Essays on Tax Administration and Tax Compliance in sub-Saharan Africa

Fabrizio Santoro

Advisor 1: Andy McKay Advisor 2: Matthew Embrey

Submitted for the degree of Doctor of Philosophy
Department of Economics
University of Sussex
September 2020

Contents

Tables

Figures

Acknowledgements

The last three years of study, field-work and writing have surely been a rollercoaster of emotions and an unprecedented time of professional growth. There are many people to which I owe much of my learning and I want to whole-heartedly thank in these lines.

Firstly, I would like to express my sincere gratitude to my advisors Andy McKay and Matthew Embrey. They represented a unique source of inspiration and provided me with all the professional and psychological support I needed. They have been able to firmly direct me in the first months of my PhD, as well as to encourage me in the challenging times of fieldwork and to guide me in the last months of writing. I could not have imagined having better advisors and mentors for my PhD experience.

My sincere thanks also goes in advance to my thesis committee, the external examinor Jukka Pirttilä and the internal one Amalavoyal Chari, for the time they spent reading my thesis and for the inputs and comments they will provide, as I expect them to be both challenging and constructive. Many thanks to Chari in particular, who always showed genuine interest towards my work on tax and development and assisted me with inputs on two chapters of this thesis.

Besides my advisors, I would like to thank a third essential mentoring figure, with whom I also wrote a chapter of this thesis - Giulia Mascagni. I have been working with Giulia for the last four years, even before starting the PhD, and I basically owe much (of the little) that I know to her. Giulia proved to be a fantastic informal advisor, a careful teacher and an encouraging friend with whom I could discuss the ups and downs of my PhD journey. I am proud to continue working with her in the future.

Relatedly, I want to thank the whole family of the International Centre of Tax and Development, for which I have been working in the last four years. ICTD represents a special space in which to grow and challenge any conventional wisdom on tax and development. While I worked on many different projects which often clashed with my PhD schedule, I do not regret that as it broadened my research interests and horizons as well as my understanding of tax in Africa. Colleagues like Mick Moore, Wilson Prichard, Vanessa van den Boogard, Giovanni Occhiali, Rhiannon McCluskey, Simon Rees, Camilla Walsh, Oli Roy, Jalia Kangave, Michael Falade, Ronald Waiswa, Rhea Millward-Thompson, Roel

Dom, Max Gallien, Martin Hearson, Adrienne Lees, James Murdoch and Kate Graham all deserve my gratitude.

More broadly, this thesis would not have been possible without the support of the Institute of Development Studies as whole, who administered my contracts and organised my flights to the field, as well as to DFID UK, which funded all three projects in this study. Likewise, I am grateful to the institutional support of the revenue authorities in Rwanda and Eswatini, whose senior management approved my research plans. For the RRA family, I would like to thank staff and dear friends such as Denis Mukama, John Karangwa, Naphtal Hakizimana. For the SRA, my thoughts go to Mbongeni Shongwe, Edward Groening, Winnie Mdluli, Lungile Khumalo and my driver Soldier.

For what concerns fieldwork, I owe my gratitude to the practitioners and field staff who physically implemented my research. They made tangible what I just had in mind. In particular, I thank Laterite Ltd Rwanda and Vanguard Economics (Kigali). Thanks to the latter, I made invaluable friends in Yves Nsengiyumva, brave comrade in the land of a thousand hills, Emmanuel Kwizera, Benjamin Byinshi, Marcel and Christopher Shema, Isabelle Mugeni, Mads Knudsen, Amin Lalui. I also thank the field staff from Eswatini: Thabile Ngwenya, Gamedze Nomasiko, Mashaba Nkosingiphile, Wandile Dlamini, Sabelo Bhembe, Sukoluhle Dlamini, Sipho Dlamini, David Mpungose, Masina Bhekumusa, Thabani Simelane and Nhlanhla Mhlanga, Mkhulisi Mdluli, Sanele Twala and Mamba Mduduzi. Most of them also became good friends of mine. In general, I am indebted to all people of Rwanda and Eswatini - unknowingly, they totally delighted my days of hard work in the field.

Last but not the least, many thanks to friends and family. Even if they knew nothing about my research, they have been there ready to support me in any situation. Thanks to Antonietta, Angelina, Davide, Luca, Kyawthann, Sara, Michele, Caterina, Alessio.

At the end of the day, while dozens of overlapping faces and names come to my mind when thinking about this venture, only one is destined to remain. Many thanks to Francesca, the best human being I could ever ask for to spend my life with. This thesis, as any step I take in this world, is dedicated to you.

UNIVERSITY OF SUSSEX

Fabrizio Santoro PhD in Economics

Essays on Tax Compliance and Tax Administration in Sub-Saharan Africa

SUMMARY

This thesis studies the determinants of tax compliance and the challenges of tax administration in a context of weak fiscal capacity and poor taxpayer attitudes, sub-Saharan Africa. The main research question regards the understanding of why African taxpayers evade taxes and which are the most effective strategies revenue authorities could pursue to foster voluntary compliance.

The thesis contributes to the debate on tax revenue mobilisation in Africa by focusing on Rwanda and Eswatini. It stems from close research collaboration with national revenue authorities and relies on rich administrative data, uniquely combined with survey data. Also, it focuses on compliance of companies and self-employed with income tax, a crucial revenue source in both countries for which opportunities of tax evasion are more common.

This thesis comprises three chapters. The first chapter evaluates the effectiveness of a tax training program on tax knowledge, perceptions and compliance. It combines pre and post-training survey data with tax returns records. I find that taxpayer education significantly improves knowledge and perceptions of complexity. I also show that the program brings taxpayers into the habit of filing returns, in a context where non-filing is widespread.

The second chapter looks at the determinants of non-filing of tax returns in Eswatini. The motivation of this analysis comes from the large extent of this under-studied behavior in SSA. I combine survey data from 1,000 entrepreneurs with their 2013-2018 income tax returns. I show that economic deterrence, compliance costs and moral factors are strongly correlated with actual filing. I also argue that tax knowledge plays a major role in explaining the decision to file.

Lastly, the third chapter implements a randomized controlled trial nudging more than 20,000 income taxpayers in Eswatini and targeting non-filers, nil-filers and active taxpayers. I find that non-filers significantly respond to nudges, while nil and active filers do not. The best performing nudges build on the deterrence and facilitation paradigms. Also, negative responses from large companies are found.

Chapter 1

Introduction

The goal of this thesis is to shed new light on tax compliance and tax administration in a developing region of the world, sub-Saharan Africa (SSA). Domestic revenue mobilisation, enshrined as one of the Sustainable Development Goals, has not reached its full potential in the African Continent. The urgent need to collect more taxes, and, more importantly, in a *better* way, in order to sustain development calls for evidence-based policy recommendations on what works and what does not in the way revenue authorities in SSA enforce and encourage tax compliance.

The central question motivating this thesis refers to how to foster tax compliance in a context where, on one side of the equation, tax administrations are constrained, both in terms of budget and of technical and staff capacity, and, on the other side, society's willingness to voluntary contribute is low. A sequence of relevant research questions naturally arises: (i) How do taxpayers differ in terms of their filing behaviour and how relevant are these compliance patterns in SSA? (ii) Which mix of economic and behavioural factors more reasonably predicts the way in which SSA taxpayers declare their taxes? (iii) After learning more about these factors, what can resource-strapped SSA tax administrations do to leverage them in order to improve both tax attitudes and behaviour?

1.1 Background

Why do taxation and tax administration matter for development? In the last 70 years, taxation has gradually become a fundamental pillar of how scholars model the economic development process. Back in the early '60s, this was probably very clear to the economist Nicolas Kaldor who, exposed to the tax system of a developing country (India), raised the question on why developing countries tax so little and whether they will eventually "learn to tax" (Kaldor, 1963) – a question that still retain its relevance today (Genschel and Seelkopf, 2016). He eventually argued that for a country to become "developed" it needed to collect in taxes 25-30 percent of its GDP, a target of which, as explained below, most SSA countries still remain short.

Since Kaldor, the international community reached unanimous agreement and full commitment in supporting taxation in developing countries, while setting more achievable revenue targets.[1] Specific Sustainable Development Goals were set in development areas for which public financing is critical, including ending poverty (SDG1) and hunger (SDG2), improving health (SDG3) and education (SDG4), achieving gender equality (SDG5), reducing inequality (SDG10), and enhancing infrastructure (SDGs 6, 7, 9, 11). Domestic revenue mobilisation has therefore been championed as the main solution to fill the gaps in available development finance.

Apart from the immediate link between increased tax revenues and more funds to devote to structural development, there are other important reasons for which taxation matters for development in SSA, which will be only sketched out in this section, but will emerge as recurring themes throughout the thesis.

First of all, taxation is mostly about State-building, which in turn is crucial for development. This *governance* argument is clearly outlined in Bird (2015), according to which "the tax system constitutes one of the major interfaces between citizens and state in any country so how taxes are administered may affect [...] public trust in government. Tax administration may thus play a critical role not only in shaping economic development

[1] With the UN Millennium Project (2005), a less ambitious target was set according to which developing countries needed to mobilise only an additional 4 percent of GDP in tax revenue beyond their current average level of about 18 percent. However, very few countries (such as India) actually met that target.

but in developing an effective state". The tax system can play a key role in building and strengthening the relationship between citizens and State, in a context, like SSA, for too long associated with violence, crime, despotism and bad governance (Chabal and Daloz, 1999). By means of effective taxation, the State necessarily has to create robust public organisations, supported by a clear legal framework and adequate bureaucratic capability – accompanied by meritocratic hiring, higher professionalism, a less capricious enforcement and judicial system. In so doing, the State becomes accountable for its actions to the eyes of taxpayers, which in turn can raise their voice for better policy and take part in public debates. This tax bargaining process eventually paves the way for the construction of a fiscal social contract (Moore et al., 2018), in which taxpayers abide to the law and the State is considered as legitimate, accountable, responsive (Prichard, 2015). Cross-country descriptive evidence confirms this argument (Long and Miller, 2017).

As a second related point, taxation fundamentally relies on the coercive power of the State and its ability to establish law and order. As shown in Besley and Persson (2013), States with poor tax performance commonly fail to protect property rights effectively and are characterised by higher informality. Properly implemented taxation could, on one side, improve the functioning of financial markets and encourage formality and, on the other side, change the way coercive power is imposed – no more enforced capriciously, but based on the rule of law.

Third, better taxation has an intrinsic economic value in the way it reduces distortions in the economy between different categories of firms, which in turn reverberates on overall economic efficiency, societal fairness and structural growth. In this sense, reliance on income taxes is encouraged, since they are perceived to be fairer and more effective in achieving redistributive goals. Income taxes are also instrumental in improving a State overall fiscal apparatus, as they require an efficient tax administration, a smart enforcement strategy, a clever use of data. The reliance of a tax system on a progressive income tax is directly correlated with State tax capacity (Keen, 2012).

Lastly, developing countries seeking to collect their own revenue would ultimately depend less on international aid and donors – with the caveat that the effect of aid on tax effort is complex and not always obviously negative (Prichard et al., 2012). A decrease in aid-dependence translates into more control of a State's political agenda and increased accountability towards the citizens. The overall aid-related debate and donors' agenda would inevitably shift on different, probably more constructive, priorities.

Given the importance of taxation in a developing context, it is no surprise that increasing investment has been directed to reforming tax administrations. As a matter of fact, the ideal tax policy or reform is inevitably bounded by what can be implemented in practice.[2] This is truer in developing countries: as Casanegra de Jantscher (1990) famously put it, "in developing countries, tax administration *is* tax policy", meaning that what actually happens in terms of revenue collected depends as much on how the tax code is administered as on the provisions in the code itself (Keen, 2012). In light of the benefits that effective taxation can bring to developing countries outlined above, the very same benefits are likely to vanish if the way in which a tax system is administered is not fair, professional, transparent and predictable.

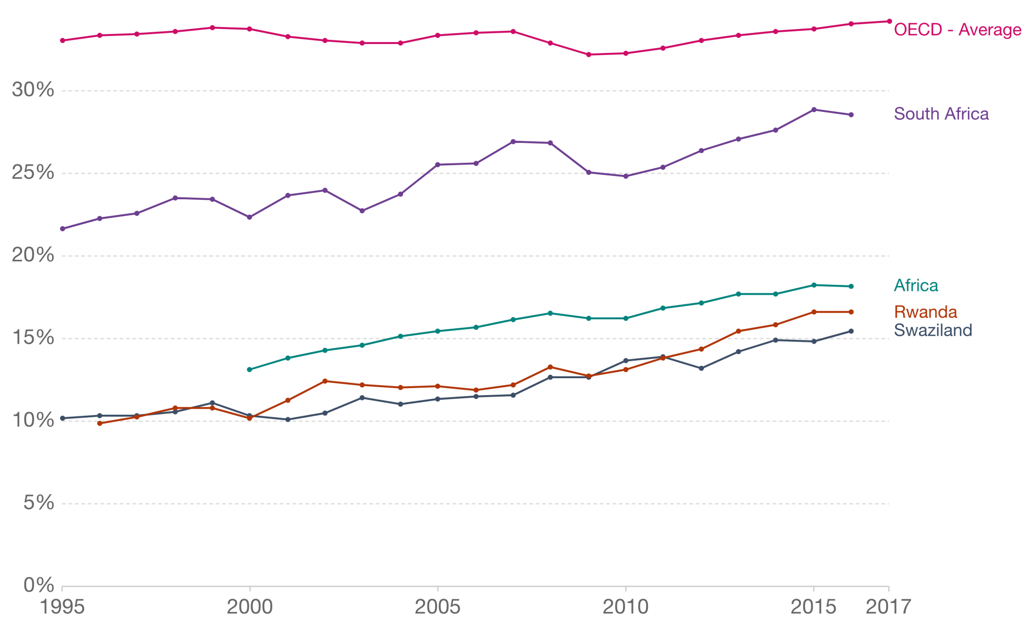
While fiscal capacity in developed countries enabled a series of tax improvements and instruments that facilitated raising revenue – such as withholding, third-party reporting or a complex-to-administer tax as the VAT (Pirttilä, 2017) –, SSA is still undergoing a period of transition (Moore et al., 2018). While there is no one-size fits all solution to the fundamental question on how to best reform tax administrations in such a diverse and complex environment as the African continent, some key ingredients are quite unanimously thought to be needed to modernise tax authorities in SSA. Apart from the political will and clear reform strategy which are at the linchpins of any structural transformation (Bird, 2015), more practical points, which will also be discussed throughout the thesis, are the following: (i) the simplification of procedures, regulations and practices (ii) a service-oriented approach in which the taxpayer is seen as client and not as a potential criminal (iii) the institutional reorganisation in order to work more independently and cost-effectively (iv) digitalisation, conceived as an opportunity to fundamentally rethink systems and procedures. This thesis attempts to provide scientific evidence on how effectively tax authorities could pursue these solutions.

Tax collection in Africa When starting to consider taxation and development in Africa, a first practical question one has to answer is: how is SSA performing in terms of revenue

[2] As Bird et al. (2008) observe, "the best tax policy in the world is worth little if it cannot be implemented effectively".

collection? The immediate reaction to this question is one of dismay. If one considers the tax raised over a country's GDP – usually called the tax take – as an indicator for fiscal performance, SSA raises about 15% of its GDP in tax revenues (ICTD/UNU-WIDER, 2020). This ratio is remarkably low when compared to the OECD economies and still short of the revenue target imposed by the SDG agenda. According to IMF (2019), SSA will need additional resources amounting to 19% of GDP to finance the SDGs by 2030. On top of that, the trend over time has been stable, with the SSA tax-to-GDP ratio rising by only 2 to 3 percentage points of GDP in the past two decades (Akitoby et al., 2019). Figure 1.1 displays the trend over time of the tax rate in SSA as compared to the OECD. It also displays the tax performance of the two countries under study (Rwanda and Eswatini), as well as of South Africa, which is clearly an outlier in the continent.

Figure 1.1: Tax revenue as a share of GDP, 1995 to 2017



Source: Ortiz-Ospina and Roser (2018) from Our World in Data

In reality, the situation is not as bleak as one may tend to think. First of all, while SSA scores poorly when compared to the OECD group, SSA tax agencies perform almost as well as their counterparts in the much wealthier environment of Latin America, and quite better than in South Asia (Moore et al., 2018). Second, contrary to general wisdom, SSA would perform relatively well when taking into account a different, less popular and documented, indicator of fiscal performance, the "tax effort" – the ratio between the actual revenues that a government collects and the potential one given the structure and features of the economy.[3] Thirdly and more importantly for this thesis, reliance on income taxes in SSA – a measure of tax performance, also indicating, as stated above, the equity and sophistication of a tax system – is reasonably high: in SSA, 31% of tax revenues are from direct taxes, compared to 35% in East Asia and Pacific region; 30% in Latin America; 25% in South Asia and the Middle East and North Africa. In SSA, income tax usually represents the second largest contributor – after VAT – to tax revenue (ATAF, 2017).

