post period	0.00	-2.58*	3.60
	(.)	(1.32)	(2.78)
Controls	Yes	Yes	Yes
Control Mean	6.014	5.687	5.614
R-sq.	0.082	0.041	0.054
N	675	16434	16286

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3g.ii DID estimates of e-Tax policy on taxable income by age group

	(1) Age<30	(2) 30<=Age<=50	(3) Age>50
E-Tax*Post	-1.15	1.69	1.82*
	(0.89)	(1.04)	(0.95)
E-Tax	-0.97	-0.17	-0.27*
	(0.71)	(0.13)	(0.15)
Post period	0.00	-2.60**	-3.22***
	(.)	(1.03)	(0.94)
Controls	Yes	Yes	Yes
Control Mean	6.634	6.095	6.008
R-sq.	0.155	0.105	0.156
N	649	16047	16028

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3g.iii DID estimates of e-Tax policy on tax assessed by age group

	(1) Age<30	(2) 30<=age<=50	(3) Age>50
E-Tax*Post	-0.58	0.35	-2.94
	(0.71)	(0.94)	(2.22)
E-Tax	-0.03	-0.14	-0.16
	(0.58)	(0.11)	(0.13)
Post period	0.00	-0.83	1.99
	(.)	(0.94)	(2.21)
Controls	Yes	Yes	Yes
Control Mean	5.580	5.175	5.093
R-sq.	0.269	0.150	0.195
N	651	15926	16028

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3g.iv DID estimates of e-Tax policy on total deductions by age

	(1) Age<30	(2) 30<=age<=50	(3) Age>50
E-Tax*Post	0.14	0.82	0.00
	(0.71)	(0.53)	(0.42)
E-Tax	-0.76	-0.47***	0.05
	(0.68)	(0.09)	(0.11)
Post period	0.00	-0.70	0.10
	(.)	(0.52)	(0.40)
Controls	Yes	Yes	Yes
Control Mean	3.364	2.581	2.473
R-sq.	0.095	0.065	0.109
N	481	13879	14236

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3g.v DID estimates of e-Tax policy on total operational expenses by age group

	(1) Age<30	(2) 30<=age<=50	(3) Age>50
E-Tax*Post	-0.28	1.44	-0.52
	(1.65)	(1.44)	(1.26)
E-Tax	-0.48	0.72**	-0.32
	(1.37)	(0.34)	(0.39)
Post period	0.00	-2.10	-0.06
	(.)	(1.45)	(1.13)
Controls	Yes	Yes	Yes
Control Mean	2.613	2.945	2.863
R-sq.	0.191	0.085	0.050
N	194	2555	2051

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a tobit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3g.vi DID estimates of e-Tax policy on nil-filing by age group

	(1) Age<30	(2) 30<=age<=50	(3) Age>50
E-Tax*Post	0.07	-0.26**	0.27
	(0.08)	(0.11)	(0.23)
E-Tax	0.06	-0.02	0.01
	(0.06)	(0.01)	(0.01)
Post period	0.00	0.32***	-0.16
	(.)	(0.11)	(0.23)
Controls	Yes	Yes	Yes
Control Mean	0.373	0.412	0.419
R-sq.	0.066	0.042	0.070
N	675	16435	16287

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a probit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A3.g.vii DID estimates of e-Tax policy on on-time filing by age group

	(1) Age<30	(2) 30<=Age<=50	(3) Age>50
E-Tax*Post	0.02	0.17	0.37**
	(0.08)	(0.11)	(0.18)
E-Tax	-0.12*	0.05***	-0.01
	(0.07)	(0.01)	(0.01)
Post period	0.00	-0.23**	-0.45**
	(.)	(0.11)	(0.18)
Controls	Yes	Yes	Yes
Control Mean	0.595	0.600	0.614
R-sq.	0.075	0.054	0.055
N	675	16435	16287

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Results derive from a probit model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A4a DID estimates of e-Tax policy on payment behaviour by TP type