In sum, the overall picture of tax collection in SSA is much more nuanced than one would generally expect, especially in light of the abundant scholarship studying how poor State capacity and governance (of whom taxation is a core function) are in the continent (Bayart, 1993; Chabal and Daloz, 1999). The reality is different: since the 1990s, SSA countries are entering a tax transition process or a new *tax era* (Moore et al., 2018), in which modernisation of tax administration is key, and reached a point in which tax (or a mix of taxes) is the dominant source of public revenue, as it happens in the rest of the world. As stated in Moore et al. (2018), the core policy issue, for both donors and SSA policymakers, is not *only* about raising more taxes, but *also* about addressing the "current challenges around who pays taxes, how they are collected, and how governments use the revenue." If the SGDs cannot be financed adequately is still because SSA has not yet untapped the full potential of its tax system (Long and Miller, 2017). There exist several margins of improvements that would, at least partially, solve the factors which are still hampering tax performance. These factors can be summarised as:

1. Informality, which is not captured by taxpayer records and totally out of the legal tax system, is still rampant in the continent. According to Schneider and Medina (2018), the shadow economy amounted to about 40% of SSA's GDP in 1991-2015, compared to just 17% of OECD countries' GDP.[4] The existence of a *ghost*, informal

[3]According to the IMF (2011), the average tax effort for the 14 SSA (mostly anglophone) countries in the sample, excluding South Africa, was 75%. By contrast, the average for 6 Latin American countries was 59%, and for 4 South Asian countries it was 51%.

[4]Only Latin America has similar extent of informality over the period, 39%, while other regions in the

economy generates a number of horizontal inequalities and economic distortions that eventually undermine tax perceptions and the moral fibre of a society as whole.[5] In tax terms, income taxes are the most likely tax types to be underperforming due to informality. This aspect makes informality an ever bigger challenge to domestic revenue mobilisation, since income tax represents a large share of the revenue pot in SSA. Relatedly, most harm to revenue is produced not by micro-enterprises, such as street vendors, who operate out of government regulations, but rather from tax-minimising schemes of professionals and high-net worth individuals (doctors, lawyers, architects, etc.) which are equally hard-to-tax but determine much larger revenue losses and damage to the fairness of the tax system (Keen, 2012; Kangave et al., 2016). From a redistributive point of view, informality also severely undermines the validity of the optimal income tax approach for redistributive policies (Mirrlees, 1971) – so far widely adopted in developed countries –, as illustrated in Kanbur et al. (2018).

2. Despite the continued efforts of SSA tax authorities in registering income-generating activities, compliance of formal taxpayers is far from being optimal, as the following can be observed: (i) non-filers, i.e. taxpayers who are liable to file a return but systematically fail to do so, are widespread both in SSA[6] and in the countries under study – in Rwanda, over three-fourths of individuals supposed to file for the fiscal year 2018 failed to do so, with about half of companies doing the same, while the last 6 years average non-filing rate in Eswatini is 57% for individuals and 43% for companies (Chapter 2 and 3); (ii) nil-filers, i.e. taxpayers filing nil tax returns (zero turnover, taxable income and tax liability) are common as well[7] – in Rwanda (2013-2018), 53%

world fare better (South Asia 34%, East Asia and MENA 25%, Europe 22%).

[5] For a review on this topic, see Joshi et al. (2012)

[6] In Uganda, the average rate of Personal Income Tax (PIT) non-filing is 86% over the period 2014-2018. In Malawi, almost 50% of income taxpayers have filed no tax return and/or made no tax payment over the period 2014-2016 (Ligomeka, 2019a). Additional descriptive evidence from Kenya shows that out of the over nine million registered taxpayers, only 3.5 million filed their 2018 returns (see `https://www.businessdailyafrica.com/lifestyle/profiles/what-to-expect-file-nil-return/4258438-5232858-oiqmom/index.html`, accessed on June 10, 2020.) Lastly, Moore (2020) notes, non-filing rates in Nigeria are exceptionally high: 98% for PIT; 94% for Corporate Income Tax (CIT); and 95% for Value-Added Tax (VAT).

[7] In Ethiopia, about 23% of CIT returns filed in 2006/2013 are from nil-filers (Mascagni and Mengistu, 2016).

and 19% of Corporate Income Tax (CIT) and Personal Income Tax (PIT) returns are nil, while in Eswatini the corresponding figures are 29.5% and 26% (Chapter 2 and 4); (iii) imperfect compliance takes place also within active taxpayers, i.e. those reporting positive tax liabilities, but who are able to minimise their tax due either through avoidance or blatant evasion (Chapter 3 and 4).

3. Although SSA revenue authorities experienced important structural reorganisations and modernisation in their processes, mostly spurred by tax reforms initially motivated by external actors, the average enforcement capacity is still low (Chapter 2, 3 and 4) due to insufficient investment in specialised skills, qualified personnel, customer relations, ICT, research and data matching – such as use of third-party reporting to detect evasion (Pirttilä, 2017).[8]

4. Taxpayers' morale and willingness to pay is remarkably low (Chapter 3), especially when compared to outliers such as Nordic countries (Pirttilä, 2017). Afrobarometer surveys provide insightful evidence on how African taxpayers see the tax system they live in. Factors such as distrust towards a corrupt, rent-taking State (Bratton and Gyimah-Boadi, 2016; Isbell, 2017), dissatisfaction with how tax revenues are spent, so that citizens cannot realise the benefits derived from tax payments (Blimpo et al., 2018), perceived unfairness of the tax system (D'Arcy, 2011) and the complexity in navigating often obscure tax regulations (Aiko and Logan, 2014) all concur in affecting tax morale, which in turn is likely to deter compliance.

5. More macroeconomic factors are at play as well (Besley and Persson, 2014), which are however beyond the scope of this thesis: (i) the particular structure of the economy – less urbanised, marketised and wealthy, more reliant on hard-to-track agricultural activities and with a lower international trade to GDP ratio; (ii) reliance on easy-to-tax natural resources, especially oil and minerals, where taxation can use royalty

Likewise, in Uganda 27% of PIT returns are nil over the period 2013-2018 and, according to Almunia et al. (2017), 15% of VAT returns in 2012-2015 are nil.

[8] Related to the last point, revenue authorities are undergoing a process of digitalisation for which coherent evidence is still missing. This thesis attempts to contribute to this point by providing convincing evidence on the benefits of systematic data analysis on the massive amounts of administrative data produced daily within revenue auhorities.

payments which act as a substitute for direct taxation; (iii) reliance on external aid, which also substitutes for internally generated tax revenue; (iv) the political economy of most SSA countries, in which vested interested of wealthy individuals and companies are lobbied by politically influential groups and more equitable tax reforms are harder to implement. On top of that, Governments use their ability to grant tax exemptions as a direct instrument of rule and political support, while democratic negotiation and collective action are undermined (Moore et al., 2018).

For the sake of this thesis, point 5 above is out of the scope of my work, while point 1 is touched only tangentially. Points 2, 3 and 4 are directly addressed. More specifically, the next subsection will deal with the behavioural drivers of compliance (point 4) more in depth. While this paragraph focused on tax administrations, the next one will provide a more thorough discussion on the other side of the compliance equation – the taxpayers.

Why do African taxpayers comply? A *behavioural* solution Against the background delineated above, it is crucial to understand why African taxpayers choose to comply with or escape from tax obligations. It is also worth stressing that tax compliance decisions of income taxpayers are particularly interesting from both a theoretical and practical point of view, since tax evasion from these agents is particularly difficult to uncover given that this group has a higher economic incentive to underreport income to reduce their tax liabilities (Allingham and Sandmo, 1972; Slemrod and Yitzhaki, 2002; Sandmo, 2005). The challenge of taxing income is even more exacerbated in a context of low tax capacity as in SSA.

The theoretical formulations produced in the literature provide a robust framework in which to conceive the taxpayer's decision. Two broad branches are usually referred to. On the one hand, the neoclassical theory of Allingham and Sandmo (1972) formulates that taxpayers are self-controlled, fully rational, utility-maximising agents and, therefore, all potential evaders. As in a gamble, taxpayers compare the benefits from evading (lower tax paid) with the potential costs (the probability of getting caught and punished). Only pecuniary motives are at play: tax rates, audit probabilities and fines.

On the other hand, the neoclassical model soon appeared insufficient to explain the complexity of the taxpayer decision-making process. The model could not justify the observed rates of compliance, except by using unacceptably high measures of risk aversion (Clotfelter, 1983; Andreoni et al., 1998). Given that more taxpayers comply than what predicted by the theory, scholars incorporated in the model a range of new, non-pecuniary factors in order to explain the observed behaviour. At least two sets of non-pecuniary factors are explored in this thesis and described in detail in Chapter 2 and 3:[9] (i) the complexity of navigating the tax system, a broad concept that encompasses compliance costs and tax knowledge, (ii) behavioural motives that usually go under the umbrella term of *tax morale* (Luttmer and Singhal, 2014) and include perceptions on the fairness of the tax system, a reciprocity mechanism through which taxpayers pay taxes if they get something in return, intrinsic attitudes to comply, the effect of peers' behaviour and descriptive norms around compliance, perceptions on State legitimacy and trust towards the authority, etc.

Only recently the main behavioural formulations have been embedded in a conceptual framework to understand tax compliance (Prichard et al., 2019), which however needs to be tested more coherently in the field, as this thesis attempts to do. The framework encourages revenue authorities in low-income countries to pursue a multi-pronged approach to foster compliance, consisting of three main paradigms: (i) the enforcement paradigm, which directly links to the neoclassical model of Allingham and Sandmo (1972), (ii) a facilitation paradigm – as in Alm (2012) – in which communication, simplification, assistance, tax education and facilitation are offered to the taxpayer-client, (iii) a trust paradigm, in which non-pecuniary, soft, factors such as professionalism, trustworthiness, transparency, wise use of revenue, rewards for honesty and social norms are all leveraged by tax agencies in order to encourage voluntary compliance.

This thesis contributes to the main gap in knowledge according to which very little is known about why African taxpayers remit their taxes. The majority of existing studies are descriptive in nature, focus on survey data only and are not backed by administrative data. Therefore, the correlational and causal evidence produced in this thesis is relevant to both scholars and SSA tax administrators. For what concerns the latter, this thesis confirms the adoption of a multi-faceted strategy, as suggested in Prichard et al. (2019), in which the *behavioural* solution plays a greater role. In addition to the traditional enforcement

[9] Additional, more theoretical, factors pertain to deviations from utility maximisation and non-standard preferences, thus affecting how the modelling problem is set up. These factors are only briefly mentioned in Chapter 2 and are out of the scope of this study.

paradigm, this thesis makes the argument in favor of the role of both facilitation and trust in understanding compliance.

1.2 Approach

Working with local tax authorities This thesis would not have been conceivable without the support of the revenue authorities in Rwanda and Eswatini, as the only way this research could be carried out has been through the joint work with tax administrators.

The three chapters presented in this dissertation take part to the recently inaugurated wave of the public finance literature which builds on close research collaborations between economists and tax administrators from the South. This new wave of tax research produced a remarkably insightful evidence based on a wealth of new, previously inaccessible, administrative data. It also carried innovative and more ambitious research questions who advanced the current debate on tax and development.

Throughout my work in Rwanda and Eswatini, the collaboration followed a similar path in both countries, which largely resembles the process described in Pomeranz and Vila-Belda (2019). After getting institutional buy-in and establishing a formal relationship,[10] I devoted much effort in running exploratory, qualitative analysis and piloting the intervention. This preliminary work greatly improved my understanding of the context and generated mutual trust with my partners.[11]

A following important phase consisted in accessing a large number of administrative

[10] High-level commitment is officially reached through the signature of a Memorandum of Understanding between the partnering revenue authorities and the International Centre for Tax and Development (ICTD), for which I work. The MoU signing undoubtedly helps in formalising the research collaboration and aligning the (apparently divergent) partners' interests over a long-term period. In most cases, the primary goal of revenue authorities is revenue maximisation. This goes often in conflict with the broader research perspective. As shown in Chapter 2 and 4, both the tax trainings in Rwanda and tax nudges to non-filers in Eswatini significantly improved the likelihood to file a return, a meaningful result per se in a context where non-filing is the norm, while had no direct impact on revenue generation.

[11] This also helped sensitising my partners on the value of research and pave the way for the introduction of more sophisticated design aspects, such as the implementation of an RCT, which can be seen with suspicion and resistance by tax administrators.

data, signing a legal agreement with the authority and understanding the content of many different datasets. In this phase, multiple iterations with local officials are needed, usually travelling to the partner country, since tax officials have deep knowledge of the meaning of each tax return item. Extracting clear knowledge and insights for interpretation from often obscure raw administrative data is a key component of the *plumber* approach for development economists, as outlined by Duflo (2017).

Subsequently, the most delicate phase probably consists in designing an effective experiment by trading-off a high level of flexibility – in order to meet the partners' expectations – with the need of implementing a robust and clean project, in which causal effects can be measured. Repeated inputs and feedback from both senior and junior staff are incorporated in the final design. If the project has been designed carefully, the implementation of it is likely to proceed as planned, even if inconveniences are likely to raise anytime. After the project has been implemented, it is the role of the researcher to produce results using the most advanced econometric tools.

As a last step, which is still ongoing in this case, it is crucial to communicate the research findings to the main stakeholders, with the ultimate goal of leading government partners to policy changes.

This thesis makes an argument in favor of setting up research relationships with local partners in Africa. As a key outcome of this approach, this type of collaboration often translates into fruitful, mutual learning. On one side, international researchers build capacity of local tax officials, who have often little time to devote to research, with the hope that they will eventually end up owning the research output.[12] On the other side, economists are introduced by tax officials to the peculiarities of the local reality they intend to study.

Administrative and survey data As repeatedly emphasised throughout the dissertation, measuring tax compliance is not an easy task. In the innovative attempt to accurately capturing tax evasion, the approach adopted in this thesis consists in combining quantitative and qualitative methods, with the overarching idea of reaping the benefits of both administrative and survey data, while controlling for the potential pitfalls in each of them.

On the one side, all three chapters of this thesis rely heavily on a wealth of de-identified

[12] A key challenge is represented by teaching often sophisticated data analysis and econometrics to staff from tax authorities, since it would then be hardly applied in their everyday job.

tax records shared by the partnering agency. Administrative data is produced on a daily basis and in massive amounts, often difficult to manage even by the collectors (the revenue authorities) themselves. At this level, the collaboration with data experts and economists offers remarkable synergies and opportunities for mutual learning. This data comes from all different types of sources, from companies filing their income tax return to VAT payers, from border trade posts to employers remitting PAYE on behalf of their employees. Even if this data is gathered for administrative and enforcement purposes, it provides exceptional opportunities for research and has gained momentum in the last decade. Most importantly, revenue authorities from the South, and SSA in particular, began to share their tax records with international researchers, spurring a new wave of tax studies in the Continent (Mascagni et al., 2016), to which this thesis aims to contribute.

On the other side, survey data produces extremely valuable information on certain features, such as tax attitudes and perceptions, which cannot be captured otherwise. The contribution of this thesis stands in the attempt to merge survey data with tax records by the use of unique identifiers. In so doing, I am able to observe both a taxpayer's filing history from tax records and a range of background information that are likely to explain that particular filing decision. This synergistic combination of different data sources takes place in Chapter 2 studying the tax trainings in Rwanda (see section 1.4), where survey data provides insights on the mechanisms likely to explain the observed compliance outcomes after the training, as well as in Chapter 3, where non-filers and active taxpayers in Eswatini, as extracted from official tax records, are explored through the analysis of detailed survey data. Lastly, in Chapter 4, the same survey data collected in Chapter 3 serves the purpose to shed more light on the impacts of a nudge experiment.