	(1) Paid Yes No: CIT	(2) Paid Yes No: PIT	(3) Amount Paid: CIT	(4) Amount Paid: PIT	(5) Full Payment: CIT	(6) Full Payment: PIT
E-Tax*Post	0.14***	0.04	1.08***	0.28	0.00	0.19***
	(0.03)	(0.03)	(0.24)	(0.23)	(0.04)	(0.03)
E-Tax	-0.02*	-0.06***	-0.23**	-0.51***	0.00	-0.03*
	(0.01)	(0.01)	(0.10)	(0.08)	(0.02)	(0.02)
Post period	-0.30***	-0.16***	-2.39***	-1.22***	0.00	-0.18***
	(0.03)	(0.03)	(0.23)	(0.23)	(.)	(0.02)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.114	0.095	0.885	0.742	0.714	0.133
R-sq.	0.025	0.046	0.027	0.059	0.022	0.021
N	18892	42173	18892	42173	3848	6600

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Paid Yes-No is a 0-1 dummy indicating whether the taxpayer paid their tax liability. Amount Paid is a log transformation. Full payment is a 0-1 dummy indicating whether the taxpayer paid their tax liability in full. Results in col. 1-2 and 5-6 derive from a ptrobit model, while results in col. 3-4 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A4b DID estimates of e-Tax policy on payment behaviour by TP size

	(1) Paid Yes No: Small	(2) Paid Yes No: Medium/Large	(3) Amount Paid: Small	(4) Amount Paid: Medium/Large	(5) Full Payment: Small	(6) Full Payment: Medium/Large
E-Tax*Post	0.07***	-0.30***	0.54***	-3.23***	0.16***	-0.05*
	(0.02)	(0.06)	(0.18)	(0.50)	(0.03)	(0.02)
E-Tax	-0.03***	-0.01	-0.27***	-0.24	-0.04**	0.15
	(0.01)	(0.10)	(0.06)	(1.04)	(0.01)	(0.09)
Post period	-0.21***	0.00	-1.59***	0.00	-0.14***	0.00
	(0.02)	(.)	(0.18)	(.)	(0.02)	(.)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.098	0.367	0.759	3.561	0.306	0.092
R-sq.	0.038	0.121	0.042	0.145	0.307	0.198
N	60322	743	60322	743	10137	311

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Paid Yes-No is a 0-1 dummy indicating whether the taxpayer paid their tax liability. Amount Paid is a log transformation. Full payment is a 0-1 dummy indicating whether the taxpayer paid their tax liability in full. Results in col. 1-2 and 5-6 derive from a ptrobit model, while results in col. 3-4 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A4c DID estimates of e-Tax policy on payment behaviour by urban/rural

	(1) Paid Yes No: Rural	(2) Paid Yes No: Urban	(3) Amount Paid: Rural	(4) Amount Paid: Urban	(5) Full Payment: Rural	(6) Full Payment: Urban
E-Tax*Post	0.08**	0.06**	0.59*	0.44*	0.05	0.13***
	(0.04)	(0.03)	(0.31)	(0.22)	(0.05)	(0.03)
E-Tax	-0.03**	-0.03***	-0.22**	-0.30***	-0.03	-0.03**
	(0.01)	(0.01)	(0.10)	(0.07)	(0.03)	(0.02)
Post period	-0.20***	-0.20***	-1.53***	-1.57***	0.00	-0.12***
	(0.03)	(0.03)	(0.30)	(0.22)	(.)	(0.02)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Control Mean	0.081	0.108	0.622	0.847	0.299	0.301
R-sq.	0.046	0.039	0.055	0.047	0.214	0.347
N	17753	43312	17753	43312	2497	7951

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Paid Yes-No is a 0-1 dummy indicating whether the taxpayer paid their tax liability. Amount Paid is a log transformation. Full payment is a 0-1 dummy indicating whether the taxpayer paid their tax liability in full. Results in col. 1-2 and 5-6 derive from a ptrobit model, while results in col. 3-4 come from a tobit regression model. E-Tax is a dummy that denotes whether the taxpayer registered for E-Tax after the mandate, $Post$ takes value of 1 if the taxpayer filed after September 2020 and 0 otherwise, $e - Tax \cdot Post$ is the interaction term between the e-Tax registration and the Post period variable. This interaction term captures the effect of being registered for e-Tax and filing after the mandate. The control variables include the firms' sectors, locations, or districts, whether the firm is in an urban or rural area, taxpayer's size, taxpayer's registration year, year of birth and age, and whether there is a missing email address or phone number. ε_{it} is an error term.