Both sources of data presents important drawbacks which are discussed at length in Chapter 3.[13] It is notoriously difficult to capture tax evasion from surveys (Pirttilä and Tarp, 2019) and scholars, as well as international organisations, have used more or less justifiable proxies for identifying compliant taxpayers. However, as elaborated more in detail in Chapter 3, these measures are rarely satisfactory. On the other hand, it is also true that administrative data do not capture the informal, not registered, economy and

[13]Relatedly, in that Chapter, I am able to show how divergent self-reported measure of compliance can be when compared with actual filing status from official records.

report only the taxable income that registered taxpayers decide to disclose. Despite these limitations, this thesis points towards a strengthened reliance on this type of data. As a supporting example, the following chapters show that registered, formal taxpayers who systematically fail to file (non-filers) show a profile – in terms of business practices, record-keeping, IT sophistication, tax awareness, etc. – which is more similar to small, subsistence-level informal entities than to active taxpayers, suggesting that researchers can learn about this type of informality indirectly through the study of chronic non-filers.[14] At the same time, another supporting argument is given by the fact that taxable income itself – as derived from administrative data –, even though self-declared, is a key outcome per se that can be shaped through different strategies and respond to a full-house of motivations whose understanding is at the core of most tax evasion research.

In conclusion, there is much to learn from this mix of quantitative and qualitative methods. The approach I follow in this study is ultimately motivated by the desire to produce evidence-based policy recommendations to enhance domestic revenue mobilisation from revenue authorities in SSA. To this aim, the evidence here produced necessarily had to be extracted from and shed light on the local realities of taxation in the countries under study, which are outlined next.

1.3 Rwanda and Eswatini

The choice of Rwanda (a low income country, ranking 167^{th} in the world for GDP per capita in 2019[15]) and Eswatini (a lower-middle income country, 118^{th}) as case studies has been motivated by a number of reasons.

First, strategically, the two countries became partner of the International Centre for Tax and Development (ICTD) in 2015 (Rwanda) and 2018 (Eswatini) after repeated informal consultations and meetings with the revenue authorities' senior management. Second

[14]In this particular case, by informals I refer to small, often subsistence-level, income-generating activities while I exclude professionals and other high-net worth individuals, who are equally likely to trade outside of the legal tax system.

[15]See World Bank webpage `https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?most_recent_value_desc=true` - accessed on August 14, 2020.

and relatedly, the two countries enjoy the service of two professional, service-oriented and modern revenue authorities, with whom the process of doing research has evolved smoothly and stimulatingly. Third, the two countries are similar in size and geography – they are both landlocked – even if Rwanda has a ten times larger population. The fact that the countries are small in size significantly facilitated doing research in the field.

Apart from these broad similarities, the two countries also share some common features of the tax system, which relies mostly on non-oil tax revenues. Above all, the trend of the tax revenues over GDP is similar, with the tax/GDP ratio hovering around 15 per cent in both countries (Table 1.1 below). While Rwanda's ratio is line with the East Africa average, Eswatini's is quite lower than the Southern Africa average, indicating more severe challenges with revenue mobilisation.[16]

As a second point in common, income taxes, the focus of this thesis, represent about 40 per cent and 35 per cent of Rwanda and Eswatini total tax revenue (ICTD/UNU-WIDER, 2020), hence their policy relevance in both countries. Despite being important in relation to total revenue, the share of income taxes over GDP is still low in both countries – and in Africa as a whole. Rwanda and Eswatini collect about 7-8 per cent of GDP in income tax, which is much lower than how the Continent outlier South Africa and OECD countries perform, as shown in Figure 1.2.

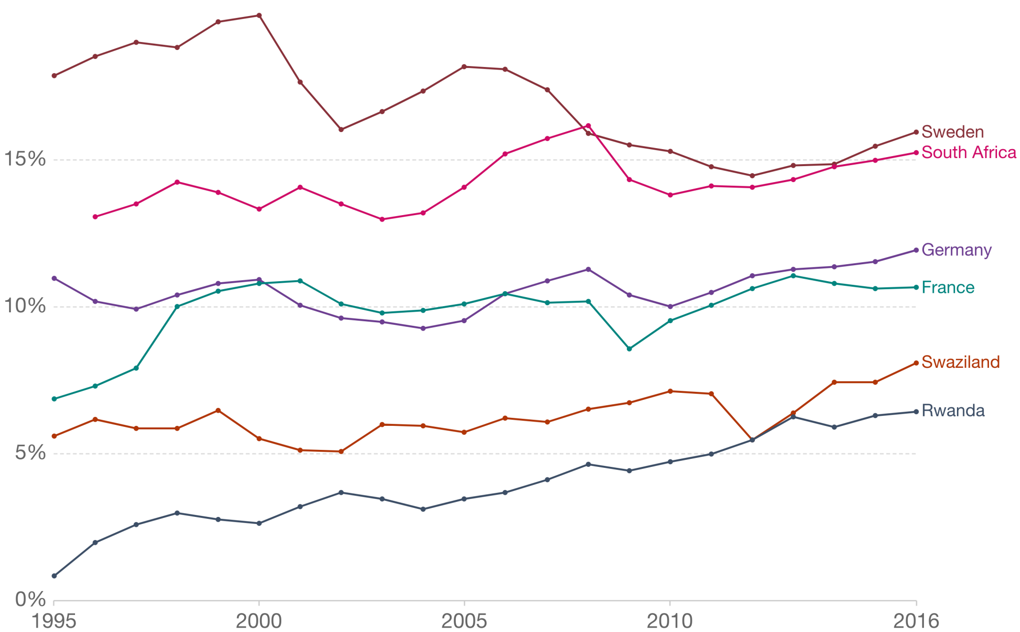
Thirdly and relatedly, similar alarming patterns of compliance with income taxes can be found in both countries. The extent of non-filers is widespread, often being the majority in Rwanda and Eswatini (and other SSA countries as well) and also nil-filers are quite common, as highlighted in section 1.1 and explained in detail throughout the thesis. Lastly and explaining the low take of income taxes, the informal sector is widespread in both economies and stands at about 36-40 per cent of national income (Schneider and Medina, 2018).[17]

Apart from these broad similarities, the institutional context of the two countries is quite different. Country indicators are reported in Table 1.1 below, which is also discussed

[16]While Rwanda is booming in terms of economic growth (10 per cent in 2019), Eswatini projected growth stands well below, at 1.3 per cent for 2019. About 40 per cent of the population in both countries live under the international $1.90 poverty line.

[17]As with the tax/GDP ratio, while Rwanda is in line with the regional average, informality in Eswatini is much higher than in the corresponding region.

Figure 1.2: Tax on incomes of individuals and corporations as a share of GDP, 1995 to 2016



Source: Ortiz-Ospina and Roser (2018) from Our World in Data

at length in Chapter 3. In a nutshell, Table 1.1 indicates that Eswatini scores worse in terms of most governance indicators. The Corruption Perception Index, which ranks countries based on how corrupt a country's public sector is perceived to be by experts and business executives, is lower in Eswatini than in Rwanda, meaning a higher perceived corruption in the former than the latter. In contrast, Rwanda's public sector is perceived to be much more transparent than the East African average. Consistently, Eswatini performs worse than Rwanda in terms of the World Bank Governance indicators and the Index of Economic Freedom. This reflects a situation in which institutions in Eswatini are perceived to be less effective, reliable and accountable. This divergence will resonate in the survey findings extracted from both countries in Chapter 2 and 3 – while perceptions of government authority, fairness of the tax system and satisfaction with public services are extremely high in Rwanda, there is much more discontent and mistrust in Eswatini. Lastly, Rwanda ranks higher in terms of the WB Doing Business indicators, which also includes proxies for the compliance costs of firms, a concept that will be explored at length in Chapter 2 and

3. Relevantly for this study, while Rwanda ranks 38^{th} in the world for the ease of doing business, Eswatini is 122^{nd}; when considering the ease of paying taxes, however, Eswatini fares slightly better (77^{th}) but still below Rwanda (38^{th}), which is also the second best rating in SSA, following Mauritius.

In sum, the two countries represent a unique scenario – both in terms of successful collaboration with the tax agency as well as for the institutional background and compliance patterns they feature – to gain a deeper knowledge of the role of taxation in a developing context. More detailed information on the two contexts will be provided in the main Chapters.

1.4 A bird's eye view of the thesis

This thesis is composed of three Chapters, from 2 to 4, while Chapter 5 highlights some conclusive arguments and directions for future research. The dissertation begins in Chapter 2 with the impact evaluation of a tax training program implemented by the Rwanda Revenue Authority, which aims at improving the knowledge of the tax system of newly registered income taxpayers. This chapter represents a longer, more elaborated version of the working paper published with the ICTD (Mascagni et al., 2019) and is a joint work with Giulia Mascagni (ICTD) and the RRA research team. The main research question I aim to address is whether and how the provision of tax education shapes tax compliance, as well as tax knowledge and perceptions. This work contributes to the literature in different ways. First, to the best of my knowledge, it represents the first study of its type exploring the link between tax education and tax compliance in a low-income country (Rwanda). The evidence on this link is limited and exclusively deriving from high-income countries (Chetty and Saez, 2013). Second, this study combines pre- and post-training survey data of 1,000 taxpayers with tax returns records. As discussed in section 1.2, this approach enables tax researchers to overcome most of the shortcomings of both survey and admin data. Third, this study contributes to the current understanding of tax compliance by considering its multi-dimensionality and, in the same vein, it implements a quite detailed survey module on tax knowledge. The results show that taxpayer education significantly increases the

Table 1.1: Governance and country indicators - Eswatini vs Rwanda

	Eswatini	Southern Africa	Rwanda	East Africa	Year
Tax to GDP ratio[a]	14.7%	22.3%	15.3%	13.8%	2018
Tax revenue per capita (USD)[a]	444	949	106	112	2018
Informality (% national income)[b]	40.1	32.3	36.3	40.3	2005-2015
CPI[c]	39	47	55	33	2017
Governance indicators[d]					
Control of corruption	-0.44	0.18	0.69	-0.59	2016
Rule of law	-0.32	0.10	0.07	-0.50	2016
Regulatory quality	-0.58	-0.07	0.11	-0.33	2016
Government effectiveness	-0.56	-0.08	0.11	-0.54	2016
Political stability	-0.49	0.19	-0.05	-0.592	2016
Voice and accountability	-1.42	0.06	-1.21	-0.75	2016
Index of economic freedom[e]	55.9	60.2	69.1	59.3	2018
Tax burden	74.8	64.9	75.8	75.6	2018
Government integrity	27	41.4	61.2	35	2018
Judicial effectiveness	35.3	52.6	79.6	44.1	2018
Business freedom	61.1	63.2	81.2	51.8	2018
Doing business indicator[f]	59.5	62.3	76.5	59.3	2018
Starting a business	77.2	79	93.2	81.6	2018
Registering property	60.8	57	93.7	63.1	2018
Paying taxes	77.1	76.2	84.6	69	2018
Bank account ownership[h]	29%	42%	38%	39.5%	2017

Southern Africa: Botswana, Lesotho, Namibia, South Africa and Eswatini.
East Africa: Burundi, Kenya, Rwanda, Tanzania and Uganda.
[a] Annual Reports of the SRA/RRA
[b] Schneider and Medina (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.
[h] Global Findex (2017). Adults (+15 yo) in labor force. Burundi excluded.

filing probability of attendees, in a context where non-filers with income tax are the norm, as outlined in section 1.1. The most plausible mechanism consists in the improvement of tax knowledge and perceptions of complexity, while other perceptions and attitudes are not affected. In sum, this study makes the argument for a more pronounced focus on taxpayer education as a key driver of compliance, especially so in a continent where most tax authorities are increasingly implementing a variety of taxpayer educational initiatives (Mascagni and Santoro, 2018).

Chapter 3 focusses on Eswatini and attempts to shed light on the often neglected behaviour of tax non-filing. By surveying a representative sample of one thousand sole traders registered for personal income tax, evenly split in non-filers and active filers, this study tries to understand which are the behavioural drivers of tax filing. The modus operandi mirrors the one adopted in Rwanda. First, this paper stems from a close research collaboration with the Eswatini Revenue Authority (SRA). Second, it merges survey data with the filing history (2013-2018) of the sample, as derived from income tax records, and gives particular attention to a mistakenly under-researched aspect, i.e. a taxpayer's past filing behaviour – by creating the category of *perpetual* non-filers. Third and relatedly, it contributes to the current knowledge on how to measure tax compliance by comparing self-reported measure of compliance with actual filing behaviour. Lastly, it tests a range of theoretically driven explanation for compliance, ranging from neoclassical to more behavioural formulations. Results show that economic deterrence, compliance costs and moral factors are strongly correlated with actual filing, supporting the framework of Prichard et al. (2019). At the same time, drivers of actual compliance are quite divergent from those correlated with self-reported willingness to pay. Consistently with what found in chapter 2, this study as well argues that tax knowledge plays a major role in explaining the decision to file.

Lastly, the thesis ends with Chapter 4, always set in Eswatini. In collaboration with the SRA and the national post office, this study implements a nation-wide randomised controlled trial nudging more than 20,000 income taxpayers with behaviourally-informed mailings. Following the same broad rationale as Chapter 3, this study attempts to answer the relevant questions on which are the drivers of African taxpayers' compliance and how can these be leveraged by resource-constrained tax authorities. While the tax nudges literature has boomed in HICs and Latin America, only a handful of studies can be found in SSA (Shimeles et al., 2017; Mascagni et al., 2017, 2020). At least three main contributions can be identified. First, following the *economists as plumbers* approach of Duflo (2017), this study builds on the research partnership with local policymakers to evaluate actual policies and highlights margins for policy improvement that diverge from textbook models of tax compliance. Second, thanks to the wealth of administrative data provided by the

SRA, this study is the first of its kind in targeting three different categories of taxpayers at the same time – non-filers, nil-filers and active (section 1.1) – while most of the existing literature focusses on positive filers. Relevantly, I tailor the content of letters to be specific to each taxpayer category. Third, I am able to target both companies and individuals and explore heterogeneity of results along a number of dimensions, including past filing behaviour. I find that non-filers significantly respond to the nudges, while nil and active filers do not. The best performing nudges build on the deterrence and taxpayer-assistance paradigms (Prichard et al., 2019). Also, perverse responses from large companies are found. With the causal evidence produced, I am able to produce policy recommendations on how to best target the multidimensional ecosystem of income taxpayers.

Chapter 2

Teach to Comply? An Experimental Evaluation of a Taxpayer Education Program in Rwanda

Abstract

The role of taxpayer education in improving tax compliance has been largely unexplored in the literature. This paper starts filling this gap by providing the first ever rigorous evaluation of the effectiveness of taxpayer education on knowledge, perceptions, and compliance. The study is an impact evaluation of the Taxpayer Training Programme run by the Rwanda Revenue Authority and targeting newly registered taxpayers. The analysis is based on a unique dataset that combines administrative and survey data. I show that taxpayer education results in significant and large increases in knowledge, which starts from a very low level at baseline. I also show that the program contributes to improving compliance behaviour. The results show that training new taxpayers helps bringing them into the habit of filing tax declarations, in a context where non-filing is widespread. In terms of policy, results show that the benefits of taxpayer education go beyond increased revenue in the short term, and include building a habit of tax compliance.

2.1 Introduction[1]

As discussed at length in Chapter 1, tax collection in low-income countries is severely underperforming (Akitoby et al., 2019). Sub-Saharan Africa (SSA) is not an exception, with direct negative consequences on the resources available to finance development (IMF, 2019). The tax revenue to GDP ratio, usually seen as an indicator for the tax effort in a country, in SSA is 20 to 30 percentage points less than that of Europe and North America.[2] On top of that, the trend across country income categories has been stable for decades (Cottarelli, 2011). This gap in tax capacity translates in fewer resources available to provide basic public goods and services, incentivise investments or fund anti-poverty programs. Consequently, without a reliable and sustained stream of tax revenues, SSA countries are still depending on external aid, which most often acts as a substitute for properly raised domestic revenues. In the country under study, Rwanda, in the last two years foreign aid has been almost equal to total revenues collected.[3]

In this context of weak fiscal capacity, it is crucial to understand which strategies SSA tax authorities could pursue to raise domestic tax revenues. While deterrence has historically been the main tool for inducing compliance, alternative service-oriented solutions are gaining momentum among SSA tax agencies (OECD, 2015; Mascagni and Santoro, 2018). Among them, initiatives providing tax education are becoming more and more popular. Tax knowledge and, more in general, compliance costs recently received attention as key explanatory factors of poor compliance (Eriksen and Fallan, 1996; Alm et al., 2010). In the context of taxpayer confusion, compliance costs refer to the cognitive effort that taxpayers make in understanding complex tax systems, and to the administrative costs they incur in fulfilling their tax obligations. Taxpayers often have a very limited understanding of the tax system. As a result, they behave in ways that are inconsistent with economic theory, and with the incentives set out by policymakers (Feldman et al., 2016). However, as elaborated more in section 2.2, the role of tax knowledge in explaining actual tax compliance has not yet been rigorously tested so far.

[1] This chapter is an extended version of a paper written with Giulia Mascagni and published as an ICTD working paper (Mascagni et al., 2019).

[2] For a visualisation of tax/GDP ratios worldwide, see Figure A1 in the Appendix.

[3] Appendix Figure A2 shows the trend of aid and revenue collected over time.

This paper aims to address this gap by answering the following question: can taxpayer education affect compliance? To address this question I attempt to measure the effectiveness of a taxpayer education program, run every year by the Rwanda Revenue Authority (RRA), on tax compliance, as well as on tax knowledge and perceptions. The program is targeted at newly registered income taxpayers and is specifically aimed at helping them to comply as they enter the tax system, with a focus on learning and setting good compliance habits right from the start. In addition to the training program, a novel educational strategy in the form of a personalised tax coaching provided by RRA officials is piloted and evaluated (section 2.3.3). To the best of my knowledge, this is the first study to explore the link between knowledge of the tax system and tax compliance in a developing country.

The context, Rwanda, is typical of a low-income SSA country, both in terms of poor tax compliance (section 2.3.2) and low literacy rates (73%) when compared with high income countries (almost 100%),[4] despite being considered a success story in terms of economic and social development. African countries, Rwanda included, also face very low tax knowledge, to the extent that the majority of Africans do not know what taxes they owe to the government or what tax payments are for (Aiko and Logan, 2014; Isbell, 2017). These facts make the analysis both urgent and relevant on top of being novel, given the near absence of other studies in this area (section 2.2).

In addition, income taxes are economically significant in the Rwandan context (and SSA as a whole) – the country collects a third of the total tax revenue from this type of tax (ATAF, 2017). However, filing of income taxes is far from being optimal. Apart from informal entities, who are totally out of the tax system,[5] filing behaviour of registered taxpayers is inadequate. It is important to consider filing behaviour in its inherent multi-dimensionality (Slemrod, 2019). In order to do that, I consider three important filing choices: the probability to declare, the probability to nil-file, and the tax amount remit-

[4]The corresponding average figure for low-income countries is 63%. The figures for Rwanda and LICs are from the World Bank's World Development Indicators (`https://data.worldbank.org/indicator/SE.ADT.LITR.ZS` accessed on 21 January 2020), while the figure for high income countries is from Our World in Data (`https://ourworldindata.org/global-education` accessed on 21 January 2020).

[5]In Rwanda, informality is rampant, representing 36% of national GDP in 1991-2015 (Schneider and Medina, 2018), indicating, on the one hand, a situation of lack of information on the tax base and weak monitoring power of the tax agency (Porta and Shleifer, 2016) and, on the other hand, a common norm of tax evasion among the population.

ted. All these choices may imply tax avoidance or evasion: while non- and nil-filing have immediate negative repercussions in terms of impaired domestic revenue generation, tax avoidance usually takes place also among positive filers. The presence of non- and nil-filers is sizeable in Rwanda, where only a portion of newly registered taxpayers remit positive taxes every year (section 2.3.2). Non- and nil-filing is the norm also outside Rwanda. Evidence on this filing patterns come from the United States[6] and Latin America.[7] Also the SSA region experiences both failure to file[8] and nil returns.[9]

The research design builds on the collection of survey data from a randomly selected group of one thousand taxpayers before and after the training. Survey data provides a rich source of information on the level of tax knowledge – thanks to a detailed quiz-like section on tax – and perceptions towards the tax system, which are investigated as the key mechanisms of impact at play. In terms of the main outcomes, I observe taxpayers' filing behaviour thanks to administrative data from tax returns provided by the RRA (section 2.4.1). Using different methods to address self-selection into the training, which I discuss in detail in section 2.4.3, I measure the causal impact of the training program on tax knowledge and perceptions as well as tax compliance.

Results in section 2.5 show that trainees perform significantly better in terms of filing behaviour. The training is explaining a 15 percentage points (or 43% of the control group mean) increase in the probability to file a tax return (extensive margin), despite not inducing taxpayers to remit more taxes (intensive margin). The impact at the extensive

[6]Erard et al. (2018) estimate the share of ghosts for the US federal individual income tax for the tax years 2000-2012 to be 7.1% of the total number of returns expected to be filed. In the same fashion, looking at the city of Detroit's individual income tax for the tax year 2014, Meiselman (2018) estimates that 48% of returns were still missing as of April 2016, two years later.

[7]In Guatemala, the share of non-filers of income tax in 2013 is of 39%. In Costa Rica, Brockmeyer et al. (2019) estimate that 25% of tax-registered firms and 19% of self-employed did not file an income tax declaration in 2014 as of February 15, 2015, two months after the filing deadline. The incidence is lower but still substantial for sales tax: 14% of firms and 19% of the self-employed. In a companion paper, Brockmeyer and Hernandez (2016) find that about 50% of tax-liable firms fail to file their income tax declaration for the period 2006-2014. Moreover, 20% of the businesses registered for the municipal income tax in a municipality of Venezuela's capital did not have any fiscal activity for at least 3 years (Ortega and Scartascini, 2016b).

[8]See footnote 6 of Chapter 1.

[9]See footnotes 7 of Chapter 1.

margin of compliance is consistent across three different econometric methods. This result is economically significant given a context where every year about half of registered taxpayers fail to file. Bringing taxpayers into the tax system, who in turn share information and remain visible to the authority, is a main goal of any tax agency. Moreover, filing tax returns could become a virtuous habit and a learned responsibility (Dunning et al., 2017), most importantly for small taxpayers who usually are less likely to file a return. There is no significant impact in terms of the amount of tax declared, largely due to lack of power and self-selection. Always due to low sample size and relevant implementation issues, the coaching intervention does not prove to be effective (section 2.5.3).

In terms of mechanisms (section 2.6.1), the most plausible one is an increase in taxpayer knowledge and a decrease in the perceived complexity of the tax system – both of which are directly related to the central themes of taxpayer confusion and compliance costs. Given the very little evidence on the role of tax knowledge and compliance costs in motivating compliance, this result is a key contribution of the paper to the existing literature, as explained more in depth in section 2.2.4. I trust that future exciting work will be carried out in this direction.

2.2 Literature Review

This study aims at producing new evidence on the determinants of tax compliance by considering a largely neglected factor, tax knowledge. The literature on the drivers of compliance is abundant and composed by different branches, which are summarised below.

2.2.1 Neoclassical standard theory on utility maximization

The workhorse model of tax compliance is the one by Allingham and Sandmo (1972) – AS from now on – who builds on the economic theory of crime (Becker, 1968) and uncertainty (Mossin, 1968).[10] In the AS setting, taxpayers are modelled as self-controlled, fully rational utility maximizing agents and all potential evaders. They face the decision to comply under

[10]See also Srinivasan (1973) for an early model of compliance.

risk, as in a gamble: the agent compares the benefits from evading (lower tax paid) with the potential costs (the probability of getting caught and punished). The main finding from this utility maximisation calculus is that the level of evasion is negatively affected by deterrence factors, such as audits and fines.

In a following contribution, Yitzhaki (1974) shows that, when the fine is assessed on the evaded tax rather than on the undeclared income as in AS, higher tax rates will lead to lower level of evasion, at odds with most empirical evidence (Clotfelter, 1983; Andreoni et al., 1998). For this and other reasons discussed below, the AS model is unanimously considered as incomplete and unsatisfactory.

2.2.2 Deviations from the standard approach

The neoclassical model has been often criticised for relying on assumptions that are generally unrealistic for explaining a compliant behavior (Andreoni et al., 1998). First, taxpayers are assumed to have full knowledge of all the key policy parameters, such as the audit risk and the size of penalties.[11] Second, taxpayers' response is predicted to be the same across all of them. Third, risk preferences are assumed to be identical across taxpayers. It results that, according to the assumption of risk aversion in the AS model, evasion should be much more than what is actually observed in the real word. In reality, the probability of getting caught is often very low and varying by type of income, while expected fines are rather small, both for political motivations and because quite infrequently imposed. Audits are even more unfrequent in administrations with inadequate staff and budget as in low-income countries.[12] For these reasons, the AS model requires coefficients of risk aversion to be much higher than normal to justify the higher rate of compliance observed (Alm et al., 1992).[13]

Alternative theoretical formulations enriched the standard model with new parameters that attempt to explain the high observed levels of compliance. The term "tax morale" has been used as an umbrella term capturing the non-pecuniary drivers of tax compliance

[11]This paper mainly aims at questioning this assumption.

[12]In Rwanda, the revenue authority audited less than 300 taxpayers in 2015-16 (Mascagni et al., 2016).

[13]For example, the estimated Arrow-Pratt measure of risk aversion in the United States is between 1 and 2, when it should be around 30 to reach the observed tax compliance level.

who are not accounted for in the AS framework (Luttmer and Singhal, 2014).[14] The theoretical contributions of this alternative approach are reported in Appendix section A1.1. Throughout the paper, these aspects will be touched only tangentially, while the main focus is directed on an unduly neglected factor driving compliance, i.e. tax knowledge.

2.2.3 The role of tax knowledge and compliance costs

This paper contributes to the scant literature on the role of tax knowledge in shaping compliance. The only study on this topic evaluates the effect of a 2-minute explanation provided by tax preparers to taxpayers who are eligible for the Earned Income Tax Credit (EITC) in the United States (Chetty and Saez, 2013). The authors find no effect on the amount of income reported, which is the only outcome evaluated. While this paper represents a key reference point for this study, it focuses on a relatively quick (2 minutes) and focussed (specifically on the EITC) intervention.[15] Also, Chetty and Saez (2013) consider a high-income country, which is very different – for example, in terms of literacy rates in the population – from the context of this study.

A number of studies have included general education as a background variable, assuming that tax knowledge is increasing with literacy (Vogel, 1974; Spicer and Lundstedt, 1976; Song and Yarbrough, 1978; Kinsey and Gramsick, 1993). More recently, knowledge about tax has been shown to promote positive tax ethics. Roberts et al. (1994) study how the general lack of understanding of the concept of progressive taxation is related to tax attitudes. When asked a general question on tax preferences, the majority of students are in favour of a progressive structure. However, when asked in concrete terms how much they should contribute in comparison to other taxpayers, the majority consider a system of flat tax to be fairer. Likewise, Eriksen and Fallan (1996) provide evidence of the impact

[14] Interestingly enough, Allingham and Sadmo themselves recognised that their model was not capturing all the factors motivating tax compliance: "This is a very simple theory, and it may perhaps be criticised for giving too little attention to non-pecuniary factors in the taxpayer's decision on whether or not to evade taxes."

[15] In the words of the authors: "While our results suggest that knowledge about the tax code cannot be easily manipulated with simple information treatments, the spread of knowledge through peer networks or other sources that affect knowledge in more persistent ways could have larger impacts on behaviour" (Chetty and Saez, 2013).

of tax information on the perceptions of the tax system. In a quasi-experimental work, students who are taught about the tax law tend to change their attitudes at the end of the year: the tax system is perceived to be fairer, while own and other's tax evasion is more likely to be considered wrong.

Despite the paucity of rigorous evidence, the lack of tax knowledge is assumed to explain a substantial part of total tax non-compliance (Richardson, 2006; Palil, 2010). Significantly, as much as 30% of all misreporting in income taxes may be due to honest mistakes (Christian et al., 1993; Erard, 1997). In practical terms, ignorance can affect tax compliance in two opposite ways. On the one hand, it can be associated with lower compliance, including both underreporting and failure to file (Palil, 2010; Lubua, 2014; Kira, 2017). On the other hand, a limited understanding of the tax system could result in higher compliance costs or even overpayment. A recent study shows that taxpayers in the US often pay more than they should due to high compliance costs and relatively complex reporting requirements (Benzarti, 2015).[16] Compliance costs due to regulatory burdens on firms are relevant in SSA too, as some growing literature shows (Bird, 2015; Dabla-Norris et al., 2017).[17] Moreover, this type of costs seem to have a regressive nature. Small taxpayers are more affected as they have less resources to dedicate to tax matters (e.g. through tax advisors), as suggested in Coolidge (2012); Yesegat et al. (2015); Mascagni and Mengistu (2016). Consistently, a recent study focussing on small firms in Finland shows that compliance costs produce larger behavioural responses than changes in the tax rate (Harju et al., 2019).

2.2.4 Contributions

The main contribution of this study is to produce causal evidence on the linkages between tax knowledge and compliance in the specific context of a low-income country in SSA, for

[16]This is in line with related findings showing how poor people often fail to benefit from public programs due to their lack of knowledge about them (Duflo et al., 2006; Feldman et al., 2016).

[17]Using data from the World Bank Enterprise Surveys on 17 SSA countries for the period 2014-2016, Ali (2018) shows that managers in a typical firm in the formal sector spend about 1.3 months dealing with government regulations every year. A strong negative correlation exists between compliance costs and and index of a government regulatory quality.

which knowledge on the topic is almost non-existent.[18] As a matter of fact, very little is known about why people in Africa pay taxes (Fjeldstad and Semboja, 2001; D'Arcy, 2011). Only descriptive, albeit suggestive, evidence on African taxpayers' attitudes exists. Using Afrobarometer data on 36 African countries, Isbell (2017) reports that the majority of respondents have difficulty figuring out what taxes they owe to the government. While small taxpayers are likely to suffer more from lack of tax knowledge, large taxpayers and business associations are also not immune to this issue (Nalishebo and Halwampa, 2014). In a previous study on Afrobarometer data, the same confusion is observed (Aiko and Logan, 2014).

At the same time, there is an increasing awareness, especially amongst African tax specialists, that lack of tax education and knowledge is one of the key obstacles to voluntary tax compliance (OECD, 2015; Mascagni and Santoro, 2018).[19] In SSA alone, Mascagni and Santoro (2018) document the existence of radio programs on tax topics; tax-themed soap operas to sensitise the public about taxpaying; tax clubs in schools where pupils learn about tax and then compete across schools; informative videos on social media, where celebrities explain to young people why is it important to get their tax affaires in order right from the start; and mobile tax units, which are essentially vans traveling to rural areas to support taxpayers.

In this setting, the goal of this study is to inform revenue authorities on how effective alternative, service-oriented, strategies could be in improving tax compliance. While much evidence has been created on the effectiveness of enforcement strategies based on deterrence,[20] much more empirical evidence is needed to support the relevance of "soft"

[18]Few exceptions are qualitative studies (Tanui, 2016; Kira, 2017). Likewise, Ali et al. (2015) find indications that tax knowledge and awareness have a significant impact on tax compliance attitude in South Africa and Tanzania, using, as a proxy for tax knowledge, perceptions on the difficulty to find out what taxes the respondents are required to pay.

[19]This evidence has been collected with in-depth interviews with officials from the taxpayer services departments of revenue authorities in Rwanda, Uganda, Nigeria, Kenya and Tanzania. I have been able to contact them thanks to the network of the ICTD.

[20]Increasing the probability of audit reduces evasion as shown in a variety of contexts: personal income tax in US (Blumenthal et al., 2001) and Norway (Bott et al., 2014), VAT payments in Chile (Pomeranz, 2015), individual municipal taxes in Argentina (Castro and Scartascini, 2013), firm taxes in Ecuador (Carrillo et al., 2017), individual public-TV fees in Austria (Kleven et al., 2011; Fellner et al., 2013), individual church tax in Germany (Dwenger et al., 2015), corporate income tax in Uruguay (Bergolo et al., 2019).

alternatives.

As a second main contribution, this research is related to the growing strand of experimental literature on informational nudges used by tax administrations to increase compliance (more on this in Chapter 4).[21] These field experiments evaluate the effectiveness of messages sent by the local tax administration to taxpayers around the time of declaration. Such messages provide information about specific aspects of the tax system, typically related to either deterrence (e.g. sanctions for non-compliance) or tax morale (e.g. public services funded by tax). While these nudges are generally effective in increasing compliance in the tax declaration that immediately follows receipt of the message, most studies do not look at longer term effects. The limited existing evidence is that they do not seem to generate any longer-term effect (Manoli and Turner, 2014). The literature is a lot more scarce for low-income countries, although two studies from Ethiopia and Rwanda largely confirm the results of the broader literature – including the lack of any effect beyond the first year (Shimeles et al., 2017; Mascagni et al., 2017). The (very) short-lived effectiveness of behavioural nudges begs the question on which other interventions can be implemented to affect compliance in the longer term, through learning. This study attempts to address this question by showing that taxpayer education can be an effective alternative that, contrary to simple messages, could affect compliance more persistently.