Table A5a Predictors of e-Tax registration using a lasso logit model

Selected	Logistic lasso	Post logit
Primary	1.2020003	1.3241835
Large	1.1637520	1.4324974
Tertiary	1.0056298	1.0623392
Secondary	0.8313043	0.9128925
Age group	0.4672408	0.5251244
VAT	0.3287131	0.3901800
Shiselweni	0.1467741	0.3047202
Manzini	0.0877971	0.1621365
Est. log annual income	0.0186098	0.0272631
Lubombo	0.0185604	0.1398299
Year of TIN registration	-0.0136383	-0.0275718
Year of birth	-0.0001461	-0.0001783
Income tax	-0.9396063	-1.0354207
Missing phone number	-1.3498674	-2.2252950
Missing email address	-2.8012252	-2.8707256
Constant	27.5915848	55.5571349

Outcome is a 0-1 dummy for e-Tax registration. The lasso2 coefficients are a result of minimising the residual sum of squares, where the command first runs a full coefficient path for a list of lambda (the tuning parameter chosen by cross validation) then runs the model selected by EBIC (a type of information criterion). The coefficient of lasso represents the predictive power of each variable, proportional with the magnitude of the coefficient.

Table A5b Predictors of digital mode of payment using a lasso logit model

Selected	Logistic lasso	Post logit
Large	1.2280261	1.3168828
Registered for e-Tax	1.0329304	1.0682257
VAT	0.5743842	0.5993364
Hhohho	0.3970454	0.4008834
Year of birth	-0.0000630	-0.0001157
Est. log annual income	-0.0010889	-0.0032980
Primary	-0.1342854	-0.3164018
Year of TIN registration	-0.2241312	-0.2308784
Lubombo	-0.3818353	-0.4189411
Shiselweni	-0.5517539	-0.6009319
Secondary	-0.5753722	-0.7541018
Tertiary	-0.6828601	-0.8118750
Income tax	-0.9641489	-1.0154440
Missing email address	-1.0928441	-1.1136900
Missing phone number	-2.0016108	-2.2662728
Constant	27.5915848	55.5571349

Outcome is a 0-1 dummy for digital mode of payment. The lasso2 coefficients are a result of minimising the residual sum of squares, where the command first runs a full coefficient path for a list of lambda (the tuning parameter chosen by cross validation) then runs the model selected by EBIC (a type of information criterion). The coefficient of lasso represents the predictive power of each variable, proportional with the magnitude of the coefficient.

Figures

Figure A1 Propensity score distribution by treatment and control groups

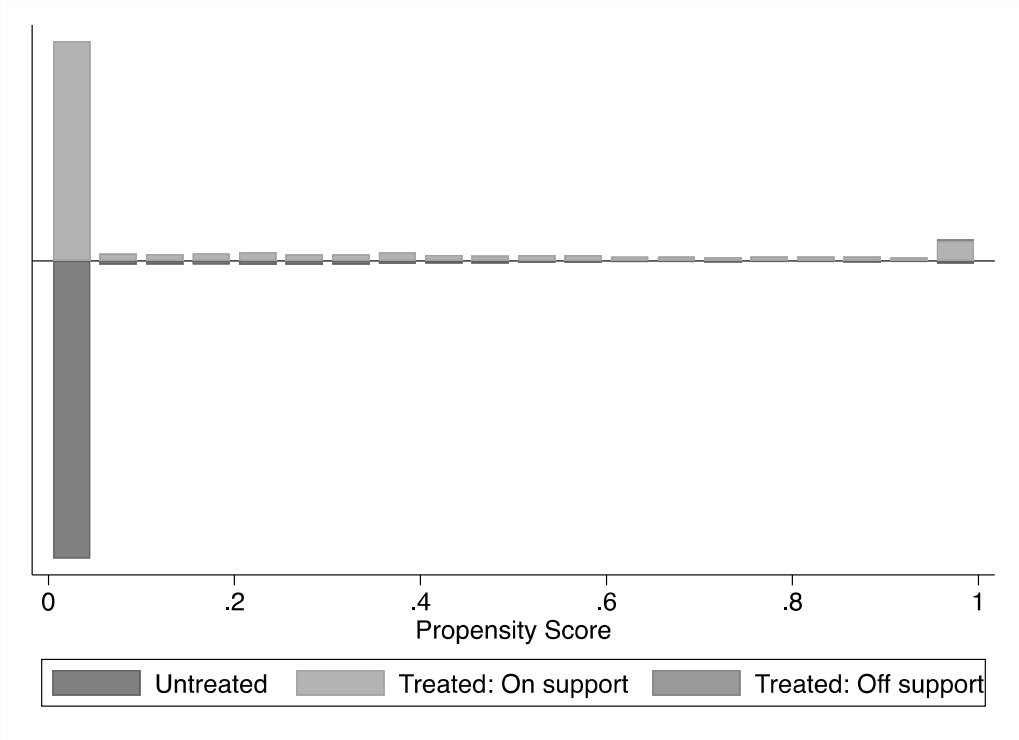
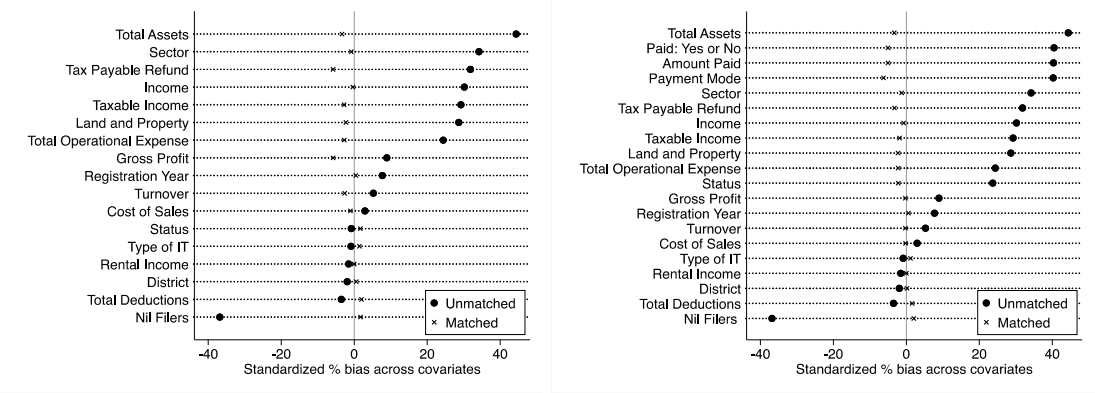


Figure A2 Matching balance for filing behaviour (left) and payments behaviour (right) respectively



References

Ali, M., Shifa, A., Shimeles, A. and Woldeyes, F. (2015) 'Information Technology and Fiscal Capacity in a Developing Country: Evidence from Ethiopia', *SSRN Journal*

Alm, J., Cherry, T., Jones, M. and McKee, M. (2010) 'Tax Information Assistance Services and Tax Compliance Behavior', *Journal of Economic Psychology 31*(4): 577-586

Alm, J. and McClennan, C. (2012) 'Tax Morale and Tax Compliance from the Firm's Perspective', *Kyklos 65*(1): 1-17

Almunia, M., Hjort, J., Knebelmann, J. and Tian, L. (2021) 'Strategic or Confused Firms? Evidence From', *SSRN Journal*, https://doi.org/10.2139/ssrn.3893340

Aragie, E., Taffesse, A. and Thurlow, J. (2021) 'The Short-term Economywide Impacts of COVID-19 in Africa: Insights from Ethiopia', *Afr Dev Rev* 33 (S1)

Arndt, C., Gabriel, S. and Robinson, S. (2020) Assessing the Toll of COVID-19 Lockdown Measures on the South African Economy, Working Paper 111, South Africa–Towards Inclusive Economic Development (SA-TIED), United Nations University World Institute for Development Economics Research (UNU-WIDER). https://sa-tied.wider.unu.edu/sites/default/ files/pdf/SA-TIED-WP-111.pdf.