Lastly, these results also speak to the literature linking taxation to state-building and accountability (Brautigam et al., 2008; Prichard, 2015; Moore et al., 2018). While weak tax knowledge can directly translate into poor compliance, it also has a number of other detrimental implications. Confused taxpayers would not know with confidence how much they should pay, thus potentially being more vulnerable to corrupt officials or to be coerced into making unofficial payments. They may also be more prone to seeing the tax system as unjust and extortionary, either because of corruption or because they might misperceive the benefits of paying taxes (Ali et al., 2015).[22] Ultimately, uninformed taxpayers are less likely to engage in a meaningful debate with the government about tax issues, thus limiting the potential of taxation to act as a catalyst for improved governance and accountability.

[21] Recent reviews of this literature are, for example, Mascagni (2018) and Hallsworth (2014).

[22] A recent study has shown that providing beneficiaries with information about the eligibility and amount of a subsidy, increases the amount of subsidy they receive by about 26%, thanks to lower leakage (Banerjee et al., 2018).

2.3 The Rwandan context

Rwanda is a small landlocked country in Eastern Africa, with a densely packed population of about 12.5 million people (2017). After going through a civil war and genocide in the 1990s, the country experienced a period of development that ensured a sustained economic growth, progress on human and social development, and a large expansion of public services. However, Rwanda remains one of the poorest countries in the world, with a per capita GDP of $702 in 2016, rising from just $342 ten years before. Between 2001 and 2016, real GDP growth averaged at about 8% per annum, while in 2017 it amounted to 6.1%.[23]

As shown in Table 2.1 (extracted from Table 1.1), tax revenues represent about 15.3 per cent of GDP in 2018, higher than the average of the East African Community (EAC)[24]. In a constant positive trend, Rwanda's total domestic revenue as a percentage of GDP rose from 8.4 per cent in 1993 to 15.3 per cent in 2015.[25] This figure is in line with other African and low-income countries, despite the absence of significant natural resources, but considerably lower than OECD's 25 per cent . More importantly for the sake of this study, as compared to other African countries, Rwanda collects proportionally more income taxes, about a third of total revenues (ATAF, 2017).[26] In a context of rising inequality such as the African continent, the higher reliance on progressive income taxes is usually seen as a way to reduce such inequalities. From the researcher's perspective, it is also more interesting to look at income tax compliance as evasion's opportunities are much higher with income taxes than other types of taxes (such as PAYE), given that the income taxpayer has to self-assess her own tax liability (Slemrod et al., 2001; Kleven et al., 2011).

Table 2.1 also provides a quick picture of the quality of governance in the country. The Transparency International's Corruption Perception Index for 2017 indicates considerable differences between Rwanda and EAC: Rwanda's score is 66% higher (better) than EAC's. This is reflected in the World Bank Governance indicators for 2016. Specifically, Rwanda

[23]World Bank website: `http://www.worldbank.org/en/country/rwanda/overview#3`

[24]Other economies in East African Community: Kenya 17.7%, Uganda 12.3%, Burundi 12% and Tanzania 11.6%. Tax revenues per capita amount to $106, lower than the EAC average, $112 (ATAF, 2017)

[25]In the year in which the genocide took place (1994), however, total domestic resources as a percentage of GDP fell to 3.6% (ATAF, 2017).

[26]This is more in line with what happened in high-income countries (33%), than low-income ones (27%).

Table 2.1: Governance and country indicators

	Rwanda	East Africa	Year
Tax to GDP ratio [a]	15.3%	13.8%	2018
Tax revenue per capita (USD)[a]	106	112	2015
Informality (% national income)[b]	36.3	40.3	2004-2015
CPI[c]	55	33	2017
Governance indicators[d]			
Control of corruption	0.69	-0.59	2016
Rule of law	0.07	-0.50	2016
Regulatory quality	0.11	-0.33	2016
Government effectiveness	0.11	-0.54	2016
Political stability	-0.05	-0.592	2016
Voice and accountability	-1.21	-0.75	2016
Fragile States index[e]	89.3	91.7	2018
Index of economic freedom[f]	69.1	59.3	2018
Tax burden	75.8	75.6	2018
Government integrity	61.2	35	2018
Judicial effectiveness	79.6	44.1	2018
Business freedom	81.2	51.8	2018
Doing business indicator[g]	73.4	59.3	2018
Starting a business	87.66	81.6	2018
Registering property	93.26	63.1	2018
Paying taxes	84.6	69	2018
Bank account ownership[h]	38%	39.5%	2017

East Africa: Burundi, Kenya, Rwanda, Tanzania and Uganda.
[a] African Tax Administration Forum (2017)
[b] Schneider and Medina (2018)
[c] Transparency International Corruption Perceptions Index. Range: 0-100.
[d] World Bank (2017). Range: -2.5 (weak) to 2.5 (strong).
[e] Fragile States Index. Min 17.9 - Max 113.
[f] The Heritage Foundation. Range: 0-100.
[g] World Bank (2018). Range: 0-100.
[h] Global Findex (2017). Adults (+15 yo) in labor force. Burundi excluded.

has much higher control of corruption than the rest of the region. Similarly, the rule of law index, which measures the degree of confidence in and compliance with the rules of society, is much larger in Rwanda. It results that the perceived legitimacy and authority of the State is better judged. Likewise, the capacity of the government to formulate and implement sound policies effectively (represented as the government effectiveness and regulatory quality) is perceived to be higher. Moreover, political stability is higher than in the

EAC – even though with a negative value. Only the index on voice and accountability is worse than the regional average, signalling that political freedom is constrained. Overall, the values of these six governance dimensions clearly indicate a higher level of institutional quality in Rwanda compared to the rest of the region.

Additionally, Table 2.1 also shows that Rwanda scores better in terms of economic freedom and ease of doing business. According to the index of economic freedom from the Heritage Foundation, while the tax burden is somehow in line with EAC, perceptions of the government integrity and judicial effectiveness almost double the ones in EAC. Overall, the index of economic freedom in Rwanda stands at 69.1, well above the EAC average of 59.3. The World Bank Doing Business indicators reflect the same story: Rwanda has the best rating among the EAC countries, 73.4 versus 59.3. Starting a business is much easier in Rwanda – it has the second best score over the whole sample of 190 economies for what concerns registering property, taking just 7 days against an average of 59 for SSA. More significantly, the best result concerns the ease of paying taxes, for which Rwanda ranks 31st out of 190 countries worldwide in 2017, and the second best rating in the Sub-Saharan Region, following Mauritius.[27] The WB Doing Business indicators suggest that the Rwanda Revenue Authority already put in place measures to reduce compliance costs and simplify the tax system.[28] For example, Rwandan taxpayers can file online through the e-tax system or, for micro taxpayers, through a mobile phone platform (m-tax). To simplify things further for small taxpayers, Rwanda also enjoys two simplified regimes with minimal bookkeeping and reporting requirements (i.e. only turnover), as explained in the next section.

[27] According to the World Bank 2018 Paying Taxes report, a Rwandan company is required to make just 8 payments a year which are below the SSA average of 37.2 p.a., and even fewer than the OCED average of 10.9 p.a.

[28] This is relevant to this study, as the intervention I set to evaluate does not vary any administrative parameter on compliance costs (i.e. simpler procedures, less strict requirements), but rather affects access to information about the tax system and the cognitive costs related to understanding the taxpaying process.

2.3.1 Statutory income taxes

Rwanda's tax system includes many of the features that one would expect to find in a modern tax system. In terms of direct taxation, Rwanda adopts three types of income tax: Personal Income Tax (PIT), Corporate Income Tax (CIT), and Pay As You Earn (PAYE). Firms above a specific threshold are also required to pay value added tax (VAT), while below the threshold they are liable for turnover tax (TOT).[29] In this study, I focus on PIT and CIT, since PAYE differs in terms of how it is collected.[30] According to ICTD/UNU-WIDER (2020), income taxes represent 40% of total tax revenues in 2018. The two types of income taxes – regulated by RRA (2017) – can be described as follows:

- PIT taxes individual businesses. Three regimes exist, depending on the business' turnover: 1) real regime, 2) lump-sum regime, and 3) flat-amount regime. While large businesses in the real regime have to submit full books of accounts to RRA, smaller businesses are subject to less strict bookkeeping rules. The declaration deadline is March 31^{st} of the following tax period. The tax rate in the PIT real regime is progressive, depending on income: turnover below RWF 360,000 is exempted from taxation; between RWF 360,001 and RWF 1,200,000 is taxed at 20%; higher incomes are taxed at 30%. Small enterprises (turnover between RWF 12 million and RWF 50 million per tax period) are subject to a lump-sum tax of 3% of turnover. Micro-enterprises generating a turnover of more than RWF 2 million and less than RWF 12 million are required to pay a flat amount between RWF 60,000 and RWF 300,000, depending on income (RRA 2017: Articles 11 and 12).

- CIT taxes the income generated from corporate business activities, which has to be declared annually by March 31^{st}, as for PIT. Corporate income is levied at a flat rate of 30%. Some reductions exist for businesses less than 5 years old, depending on how many shares the public possesses (RRA 2017: Articles 41 to 43). Moreover, small and micro companies follow the same tax structure as in the PIT scheme.

[29] The threshold is 20 million RWF, or about 21,400 USD. Firms can voluntarily opt in the VAT system even if they are below the threshold.

[30] PAYE tax of employees is usually withdrawn from the employer at the time of paying salaries. Rwanda, like many other countries, also applies a number of other taxes that I will not mention here because they are not relevant to this study, such as property taxes, taxes on capital gains, etc.

A descriptive analysis of RRA tax returns data provides some stylised facts on the relative importance of the two income taxes. More details are reported in Appendix Table A1. In 2017, over a total of 50,346 declarations, PIT represents the 44% while CIT the 56%. However, CIT accounts for the 94.5% of the total tax take. Indeed, PIT businesses are much smaller: they report an annual average of RWF 1.7 million ($1,921), versus RWF 144 million ($161,645) declared by CIT taxpayers. The amount of tax declared is also different in magnitude, with PIT averaging RWF 345,000 ($388) while CIT more than RWF 2.3 million ($2,600). The location of businesses is relevant too: while 54% of all declaring businesses are registered in the capital Kigali, they account for 94% of the tax take, while taxpayers outside the province of Kigali contribute for 6% only, as shown in Appendix Figure A3. This is consistent with the fact that most CIT firms are registered in Kigali. At the same time, Appendix Figures A4 displays the share of income tax revenue contributed by small and large firms, where large firms have business income greater than the 90^{th} percentile. As often happens in less developed economies, most of the revenues are raised from the top-decile of firms who account for 94% and 82% of CIT and PIT take, respectively, in the Rwandan economy.

2.3.2 Non- and nil-filers

Despite the ease in paying taxes discussed above, failure to file and zero filing of income taxes remain sizeable issues in Rwanda, especially for new registered taxpayers. Failure to file happens when a registered taxpayer misses his declaration by March 31^{st} of the next year, so becoming a "ghost" to the eyes of the authority (Erard and Ho, 2001). Instead, nil-filing implies filing a declaration, but declaring zero turnover and zero income tax.

While failure to file unequivocally represents a wrongdoing and a severe cause of inequality (competitiveness' gaps, economic distortions and lower tax morale among similar, filing, taxpayers), nil-filing is harder to explain. At the moment, three alternative hypotheses for nil-filing can be considered: i) tax evasion from a business hiding its income ii) legitimate declaration from a business who registered but is not operating yet, thus having zero income to declare, and iii) businesses who ceased operations but are still registered due to the bureaucratic complexity of deregistering from the authority records. A parallel work in Rwanda (Mascagni et al., 2020), attempts to shed light on nil-filing, combining a descriptive analysis of tax returns data, a randomised controlled trial (RCT), and qualitative interviews with taxpayers and tax officials. Mascagni et al. (2020) argue that evasion seems to play only a relatively small role in explaining nil-filing. Instead, a major reason for nil-filing seems to refer to aggressive recruitment campaigns by the RRA and taxpayers' response to a complex and often confusing tax system.

Administrative data help quantifying these two phenomena. If I consider new registrations for the period 2015-2017, it results that 48% of income taxpayers failed to file in their first year and, among those that filed, 54% nil-filed in their first year. It is also noticed that failure to file is a bigger problem for smaller, individual businesses registered for PIT, which probably do not have proper accounting skills, while nilfiling is more prominent in CIT businesses, who may manage to declare but nevertheless report nil as a tax avoiding strategy (Table A1).[31]

2.3.3 RRA tax education strategy and registration campaigns

In the context delineated above, the RRA operates since its establishment as a semi-autonomous agency in 1997.[32] In recent years, RRA successfully embraced the key principles of a modern tax administration and adopted a strategic orientation towards providing tax education and assistance. Many different civic education initiatives are held each year, ranging from the National Taxpayer's Day to tax taught in schools, seminars and workshops. Among these initiatives, the service-oriented approach is well represented by the Tax Education Program, the focus of this study, funded and implemented by the Taxpayer Service Department (TPS) within RRA.

The program has a national coverage and consists of about 30 trainings events in every district per year. In the year under study, trainings started in August 2017 and ended in March 2018 (see Table A2). Trainings in big urban districts are held twice a year given the large number of attendees. The program targets specifically new registered taxpayers, as they are believed to need more information to navigate through the system and abide to the law. They have typically registered for income taxes (PIT or CIT) and obtained their

[31] The larger extent of nil returns among companies is also documented in Eswatini (see Chapter 3 and 4).

[32] After a wave of structural reforms in the 2000s, today the RRA has offices in all the provincial headquarters, and in 11 out of the 30 districts, employing about 1300 staffs

TIN for the first time.[33] These are all new businesses, either individual non-incorporated ones (PIT) or corporations (CIT), and not taxpayers who only have employment incomes.

The program is delivered by TPS tax officials. The content covers the basics of taxpaying, including for example explanations of the various taxes and duties, such as deadlines, what is a TIN and why it is needed, procedures for tax declaration and payment, services available for simplifying declarations and payments, such as online services, amongst others. As such, the content is more focussed on providing practical information on taxpaying, rather than taking a broader approach based on accountability and citizen engagement. I would therefore expect it to have some impact on compliance behaviour, not least because it is specifically targeted at increasing taxpayer knowledge to this respect. Trainings are mainly conducted in hotels, last for a half a day and lunch is provided. Invitations normally happen through SMS, official letters posted at the office of the relevant tax centre and the local branch of the Private Sector Federation, and some phone calls from RRA officials to taxpayers to remind them of the session.

Although the invitation process is comprehensive in principle, practical difficulties and administrative constraints mean that in reality it does not reach all intended beneficiaries. Nonetheless, the explicit intention of the RRA is to invite all new taxpayers and, potentially, keep the training accessible to any other taxpayer even if they were not specifically invited. In practice, the vast majority of attendees are new taxpayers from the relevant district. However, this policy intention means that I could not randomise the attendance to the trainings.

On top of this initiative, a new alternative program is evaluated in this study. A personalised, coaching program is designed where TPS staff provide assistance to sensitise and educate taxpayers on timely and accurate filing. The content was open in the sense that tax officials would just answer to any query the taxpayer might have, with a maximum of 3 questions or about 15 minutes of duration. Initially thought as a pilot, this 1-1 assistance service may be extended in coming years if proved to be effective. More details are provided in section 2.4.2.

As a last consideration, it is worth stressing that the RRA has been devoting increas-

[33] A taxpayer may subsequently register also for VAT, with the same TIN number for the same business, and thus would not be included in the definition of new taxpayer used here.

ing energies to expanding the tax base and registering new taxpayers. In a State-led push towards facilitating entrepreneurship, mass registrations and door-to-door recruitment campaigns happen frequently and are often implemented aggressively. In 2008, the Rwanda Development Board (RDB) introduced simplified procedures for business registration which reduced the time for registering a business from 16 days to 6 hours.[34] In a context in which tax officials aim at increasing the size of the taxpayer registry, many of the new taxpayers are often recruited without having actually started a business and therefore are unlikely to generate any income. The trend in registration is depicted in Appendix Figure A5. Registrations peak in 2014 with more than 23,000 taxpayers entering the system, representing an increase of 66% with respect to 2013. After 2014 the number of registrations rapidly decreases and stabilises around an average of 18,000 in 2015-2017. Importantly for this study, Mascagni et al. (2020) highlight that very little guidance and information are provided to new entrants at the time of registration. New taxpayers are not given any written confirmation, such as a certificate or proof of registration. This is likely to generate confusion. Unsurprisingly, as shown in section 2.5.1, many new taxpayers do not even know what tax type they are registered for.