ATAF (2021) *Efficient Acquisition, Implementation, and Maintenance of Integrated Revenue Administration ICT Systems in Africa*, African Tax Administration Forum Guidebook

Bachas, P., Gadenne, L. and Jensen, A. (2020) *Informality, Consumption Taxes and Redistribution*, IFS Working Paper W20/14, Institute for Fiscal Studies

Benzarti, Y. (2020) How taxing is tax filing? Using revealed preferences to estimate compliance costs, *American Economic Journal: Economic Policy* 12(4): 38-57

Bertrand, M., Duflo, E. and Mullainathan, S. (2004) 'How Much Should We Trust Differences In-Differences Estimates?', *The Quarterly Journal of Economics*, 119(1): 249–275, https://doi.org/10.1162/003355304772839588

Brockmeyer, A., and Hernandez, M. (2019) *Taxation, Information, and Withholding: Evidence from Costa Rica, World Bank Group*, https://doi.org/10.1596/1813-9450-7600

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

Chalendard, C., Duhaut, A., Fernandes, A., Mattoo, A., Raballand, G. and Rijkers, B. (2020) *Does Better Information Curb Customs Fraud?*, World Bank Policy Research Working Paper 9254, World Bank Group

Efobi, U., Belmondo, T., Orkoh, E., Atata, S., Akinyemi, O. and Beecroft, I. (2019) 'Environmental Pollution Policy of Small Businesses in Nigeria and Ghana: Extent and Impact', *Environ Sci Pollut Res* 26 (3): 2882–97, https://doi.org/10.1007/s11356-018-3817-x

Eilu, E. (2018) 'Adoption of Electronic Fiscal Devices (EFDs) for Value-Added Tax (VAT) Collection in Kenya and Tanzania: A Systematic Review', *AJIC* 22: 111-34

Eissa, N. and Zeitlin, A. (2014) 'Using Mobile Technologies to Increase VAT Compliance in Rwanda', in *Centre for the Study of African Economies (CSAE)*, Centre for the Study of African Economies (CSAE)

ERS (2021) *Integrated Annual Report*, Technical report, Eswatini Revenue Authority

—— (2019) *Integrated Annual Report*, Technical report, Eswatini Revenue Authority

Gwaro, T., Maina, K. and Kwasira, J. (2016) 'Influence of Online Tax Filing on Tax Compliance among Small and Medium Enterprises in Nakuru Town, Kenya', *IOSR* 18 (10): 82-92, https://doi.org/10.9790/487x-1810028292

Jouste, M., Nalukwago, M. and Waiswa, R. (2021) 'Do Tax Administrative Interventions Targeted at Small Businesses Improve Tax Compliance and Revenue Collection? Evidence from Ugandan Administrative Tax Data', UNU-WIDER Working Paper 2021/17.

Kamara, A., Meriggi, N. and Prichard, W. (2020) 'Freetown just implemented a new property tax system that could quintuple revenue', https://www.ictd.ac/blog/freetown-new-property tax-system-quintuple-revenue/

Knebelmann, J. (2019) *Taxing Property Owners in Dakar*, International Growth Centre Policy Brief 50415, IGC

Kochanova, A., Hasnain, Z. and Larson, B. (2020) 'Does E-Government Improve Government Capacity? Evidence from Tax Compliance Costs, Tax Revenue, and Public Procurement Competitiveness', The World Bank Economic Review 34 (1): 101-20

Lechner, M. (2011) 'The Estimation of Causal Effects by Difference-in-Difference Methods Estimation of Spatial Panels', *Foundations and Trends in Econometrics* 4: 165-224, 10.1561/0800000014

Ligomeka, W. (2019) *Assessing the Performance of African Tax Administrations: A Malawian Puzzle*, ICTD African Tax Administration Paper 14, Brighton: Institute of Development Studies

Mas'ud, A. (2019) Acceptability of e-Filing of Taxes by Micro-Entrepreneurs in Northwestern Nigeria, Working Paper 96, International Centre for Tax and Development. Brighton: ICTD.

Mascagni, G., Dom, R. and Santoro, F. (2021b) The VAT in Practice: Equity, Enforcement and Complexity, ICTD Working Paper 117, Brighton: Institute of Development Studies

—— and Lees, A. (2021) *Using administrative data to assess the impact of the pandemic in low-income countries: An application with VAT data in Rwanda*, ICTD African Tax Administration Paper 22, Brighton: IDS

—— Mengistu, A. and Woldeyes, F. (2021a) 'Can ICTs Increase Tax Compliance? Evidence on Taxpayer Responses to Technological Innovation in Ethiopia', *Journal of Economic Behavior & Organization* 189: 172-93

—— Santoro, F., Mukama, D., Karangwa, J. and Hakizimana, N. (2022) 'Active Ghosts: Nil-filing in Rwanda', *World Development* (forthcoming)