2.4 Evaluation Design

2.4.1 Data

Data includes information from three main sources: (i) phone-based baseline and endline surveys, (ii) attendance recording of all trainings in the country, and (iii) RRA administrative data on income tax returns. Importantly, I can connect these three sources of information thanks to unique identifiers (TIN).

[34] According to the World Bank Doing Business report, Rwanda ranks fifty-first worldwide and third in sub-Saharan Africa (SSA) for the ease of starting a business. Simplification in registration, with a score of 91.4, is much higher than the average for SSA, 78.5. Online company registration has become mandatory as of 17 February 2014 and is free of charge.

Survey data

Survey data is collected by an independent team of enumerators from a research company based in Rwanda. The first wave of survey was conducted in August 2017, one week before the trainings. The endline survey was implemented one week after the trainings, in September 2017. In this fashion, information before and after the intervention is gathered from the same respondents.

Surveys have been administered by phone, with enumerators contacting taxpayers on the phone numbers extracted from RRA administrative records.[35] Respondents are replaced in case they are not reached on their phone (see section 2.4.2). Survey coding and data inputting took place through SurveyCTO software. Survey protocol also includes oral consent to take part in the survey and information about confidentiality, delivered at the beginning of the interview.[36]

For what concerns the questionnaire content, the final version went through repeated testing and a meticulous translation in Kinyarwanda. The survey was designed with inputs from RRA researchers and TPS team. The baseline surveys took 42 minutes to complete on average, and consisted of six modules: (i) respondent's demographics (5 questions), (ii) business characteristics (14 questions), (iii) reasons to register or remain informal (2 questions), (iv) tax attitudes and perceptions (14 questions),[37] (v) a quiz on tax knowledge (19 questions),[38] and (vi) intent to attend and other tax training experiences (5 questions).

The endline survey was shorter in length, with an average duration of 23 minutes, consisting of the same questions of modules 4 and 5, plus ten new questions for attendees only, where the enumerators asked the attendees for feedback on the training. For what relates

[35] The choice to use cell phones for data collection has been mainly motivated by budget reasons, with in-person surveys being more expensive. Moreover, it can be assumed that the distance created using cell phones also put the respondents in a more comfortable position to answer some sensitive questions related to attitudes towards compliance and perceptions about the tax system.

[36] Permission to conduct the survey has been obtained by the National Institute of Statistics in Kigali, which also granted the Ethics Approval in July 2017.

[37] Most of the questions on attitudes and perceptions are derived from the most recent wave of Afrobarometer survey (2017). However, Rwanda is not covered by it. Unfortunately, Rwanda lacks nation-wide perception surveys, which would have offered the possibility to compare and validate the data collected in this study.

[38] Knowledge questions have been formulated so to reflect the topics taught in the training.

to non-attendees, an additional module enquired about the reason of non-attendance.[39]

With the available survey data, I create a set of variables, which will provide valuable information on the underlying mechanisms of impact. These can be considered as intermediate outcomes in the theory of change that attempts to shape filing decision through changes in tax knowledge and perceptions. In order to efficiently summarise different variables and reduce concerns on multiple hypotheses testing, the solution is to form aggregate index measures. More specifically:

- Knowledge: out of 19 questions, a first index is built as the average of 19 dummy variables, each taking value 1 if the answer is correct. Appendix Table A4 describes the 19 questions. Using this index, three key measures are created: (i) a dummy for the ex-post increase in the knowledge, (ii) a continuous knowledge gain variable as the fraction of the post-pre difference over the baseline knowledge score, and (iii) a re-scaled index from 0 to 10, thus representing the percentage fraction of questions that the respondent answered correctly. Additional knowledge indicators are used as well: (i) a knowledge difference index was built as the difference of the post-training minus the pre-training scores (the numerator of measure (ii) above), and (ii) a standardised Kling index (Kling et al., 2007).[40]

- Perceptions: 13 statements on attitudes and perceptions are proposed to the respondents, with answers ranging from 1 *totally disagree* to 5 *totally agree*, as summarised in Appendix Table A5. An additional question on satisfaction with a range of public services allowed answers going from a minimum of 1 to a maximum of 5.[41] Binary perceptions variables are created with value of 1 if respondents agreed with the statement and 0 otherwise. In the same fashion, a binary satisfaction variable is derived

[39] Questionnaire forms are available upon demand.

[40] The Kling normalisation helps translate the magnitudes of different measures into standardised units. The index is defined to be the equally weighted average of z-scores of its components (the 19 knowledge variables), with the sign of each measure oriented so that more positive outcomes have higher scores. The z-scores are the normalised transformations of each variable and are calculated by subtracting the control group mean and dividing by the control group standard deviation. Thus, each component of the index has mean 0 and standard deviation 1 for the control group.

[41] Six public services are included: health, education, water and sanitation, electricity, security/police and infrastructures.

that takes value 1 if the respondent is somehow and very satisfied with a public service (score of 3-4-5) and 0 otherwise. Seven indexes of perceptions are created as group averages of binary variables (see Table A6): complexity costs, enforcement, attitude to evade, fairness of the tax system, government authority, social duty. Finally, an index from principal component analysis summarises the satisfaction with public services.

Attendance data

A dedicated team of enumerators, with the support of RRA staff, registered attendance at the training venue.[42] A pre-populated Excel list of all new registered taxpayers invited to the trainings was provided. Enumerators registering attendance were equipped with laptops to speed up the procedure. Pieces of information such as TINs, business name, tax centre, name of the legal representative, phone numbers, were already in the list. Therefore, attending taxpayers simply had to identify themselves and tick for attendance.[43]

The new registration activity produced more reliable attendance lists.[44] Overall, attendance data from all trainings (see Table A2) held in the country is available:

i 3 trainings in the *survey sample*, Kicukiro, Musanze and Rubavu

ii 4 trainings in the rural areas, implemented in November–December 2017, in Huye, Muhanga, Nyagatare and Rusizi

[42] Previously, attendance was tracked with a paper form passed in the room where attendees were supposed to write by hand few personal details, such as the attendee and business' name, but not the Tax Identification Number (TIN). A first inspection of the previous hard-copy attendance forms revealed that this registration procedure was likely to be flawed and produced unclear data on attendance. More importantly, the absence of TIN information meant that RRA could not easily identify the business and the taxpayers as registered in its records, with no possibility to link attendance to tax returns data.

[43] Separate variables were indicating whether the taxpayer himself was the one surveyed at the baseline or a delegate for him. Given that access to the trainings is open to anyone interested, a separate Excel sheet was used to collect information on all those attendees who had been not surveyed, for example taxpayers from other tax centres, or those registered after July 31 or in previous years.

[44] As a result, the measurement of attendance in each district has become much more precise and the new procedure has been extended to all the RRA trainings for the rest of year.

iii 3 trainings in 3 tax centres in Kigali (Kicukiro, Gasabo and Nyarugenge), held (twice) in January–February 2018.[45]

Data from ii) and iii) refer to the population of new taxpayers invited to the trainings (see Table 2.2 below) and are used to question the external validity of the impacts measured in the survey sample (see section 2.7.1).

Administrative data

Access to administrative data has been officially granted by RRA. It consists of different types of anonymised tax data.[46] First, the CIT and PIT taxpayer registry, accessed at the beginning of August 2017, contains the information needed to select the study sample, such as registration date, tax centre and type of business. Each taxpayer is assigned a TIN thanks to which I am able to observe his filing behavior and link survey data both with attendance data and declaration data. Second, data from CIT and PIT declarations for 2017 fiscal year constitute the key source to analyse the impact on filing behaviour, namely the amount of tax declared, if any.[47] However, this also implies that I have no baseline administrative data for these new taxpayers, as they never filed a declaration before.[48] Declaration data has been available only after the end of the fiscal period. With the filing deadline being March 31, I received the data by mid-April 2018. Additional sets of data have been used to clean the main registry and run side checks: moto-taxi, VAT registered taxpayers, prepayments datasets (see section 2.4.2).

A key feature of this study is the ability to merge survey data with administrative data. Any impact found in the administrative data can be corroborated by the analysis of survey responses. In fact, tax returns capture filing behavior in a more accurate way than what can be measured through questionnaires. Survey respondents often do not provide honest

[45] In Kicukiro, taxpayers registered from August 2017 are invited, so to exclude those already invited to the training in August 21, included in the *survey sample*. All taxpayers registered in Gasabo and Nyarugenge from January 2017 are invited in those tax centers.

[46] While taxpayer's name and contact details are removed, I can access all information included in the registration/declaration forms.

[47] Full information on turnover, costs of goods and services sold, gross profits, deductions, allowances and expenses is available too.

[48] At least, they have not with the TIN number they recently obtained. I cannot observe whether they previously had another TIN number, which however could refer to a completely separate business.

answers about income.[49] Administrative data can solve these problems because it include information on declared income and actual tax payments. At the same time, one of the drawbacks of administrative data is the limited range of variables available in tax returns, typically only sector and location (or a few others) in addition to financial variables related to income, expenses and the tax base. Additionally, it must be acknowledged that the the amount of tax declared only captures the information that taxpayers decided to disclose to the authority, while income from the informal economy or just misreported (blatant evasion, legal but informal activities, etc.) is unobserved.[50]

With these data sources available, three specific sub-samples are created. For surveyed taxpayers, I have all information from all three data sources described above (administrative, attendance, and survey data). Beyond the survey sample, I can still rely on attendance information and administrative data for the whole population of new taxpayers who registered in 2017. Table 2.2 summarises the key data sources by groups of taxpayers that I will use in the analysis.

Table 2.2: Available data by evaluation groups

Group	Attendance	Admin data	Survey
Population of new TPs: all TPs registered in any district anytime in 2017	✓	✓	×
Survey reference population: all new TPs registered in Kicukiro, Musanze and Rubavu in Jan-July 2017 (encouragement design)	✓	✓	×
Survey sample: random selection from survey reference population (naive estimation and PSM)	✓	✓	✓

[49]Obtaining honest answers about taxpayers' compliance is typically one of the major challenges in tax compliance research, especially in studies based on survey data.

[50]It still possible to infer compliance, in terms of income under-reporting, by comparing taxpayers who received a certain intervention to those who did not, for example a letter in the typical nudging RCT (Mascagni et al., 2017). In this case, I compare attendees and non-attendees.

2.4.2 Sample selection

Survey sample

The trainings under study are the first three trainings from the 2017/2018 RRA plan. They include one district from Kigali (Kicukiro) and two from the Northern Province (Musanze and Rubavu districts).[51] The training in Kigali was held on August 21 2017, while trainings in Musanze and Rubavu took place at the end of the month, August 29 and 30 2017, respectively.

The target population is derived from the list of taxpayers who registered for income tax, both personal and corporate, in 2017, as provided by RRA (section 2.4.1). Administrative data contain details of a total of 16,260 new taxpayers, registered across the country from January 1 2017 to June 31 2017. After restricting eligibility to the three districts of interest, respondents without phone number, the 4% of the population, are dropped. Likewise, a small number (2%) of moto-taxi taxpayers, who follow a different tax regime, are removed. The final target population for the three districts under study amounts to 2,551 taxpayers. Of these, the target sample size was set at one thousand, the maximum number of individuals that it was possible to survey with the available budget.[52] The population is thus composed: 64% taxpayers are from Kicukiro, 20.4% from Rubavu and 15.6% from Musanze.

Enumerators surveying taxpayers also invited them to the training. As already explained in section 2.3.3, RRA usually invites the whole pool of new registrations with a SMS, while taxpayers in this sample received an additional, more formal, invitation by the enumerators.

Crucial to the identification strategy of this study, it is worth stressing that the randomisation happens at the invitation-to-the-training level. Indeed, the survey company

[51] The first location, Kicukiro, is a busy urban district in Kigali with a population of 320,000 people (2012), where a big number of taxpayers register every year: in the first six months of 2017, 13% of all new registered taxpayers in Rwanda are from Kicukiro, making it the third largest tax center in the country. Musanze and Rubavu are much smaller centres outside of Kigali (3% of the total pool of new registrations each), with both urban and rural population totalling about 86,000 people in each city (2015).

[52] The survey company was provided with the full list of 2,551 taxpayers with the instruction to stop surveying when reaching a sample size of 1,000.

randomised the call order. I use this random allocation to the survey sample as the basis of the encouragement design, which is discussed in more detail in section 2.4.3. This type of randomisation represents the best feasible option in a setting where attendance to the training cannot be denied.

A total of 1,007 respondents are interviewed at the baseline, while all the remaining contacts in the list are left out. Enumerators had to attempt to call 1,621 taxpayers to achieve the target, randomly replacing unreachable taxpayers. The main reasons individuals are not interviewed are that phone numbers were not valid or not available at the time of the call, while refusals to join the survey are pretty low.[53] In other words, the success rate of the baseline survey is 62% (1,007/1,621) with a failure rate of 38%.

Appendix Table A7 provides t-test statistics on the differences between business in the sample and those left out.[54] An additional column reports the Δ, the so-called *normalised difference* from Imbens and Rubin (2015). Δ is a scale-free measure of the difference in locations, equal to the difference in means, scaled by the square root of the average of the two within-group variances (Imbens and Rubin, 2015).[55] While businesses' life in month is well balanced across surveyed and non-surveyed unit, it was more difficult to contact respondents from Kigali and who remit CIT. Δ statistics for location dummies is beyond the 0.25 threshold of significance. For this reason, over-representation of rural areas is addressed using sample weights throughout the analysis, which restore the balance at the location level and the representativeness of the sample.

One week after the training, enumerators attempted to reach the same respondents for a follow-up. A subset of 186 taxpayers is not found at the endline, implying an attrition rate of 18% (186/1,007).[56] Overall, attrition can be considered relatively low and implies a final sample of 821 taxpayers for which data is available both pre- and post-intervention.

[53] Two percent only.

[54] Unfortunately, very little information is available from the taxpayer registry (only 3 covariates).

[55] As Imbens and Rubin (2015) report, the t-statistic may be large in absolute value simply because the sample is large and, as a result, small differences between the two sample means are statistically significant even if they are substantively small. Large values for the normalised differences, in contrast, indicate that the average covariate values in the two groups are substantially different.

[56] Common types of attritors include (i) 143 respondents who cannot be reached on the phone, (ii) 33 respondents who do not have time for the survey, postpone it and do not pick up when called back, and (iii) 10 respondents who say they do not wish to participate without giving a reason.

Appendix Tables A8 and A9 report balance tests comparing non-attrited and attrited taxpayers. Women and less educated respondents are slightly less likely to be found, together with people who are not owning the business. Likewise, when considering outcome variables, attrition is more likely to happen for people with a lower complexity score and less satisfaction with health and security. From a statistical perspective, attrition does not pose an insurmountable problem: only one variable (*owner*) among the many tested is above the 0.25 threshold of the normalised difference.

Coaching sample

As mentioned in section 2.3.3, this study compares the main RRA education program with an alternative input: the one-to-one coaching program. The motivation for piloting this activity is twofold. First, when collecting feedback on the program at endline, many taxpayers signalled that they would benefit from some form of complement to the training program. About 40% of attendees mentioned that the session was too short, and that there was not sufficient time for questions. Second, a growing number of studies have shown that a more intensive and personalised approach might be beneficial. In the financial education literature, personalised counselling is shown to be a useful alternative to more standard financial education, to increase real financial outcomes (Carpena et al., 2015). The very scarce literature on tax coaching does not seem to be fully conclusive on this respect, and it only concerns high-income countries (Chetty and Saez, 2013; Gangl et al., 2014).

The sample for the 1-1 coaching treatment is derived from the subset of respondents who did not attend the training – in order to compare them with the trainees. As will be described in section 2.5.2, 572 taxpayers did not attend. Half of them (293) are randomly selected to be in the coaching group. The strata used are four: gender, location (in/out Kigali), business' size and baseline tax knowledge. Appendix Table A10 shows that the coaching sample and the control group (Training group excluded) share the same characteristics across most covariates as well as knowledge and perceptions outcomes.[57]

Two rounds of coaching calls are made, one in mid-February and the other in mid-

[57] In the balance tables, the non-coaching sample consists of non-attendees not selected for the coaching. Strata are included too (Bruhn and McKenzie, 2009). Just 2 out of 23 covariates are unbalanced (no school or primary and perception of taxation as social duty) and I assume it is just a random result.

March. The first wave reached 216 taxpayers (out of 293 assigned to this treatment), 57% of whom had at least one tax-related question. Similarly, the second wave contacted 182 taxpayers, with a third having a question. In order to be conservative, I consider as treated those taxpayers receiving at least one call in any of the two rounds, even if they do not have a tax-related question.[58] In conclusion, the coaching treated sample consists of 160 taxpayers.[59] Those taxpayers who were not reached at all are included again in the control group of respondents who did not receive any of the two alternative education treatments.