Mayega, J., Ssuuna, R., Mubajje, M., Nalukwago, M. and Muwonge, L. (2019) *How Clean is our Taxpayer Register? Data Management in the Uganda Revenue Authority*, ICTD African Tax Administration Paper 12, Brighton: Institute of Development Studies

Mayega, J., Waiswa, R., Nabuyondo, J. and Isingoma, M. (2021) How Clean Are Our Taxpayer Returns? Data Management in Uganda Revenue Authority, ICTD African Tax Administration Paper 24, Brighton: Institute of Development Studies

Medina, L. and Schneider, F. (2018) *Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?*, Working paper 18/17, International Monetary Fund

Mittal, S. and Mahajan, A. (2017) *VAT in Emerging Economies: Does Third Party Verification Matter?*, Working Paper, Anderson School of Management, Los Angeles

Ndayisenga, E. and Shukla, J. (2016) 'Effect of Electronic Tax Management System of Tax Collection in Rwanda: Case Study of Rwanda Revenue Authority (RRA)', *The International Journal of Business & Management* 4: 38-49

Nyanga, M. (2021) TADAT INSIGHTS. Perfomance Outcome Area (POA) 1: Integrity of the Registered Taxpayer Base – Essentials of a High Integrity Taxpayer Registry, TADAT

Obert, S., Rodgers, K., Tendai, M.J. and Desderio, C. (2018) 'Effect of E-Tax Filing on Tax Compliance: A Case of Clients in Harare, Zimbabwe', *African Journal of Business Management* 12.11: 338–342

Okunogbe, O. (2021) *Becoming Legible to the State: The Role of Detection and Enforcement Capacity on Tax Compliance*, World Bank Policy Research Working Paper

—— and Pouliquen, V. (2022) 'Technology, Taxation, and Corruption: Evidence from the Introduction of Electronic Tax Filing', *American Economic Journal: Economic Policy*

—— and Santoro, F. (2022) *The Promise and Limitations of Information Technology for Tax Mobilisation*, ICTD Working Paper 135, Brighton: Institute of Development Studies

Rosenbaum, P. and Rubin, D. (1983) 'The central role of the propensity score in observational studies for causal effects', *Biometrika* 70(1): 41-55 https://doi.org/10.1093/biomet/70.1.41

Santoro, F. (2021) 'To file or not to file? Another dimension of tax compliance – the Eswatini Taxpayers' survey', *Journal of Behavioral and Experimental Economics*, 95: 101760, ISSN 2214-8043, https://doi.org/10.1016/j.socec.2021.101760

—— Groening, E., Mdluli, W. and Shongwe, M. (2020) *How to Best Nudge Taxpayers? The Impact of a Tailored Letter Experiment in Eswatini*, ICTD Working Paper 112, Brighton: Institute of Development Studies

—— Lees, A., Carreras, M. and Mukamana, T. (forthcoming) *Digital divides and tax compliance: Adoption and impacts of e-services in Rwanda*, ICTD Working Paper, Brighton: Institute of Development Studies

—— and Mdluli, W. (2019) *Nil-Filing in Eswatini: Should the Revenue Administration be Concerned?*, ICTD African Tax Administration Paper 6, Brighton: Institute of Development Studies

—— and Waiswa, R. (forthcoming) *Small Nets for Big Fish: Taxing the Richest in Uganda*, ICTD African Tax Administration Paper, Brighton: Institute of Development Studies

Shah, S. (2020) *Using Computerized Information to Enforce VAT: Evidence from Pakistan*, job market paper

Slemrod, J. and Weber, C. (2012) 'Evidence of the invisible: toward a credibility revolution in the empirical analysis of tax evasion and the informal economy', *International Tax and Public Finance 19*(1): 25-53

World Bank (2018) *Eswatini – overview*, https://data.worldbank.org/country/SZ

—— (2016) *World Development Report 2016: Digital Dividends*, World Bank Group

—— (2015) *World Development Indicators 2015*, Washington DC: World Bank, https://openknowledge.worldbank.org/handle/10986/21634 License: CC BY 3.0 IGO

Yilmaz, F. and Coolidge, J. (2013) *Can E-filing Reduce Tax Compliance Costs in Developing Countries?*, World Bank Policy Research Working Paper 6647, World Bank Group