2.4.3 Identification strategy

Naïve estimation

The first estimation strategy can be labelled naïve. It regresses outcomes on treatment indicator variables using an OLS estimation so to get the difference in outcomes between attendees and non-attendees. The naïve estimates represent the benchmark to which to compare more robust models. When looking at tax compliance outcomes, I estimate pseudo-ATEs with the following specifications:

$$Y_i = \alpha_i + \beta_1 \mathbb{1}\{Training\}_i + X_i\Gamma_i + Z_i\Phi_i + \epsilon_i \tag{2.1}$$

$$Y_i = \alpha_i + \beta_1 \mathbb{1}\{Coaching\}_i + X_i\Gamma_i + Z_i\Phi_i + \epsilon_i \tag{2.2}$$

It is worth stressing that the naïve estimation of the training impact will be improved with more robust methods, since attendance to the training is not random. On the other hand, the estimation of coaching effects is more straightforward since it builds on the stratified random assignment to the treatment. Hence, coefficients from equation 3 can be considered as causal estimates, with no further improvement required.

[58] I assume that the coaching treatment could be effective even just signalling that the authority is client-oriented enough to offer a free service to taxpayers. Plus, the perceived audit risk may be affected even just by the fact that RRA is calling and knows that the taxpayer exists.

[59] The low compliance with the treatment and lacking records are mostly due to capacity constraints on RRA's side, as the staff allocated to this task were few and spread thin across a number of other duties.

In the specifications above, I add both the vector of individual-level covariates X_i (more on this below) and the vector Z_i of tax knowledge and attitudes at the baseline, which are likely to be key determinants of subsequent filing behavior. Covariates used as strata in randomising the assignment to coaching are included in the X_i vector in equation 3 as well. The error terms are robust and clusterised at the individual level.

In terms of reporting outcomes, tax compliance is a multi-faceted concept that encompasses various different aspects, including registration, filing a declaration, doing so in time, reporting the correct amount of tax base, and paying in full and on time (Slemrod, 2019). It is therefore important to consider multiple outcomes in the analysis. Here I focus on three aspects that are particularly relevant to new taxpayers: the probability to declare, the probability to nil-file, and the amount of tax declared.

When it comes to investigating possible mechanisms, two sets of outcomes from the survey sample are considered: i) tax knowledge and ii) tax perceptions. The estimating equation reads:

$$Y_i = \alpha_i + \beta_i \mathbb{1}\{Training\}_i + \gamma_i Y_{0i} + X_i \Gamma + \epsilon_i \tag{2.3}$$

Where $\mathbb{1}\{Training\}_i$ is an indicator variable that equals 1 if individual i attended the training and β_i is the coefficient of interest. In order to improve statistical precision, I include the baseline value Y_{0i} of the outcome (McKenzie, 2012). In equations (1) to (3), the vector of individual-level covariates, X_i, includes age (in years), a female gender indicator variable, a dummy for businesses of large size (more than 20 employees), time in months since registration and a number of dummies for primary education only, university degree, being CIT registered, business location in Kigali, use of email, bank account, books of accounts, whether the taxpayer manages taxes by himself (and the number of days spent in tax matters in a year), as well as whether he had a previous business or a previous training.[60] As above, the error term ϵ_i are robust to heteroskedasticity and clusterised at the individual level.

Different tax knowledge outcomes are considered: i) a dummy for knowledge increase after the training, ii) a continuous variable for knowledge gain in percentage points and iii) the 0-10 index. For attitudes and perceptions, I will provide results on: i) seven indexes,

[60] Appendix Table A3 lists the covariates used.

and ii) 13 separate attitudes variables. The creation of such outcomes has been described in detail in section 2.4.1.

Dealing with self-selection Estimating the equations above poses a number of technical challenges. First, training attendance is not randomised, due to RRA's explicit policy intention to keep it open to all new taxpayers who might find it useful. Therefore, I cannot assume the two groups I compare to be randomly equal. Second, always due to the lack of randomisation, taxpayers may self-select into the training. This implies that there could be unobservables in the error term explaining the decision to attend and also influencing the outcomes, thus biasing the results. My approach to tackle these issues is to adopt different econometric methods that, at least in part, take care of self-selection.

Propensity score matching

Propensity score matching (PSM) estimators (Rosenbaum and Rubin, 1983) have become increasingly popular in medical trials and in the evaluation of economic policy interventions mostly because matching on the propensity scores helps approximate randomised trial.

Rosenbaum and Rubin (1983) define the propensity score as the conditional probability of assignment to a treatment given a vector of covariates. In this setting, the probability p_i of attending the training is allowed to depend on covariates X_i and the vector of baseline outcomes Y_{0i} (see section 2.4.3).[61] These individual-level assignment probabilities p_i are called *propensity scores* and derived with a Logit function.

A first identifying assumption is *unconfoundedness* or *strong ignorability* (Rosenbaum and Rubin, 1983) and requires that treatment exposure is independent of potential outcomes. Ideally, if all confounders are included in X_i, the independence holds after controlling for X_i. Under this assumption, adjusting for the propensity score only is enough to remove all confounding. Unconfoundedness is untestable. However, as I show below, attendees and non-attendees are similar based on X_i (and baseline outcomes), thus supporting the adoption of PSM.

A second relevant assumption is the *overlap* or *common support* condition. It states that the probability of assignment is bounded away from zero and one, meaning that

[61]Therefore, I suppose that adjusting for a set of covariates is sufficient to eliminate selection bias.

individuals with the same score can be either attendees or non-attendees, as if a completely randomised experiment were carried out. In this case, the overlap assumption nicely holds: Figure A6 shows how the probability distributions of attendees and non-attendees are mostly overlapping.[62]

Next, I estimate average treatment effects using different matching algorithms. These are: i) nearest-neighbor matching with replacement,[63] ii) radius-caliper matching with a radius of 0.5,[64] and iii) Epanechnikov kernel matching with bandwidth 0.06.[65] As usually done in the literature, standard errors are estimated with the bootstrapping technique.[66]

Moreover, the quality of the matching is evaluated. Ideally, the matching procedure should be able to balance the distribution of the relevant variables in both the control and treatment group. Many different techniques to test the matching quality exist. The basic idea of all approaches is to compare the covariates' distributions before and after matching and check if any differences remain after conditioning on the propensity score. In the results tables, I report two estimates of balance: (i) the pseudo-R^2s before and after matching from Sianesi (2004),[67] and (ii) a likelihood ratio test on the joint significance of

[62] Only two taxpayers are out of the common support and are removed from the analysis

[63] The taxpayer from the non-attendees group is chosen as a matching partner for a trained taxpayer that is closest in terms of the propensity score. With replacement, an untrained taxpayer is used more than once as a match.

[64] Applying caliper matching means that a taxpayer from the untrained group is chosen as a matching partner for a treated taxpayer that lies within a given distance of p_i (caliper) and thus is closest in terms of propensity score.

[65] Kernel matching (and local linear matching) are nonparametric matching estimators that use weighted averages of (nearly) all individuals in the control group to construct the counterfactual outcome. How many individuals are chosen from the control group depends on the kernel function. Weights depend on the distance between each individual from the control group and the participant observation for which the counterfactual is estimated.

[66] An early example of bootstrapping with PSM is found in Heckman et al. (1997) who report bootstrap standard errors for LLM estimators. Other application examples are Black and Smith (2004) for nearest-neighbour and kernel matching estimators or Sianesi (2004) in the context of caliper matching. Each bootstrap draw implies the re-estimation of the results, including the first steps of the estimation (propensity score, common support, etc.). Repeating the bootstrapping N times leads to N bootstrap samples and N estimated average treatment effects. The distribution of these means approximates the sampling distribution (and thus the standard error) of the population mean.

[67] Sianesi (2004) recommends to reestimate the score on the matched sample, i.e. only on attendees and matched non-attendees, and compare the pseudo-R^2s before and after matching. In this case, the pseudo-

all regressors in the Logit model. According to (i), the pseudo-R^2s after matching should be close to zero,[68] while, in terms of (ii), the test should not be rejected before (since balance is not achieved), and should be rejected after (when the score balances the two groups), matching. It can be seen that it is the case in all PSM results below.

Finally, PSM can be implemented on the survey sample alone (see Table 2.2) for which the features to build the score are available.

IV strategy

The intuition behind the adoption of the IV strategy is that, while I cannot randomise training attendance, I can at least randomise the invitation to the training and then use the random invitation as an instrument for the training attendance.[69] The encouragement consists of the random survey itself, at the beginning of which enumerators made sure that respondents knew about the date and location of the training.

The encouragement can be considered as *unintended* because this information would have been redundant in presence of a comprehensive invitation process by RRA, which should have in principle reached all newly registered taxpayers. However, administrative constraints in the RRA department in charge of invitations meant that, in practice, many taxpayers did not receive the invitation and, thus, did not know about the program. As a result, exposure to the program was much higher for the surveyed taxpayers (49% attended) than for the remaining ones (15%). It can be argued that the main, or perhaps even the only, reason for this difference is that the first survey round happened a week before the training, and therefore served as a reminder to attend. The questionnaire did not include any incentive for attendance, and respondents were informed that data was being collected by an independent company and used in anonymised format only for research purposes.

In this setting, two groups are created at random: 827 taxpayers who are called and

R^2 indicates the explanatory power of the covariates X_i and Y_{0i} on the probability to attend the training. After matching there should be no additional systematic differences in the distribution of covariates between the two groups and therefore the pseudo-R^2 should be fairly low.

[68]Or, at least, smaller than the pseudo-R^2s on the unmatched sample.

[69]This is in line with the frequent setting of encouragement design where it is impossible to force people to take up at treatment but the encouragement is randomly assigned. For a theoretical discussion see Angrist and Pischke (2009).

surveyed and 1551 who are left out.[70] Using the random invite as an instrument, I can purge the training treatment T of its correlation with the error term ϵ in the main specification (section 2.4.3) and produce less biased estimates.[71]

In order to be valid, the instrument must be: (i) correlated with the endogenous explanatory variable (relevance condition); (ii) but uncorrelated with the error term (exclusion restriction). While it is possible to test (i), it is not possible to verify directly (ii), since the error term is unobservable. In terms of testing (i), the first-stage result confirms that the instrument is relevant and significantly affecting training attendance (more in section 2.7.5) . Relatedly, while actual attendance is 49% for the survey sample, it falls to 15% for the out-of-the-survey sample. For what concerns (ii), section 2.7.5 shows different tests on the plausibility of the exclusion restriction.

Although the encouragement design allows to rigorously evaluate the effectiveness of the program, the fact that it relies on both surveyed and non-surveyed taxpayers means that I cannot use the full dataset. More precisely, I can only use administrative and attendance data, since the survey is only available for part of the survey reference population that I use for the encouragement design (see Table 2.2). At the same time, PSM is limited by data availability and it cannot control for unobservables which, in turn, are addressed by the random invitation used in the IV strategy. For this reason, I consider the IV as the most robust and conservative strategy to measure impacts in this study.

2.5 Results

2.5.1 Anatomy of invitees at baseline

Since I have no baseline administrative data, in this section I rely mostly on the survey to provide an anatomy of new taxpayers before the program. In the first three columns

[70] The group of surveyed taxpayers exclude those assigned to the coaching treatment, which is not covered in this section. I will focus on instrumenting attendance to the training only.

[71] More specifically, instrumenting gets rid of the *bad* variation in T, which compromises identification because it entails that $Cov(T, \epsilon) \neq 0$ - while keeping only the *good* variation in T - that part of the variation in T that is uncorrelated with ϵ.

of Table 2.3, summary statistics of the sample are displayed.[72] First, as expected, new taxpayers are a relatively homogeneous group of small businesses: over 90% have less than 5 employees and many have no employee (besides the owner).[73] In almost 80% of the cases, the invitee is the owner of the firm. Average business practices are poor: only 25% use emails for business-related communications, 44% have a bank account,[74] while 47% make use of books of account. Second, for what concerns invitees' demographic background, the average respondent is 33 years old and a woman in a third of the cases. About 20% have primary education at most, while a sizeable 44% report to have university education.[75] Third, when looking at tax-related characteristics, it results that only 6% of the sample had a previous tax training. When asked about the main reason they had to formalise, obeying the law is mentioned by almost 70% of the sample while only rarely the fear of sanctions from being illegal is the key reason to register.[76] Strikingly enough, 94% of the sample acknowledges that the main reason to attend the RRA training is that they "need to know more about taxes".

In the same fashion, Table 2.4 provides baseline values of the variables of interest, knowledge and perceptions. On the one hand, knowledge at baseline is very low. Out of 19 questions in our knowledge module, respondents in the sample answer correctly to only 6 of them, or 3.4 correct answers once I re-scale the knowledge index to take values from

[72] Note that the sample here is defined as the 971 taxpayers for which survey data could be matched with TINs in the attendance sheets. Due to some issues in the logistics of attendance registration, the survey company was unable to track all the 1007 taxpayers invited at the baseline in the data collection process. For 36 taxpayers TIN could not be found.

[73] More specifically, about 48% of the sample have no other employee, while another 39% have less than 5 employees.

[74] According to the 2017 Global Findex, 38% of Rwandan adults (aged 15+) in labor force have an account with a regulated financial service provider, declining from 42% in 2011 and 45% in 2014. In comparison, the average rate of account ownership in the East Africa Community (Burundi excluded) is 39% as of 2017.

[75] This result is not surprising. Anecdotal evidence from RRA tells a story of a high number of young business enthusiasts responding to Government calls for them to start their own business after completing university. The Government encourages business creation for the overall goal of sustaining the economic growth of the country.

[76] At the same time, taxpayers are aware that many constraints prevent business to be legal: too small size (27%), too high taxes (25%) and lack of knowledge about the benefits of regulating their position with the State.

Table 2.3: Baseline Characteristics and Mean Differences by Attendance

	Baseline results			Non-attendees		Attendees			
	Mean	St.dev	Obs	Mean	Obs	Mean	Obs	Diff.	Δ
Covariates									
Age	33.08	9.37	971	32.48	604	34.08	367	-1.61***	0.17
Female	0.32	0.47	971	0.31	604	0.34	367	-0.02	0.05
No school or primary	0.19	0.40	971	0.21	604	0.17	367	0.03	-0.08
University degree	0.45	0.50	971	0.46	604	0.44	367	0.02	-0.04
Owner	0.78	0.42	971	0.76	604	0.80	367	-0.03	0.08
CIT	0.47	0.50	971	0.44	604	0.51	367	-0.07	0.13
>5 employees	0.09	0.29	971	0.08	604	0.11	367	-0.04**	0.13
Life in months	4.62	2.01	952	4.56	598	4.64	382	-0.08	0.04
Kigali	0.49	0.50	951	0.52	569	0.43	382	0.09***	-0.17
Email use	0.25	0.44	971	0.25	604	0.26	367	-0.01	0.02
Bank Account	0.44	0.50	971	0.42	604	0.48	367	-0.05	0.11
Books	0.47	0.50	971	0.46	604	0.50	367	-0.04	0.08
Had previous business	0.21	0.41	971	0.20	604	0.22	367	-0.02	0.04
Had previous training	0.07	0.25	971	0.07	604	0.06	367	0.01	-0.06
Tax time use days	1.86	4.99	710	1.90	432	1.78	278	0.12	-0.03
# RRA visits	1.36	2.31	845	1.44	521	1.23	324	0.21	-0.09
Reasons to formalize									
Obeying the Law	0.70	0.46	971	0.70	604	0.70	367	-0.00	0.00
Better reputation	0.26	0.44	971	0.24	604	0.28	367	-0.03	0.08
Constraints to formalize									
Business too small	0.27	0.44	971	0.25	604	0.31	367	-0.07**	0.15
High tax rate	0.25	0.43	971	0.26	604	0.24	367	0.02	-0.04
No knowledge abt process	0.13	0.33	971	0.12	604	0.14	367	-0.02	0.05
No knowledge abt benefits	0.15	0.35	971	0.15	604	0.14	367	0.01	-0.02
Reasons to attend the training									
I need to know more	0.94	0.24	929	0.93	565	0.95	364	-0.02	0.08
Compulsory to attend	0.26	0.44	929	0.21	565	0.32	364	-0.11***	0.24
I'm curious about training	0.16	0.37	929	0.17	565	0.15	364	0.01	-0.03

Observations weighted by sampling weights. T-test are computed on the *Difference* across the two groups. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Δ is the normalized difference from Imbens and Rubin (2015), equal to the difference in means, scaled by the square root of the average of the two within-group variances.

0 to 10.[77] No respondent gave the right answer to all questions. Interestingly, 37% of the respondents do not know what tax type they registered for. This clearly confirms the importance of organising taxpayer education trainings in Rwanda, as well as the presence of large margins for improvement. Low knowledge of the basic parameters of the tax system is also consistent with the fact that most new taxpayers in Rwanda fail to make a declaration in the first year since registration, as explained in section 2.3.2.

On the other hand, perceptions are generally good at baseline. Virtually all taxpayers agreed with the tax attitude statements I provided, for example on the government's authority to make people pay tax (95%), fairness of the tax system (98%), and tax as a social duty (98%). These figures are fully in line with the Rwandan context and the government rhetoric of self-reliance but also more objectively corroborate by international assessment (Table 2.1). The perceived enforcement likelihood is also very high (98%). At the same time, complexity of the tax system matters: 70% of the sample agree that filing tax is complex. Taxpayers with lower tax knowledge find tax filing more difficult.[78] Satisfaction with public goods is also very high, with security ranking first. Appendix table A5 reports the distribution of answers for the 13 questions in the perceptions module. Having very high percentages of agreement with many of the statements included means that I should not expect the training to have much effect in increasing them further.[79]

2.5.2 Attendance and self-selection

When surveyed, as many as 97% of the taxpayers say that they will attend the training, as shown in Table 2.3. However, actual attendance rates were not as high. A total of 367 taxpayers out of 971 surveyed actually attended, or 37.8%.[80] To be conservative, I

[77]Lower education, location outside of the capital, being female, and smaller business size are all associated with lower tax knowledge - although not all of these differences are statistically significant.

[78]Taxpayers who agree with the "difficult to file statement" have an average knowledge score of 3.2, while taxpayers disagreeing have a score of 3.6, with a p-value of the t-test of the difference in averages of 0%.

[79]Relatedly, it must be emphasised that all types of survey data may be subject to social desirability bias and other forms of measurement error, in particular when asking about sensitive topics, such as politics and tax evasion (Mullainathan and Bertrand, 2001; Krosnick, 1999).

[80]As already stated in section 2.5.1, note that 971 are those taxpayers the survey company was able to track in the data collection process. See footnote 72. Also, note that the attendance rate is lower than the what mentioned in section 2.4.3 in the context of the IV strategy, since, in that case, the coaching

Table 2.4: Baseline Outcomes and Mean Differences by Attendance

	Baseline results			Non-attendees		Attendees			
	Mean	St.dev	Obs	Mean	Obs	Mean	Obs	Diff.	Δ
Baseline Knowledge Index	6.35	2.92	971	6.29	604	6.46	367	-0.17	0.06
Complexity	0.70	0.46	939	0.69	579	0.72	360	-0.03	0.06
Evasion attitude	0.33	0.47	924	0.33	572	0.34	352	-0.01	0.02
Fairness	0.98	0.14	932	0.98	578	0.98	354	0.00	-0.00
Govt Authority	0.95	0.21	950	0.95	591	0.96	359	-0.01	0.04
Social duty	0.98	0.14	969	0.98	602	0.99	367	-0.01	0.06
Enforcement	0.98	0.15	969	0.98	604	0.98	365	0.00	-0.03
Public Service Satisfaction									
Health	3.39	1.17	961	3.38	597	3.42	364	-0.04	0.04
Education	3.50	1.16	957	3.47	595	3.54	362	-0.08	0.02
WASH	3.35	1.23	965	3.35	599	3.34	366	0.00	0.08
Electricity	3.81	1.03	967	3.83	600	3.78	367	0.06	0.10
Security	4.69	0.60	968	4.68	601	4.70	367	-0.03	0.10
Infrastructure	4.16	0.82	965	4.17	600	4.16	365	0.01	0.01

Observations weighted by sampling weights. T-tests are computed on the *Difference* across the two groups. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Δ is the normalized difference from Imbens and Rubin (2015), equal to the difference in means, scaled by the square root of the average of the two within-group variances.

assign the attendance status to a given taxpayer if both her reported attendance in the endline survey and the actual attendance as collected during the training are the same. Among the reasons for not attending, the most common are *Busy at work - No time* (32%) and *Sickness (his or his parents')* (30%), followed by *Lack of information/reminder - Forgetfulness*(13%). While I do not argue that attendance is random, the fact that these reasons are rather ad-hoc may suggest that there is some element of randomness in the decision to attend. As a matter of fact, very few non-attendees gave answers that would make me think attendance is related to unobserved factors (e.g. didn't think it would be useful, or prefer to learn from other sources).

Next, the remaining columns of Table 2.3 display the differences between attendees and non-attendees. From the observable characteristics on hand, it seems that self-selection into the training does not represent an issue. The table shows that most variables are balanced.

sample (formed by non-attendees) was removed from the calculations, hence the higher rate.

Only three background covariates appear to have a statistically significant difference between the two groups: *age*, *large*, and *Kigali*. While the latter reflects higher attendance rates outside of the capital, differences in the former two variables are quite small in magnitude. Attendees are on average 1.6 years older and 4% more likely to be large. The difference in the *compulsory* reason for attending is also predictable: non-attendees are much less likely to be required to attend the training due to their role in the business. Likewise, I do not believe that a difference of 7% in one of the reasons to formalise is relevant to create imbalance, given that respondents are already registered and all visible to the authority. Additionally, following Imbens and Rubin (2015), if I examine not just the statistical significance of balance tests but the *normalised difference* Δ across treatment and control, there are no normalised differences greater than 0.25, the benchmark they suggest in testing for balance. This confirms that the two groups are balanced and the differences of unbalanced covariates are not significant once I make them scale independent. Nonetheless, to be conservative, all covariates, including the three imbalanced ones, are included in the X_i vector in the main specification (see section 2.4.3).

More importantly, there is no difference in baseline knowledge and perceptions between attendees and non-attendees. In principle, it was correct to expect that taxpayers with lower knowledge were more likely to attend a training. As shown in Table 2.4, key perceptions that are expected to be correlated with the decision to attend, such as the fear of audits or State legitimacy, are similar across groups. These attitudes are also likely to be correlated with other unobservables, such as intrinsic motivation to learn about taxes. The *as if random* balance between attendees and non-attendees gives some relevance to the naïve and PSM estimation, as well.

2.5.3 Did the training and coaching increase tax filing?

A unique feature of this study is that it combines survey data with administrative data (section 2.4.1). The analysis of tax returns for the fiscal year 2017 provides direct evidence on the capacity of education strategies to enhance revenue collection. In this section, I start the analysis by addressing my main research question: does taxpayer education affect compliance?

Training Table 2.5 Panel A reports naïve impact estimates on the probability to declare (column 1) and positive filing (the opposite of nil-filing, column 2) as well as on the amount of tax declared (column 3). Outcomes for declaring and positive filing are dummy variables, while tax declared is operationalised by log(tax+1).[81] By construction, non-filers and nil-filers are assigned a tax amount of zero. All treatment effects in the table are to be interpreted relative to the control group who did not receive any educational inputs. Coefficients in columns 1-2 are marginal effects from probit regressions evaluated at the mean of other covariates, while columns 3 reports marginal effects from tobit regression. Given the large share of nil-filers reporting zero income, the tobit model is well suited to censored data since it allows to include nil-filers in the sample.[82] Specifications include two sets of control: taxpayer's covariates (section 2.4.3) and a set of tax knowledge and attitudes variables from at baseline.

According to the naïve estimates, the training is highly effective across all three tax compliance outcomes: it increases declaration rates by over 9 percentage points (column 1), decreases nil-filing by 7.5 percentage points (column 2), and increases log remitted tax by 0.372 (column 3). All these effects are statistically significant and economically large. When estimates from PSM are considered in Panel B, it can be noticed that, while impacts are consistent and significant for filing and positive filing probability, the significance vanishes for the amount of tax declared. This finding is suggestive of some potential self-selection bias distorting the naïve estimates: OLS specification could be overestimating impacts on tax declared probably because those taxpayers who are more likely to remit more taxes are more likely to self-select into the training.

The more robust estimates from the IV strategy in Panel C further address a possible selection bias. Once instrumenting training attendance with the random invite, only the impact on the probability to file remains significant, with an increase of 15 percentage points, while I lose evidence on the other two outcomes. This may be due to the larger

[81] I do not consider log(tax), as log(0) is undefined. This involves throwing away information, and ignoring a significant segment of the sample who is nilfiling (62%). Logs are also the best option to control for highly skewed distribution, such as the one of tax declared.

[82] For the same reason that motivates me to adopt it here, the tobit model has been used in other studies in the tax experiments literature, (Slemrod and Weber, 2012; Alm and McClellan, 2012; Alm et al., 2010; Coricelli et al., 2010).

standard errors that are inherent to the two-stage IV estimation – or it could be a genuine reflection of ineffectiveness of the program on these outcomes, once a more rigorous estimation method is used. Importantly, the lack of a detectable effect on the tax amount is in line with the results of the only other study on a similar topic, Chetty and Saez (2013).

In sum, the training seems to be highly effective in increasing the probability to file when considering the most conservative method. For the comparison group, non attendees, the declaration rate is 35%. Compared to this level, an increase of 15 percentage points corresponds to a 43% increase in the probability to declare, which I attribute to the program.

Coaching Once gauged the impact of the training, I now turn to the coaching intervention. Capacity constraints affected the implementation of the service, as this pilot program resulted to be relatively burdensome for the tax administration, even if it only involved a relatively small number of taxpayers. About half of the taxpayers assigned to the coaching actually received it. To take this partial take-up into account, Table 2.6 reports results using both ITT (intention-to-treat) and TOT (treatment-on-treated) specifications. While the former considers all taxpayers randomly assigned to coaching as intentionally treated, the latter uses the random allocation as an instrument for treatment status (i.e. the 160 that were effectively called in both rounds).

Although anecdotal evidence suggests that taxpayers appreciated the coaching call, results do not seem to show that this pilot was successful, especially when compared to the main program's effects (see Table 2.5).[83] This is potentially due to low compliance with the treatment, both because of the limited number of taxpayers who were actually reached (even if this is taken into account with TOT) and the fact that some of those reached did not have any question or only spent a very short time on the call, as explained in section 2.4.2.

[83] Note that in Table 2.6 I have dropped the group of attendees, which is why the the number of observations is smaller than in Table 2.5. However, results are still not statistically significant if I include them and control for their attendance status.

Table 2.5: Program impact on tax outcomes

	(1) Declare	(2) Positive	(3) Log Tax
		Panel A: Naïve estimation	
Training	0.094**	0.075***	0.372***
	(0.041)	(0.024)	(0.084)
Controls	Yes	Yes	Yes
Control group Y	0.35	0.11	0.860
Observations	969	969	969
		Panel B: Kernel Matching	
Training	0.08**	0.09**	0.34
	(0.031)	(0.030)	(0.712)
Common Support	807	807	807
Ps-R^2 matched	0.002	0.046	0.056
LR chi^2 matched	2.14	25.78	29.18
		Panel C: IV Strategy	
Training	0.15**	-0.01	0.45
	(0.05)	(0.07)	(0.68)
Controls	Yes	Yes	Yes
R^2	0.22	0.03	0.01
Observations	2378	2378	2378

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *Declare* and *Positive* are dummy variables with value 1 if the taxpayers declared and positively filed, respectively. *Log Tax* is the log(tax+1) transformation of the raw tax amount variable. In Panel A, coefficients are marginal effects from probit (1-2) and tobit (3) regressions evaluated at mean. The control group are taxpayers who did not receive any educational input. Covariates and Knowledge and Perceptions used are those indicated in section 2.4.3. All regressions weighted by sampling weights. In Panel B, standard errors are bootstrapped from 999 repetitions. In Panel C, Training is instrumented by the invite call. Coaching group has been dropped. Refer to Table A26 for first-stage results.

Table 2.6: Impacts of Coaching on Tax Outcomes

	(1) ITT Declare	(2) TOT Declare	(3) ITT Positive	(4) TOT Positive	(5) ITT Tax	(6) TOT Tax
Coaching	0.02	0.03	-0.13*	-0.06	0.04	0.23
	(0.05)	(0.06)	(0.07)	(0.08)	(0.09)	(0.93)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Control Y	0.38	0.38	0.68	0.68	2.53	2.53
R^2		0.312		0.352		0.283
Observations	618	618	618	618	618	618

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Coaching group is compared to control group, i.e. taxpayers not attending. Training group is dropped. Covariates used are those indicated in section 2.4.3.

2.5.4 Longer-term effects: the role of habit

As discussed in section 2.2.4, taxpayer education provides information in a more in depth and comprehensive way than other interventions, such as informational messages that have been used in the literature on behavioural nudges. As such, I would expect it to generate learning effects that extend beyond the first year post-intervention. To test this expectation, this section investigates whether the effects on compliance last over time. In addition to the data relative to fiscal year 2017, which I used for my main results, I also obtained data for declarations relative to fiscal year 2018, filed between January and March 2019. I use these data to track taxpayers in the survey and encouragement design groups, and check whether the program effects last in the second year after registration. The estimation methods are the same as those used in section 2.5, but the outcome is now measured in year two, that is fiscal year 2018.

Appendix Table A14 reports results only on declaration rates, since it is the most robust result from the main analysis. Table A14 provides some initial evidence that the program continues to have a significant impact on the probability to declare in the second year after registration.[84] In Panel A, the coefficient of interest remains statistically significant,

[84]In the same fashion, Appendix Table A15 shows the results on outcomes measured in year 2 for the whole population of new taxpayers.

whether I control or not for filing status in year 1. The effect size of 10 percentage points is comparable with the one from year 1 (see Table 2.5). The IV strategy produces a similar impact estimate in column 1, which however turns insignificant when I control for filing status in year 1.

I reinforce these results in two ways. First, I provide some descriptive evidence on the persistence in declaration behaviour. I look at administrative data for taxpayers registered before the study, for which I have a longer time series. A taxpayer who registered in 2015 and made a declaration in the first year (year 1) has a 55% probability of filing again the following year (year 2) and a 86% probability for the year after that (year 3) – conditional on having declared in year 2. On the contrary, a taxpayer who failed to file in the first year has a negligible probability of ever making a declaration: 1% in year 2 and 0.2% in year 3. The same pattern can be shown for the cohort of new taxpayers who registered in 2017: those who declare in the first year are much more likely to keep declaring (77%), while others are very unlikely to ever submit a declaration (6%).

Second, I recur to mediation analysis to quantify the contribution of filing behaviour in year 1 caused by the intervention to its overall impact on filing behaviour in year 2. I estimate the average controlled direct effect (ACDE) of the training (Vansteelandt, 2009; Acharya et al., 2016). The mediation analysis framework implies that the training can have both a direct impact on fling behaviours and an indirect impact which is channeled through a mediator – in this case, filing behaviour in year 1. In this framework, the ACDE is defined as the direct effect of the training, that is, the effect that the training would have if the mediator (filing behaviour in year 1) was not allowed to respond to the intervention and hence the indirect effect was removed.

The main assumption in order to identify the ACDE is the *sequential unconfoundedness*. According to Acharya et al. (2016), this amounts to assuming that there are no omitted variables which confound the effect of the mediator on the outcome, conditional on treatment and a set of pre-treatment controls (Acharya et al., 2016). The identification of the ACDE proceeds in two steps. First, I regress the outcome (the filing probability in year 2) on the mediator (the filing probability in year 1), the training dummy, a set of controls, and the interaction between the mediator and all other variables. I then derive the predicted value of the outcome fixing all mediators to zero. This is the *demediated* outcome, in the words of Acharya et al. (2016). In the second step, I regress the demediated outcome on the training dummy. The coefficients from this regression represent the ACDE estimate. I obtain that the proportion of the total effect of the trainings on filing behaviour in year 2 that is mediated by filing behaviour in year 1 is quite high – 59%.

These figures are highly suggestive that getting taxpayers into the habit of declaring is crucial to their future compliance behaviour. A recent paper has shown that habit can be a powerful driver of compliance – a finding that is consistent to the separate, but related, literature on get-out-the-vote (Dunning et al., 2017). Further research is needed on the medium and long-term effects of similar interventions.

2.6 Unpacking the training impacts

2.6.1 Evaluating alternative mechanisms

The previous section established the effectiveness of the training program on tax compliance, particularly on the extensive margin (i.e. declaration rates). I now investigate what is the channel through which this effect comes about. If compliance costs and taxpayer confusion are the main mechanisms (as discussed in section 2.2.3), I would expect to see program impacts particularly on knowledge and taxpayers' perceptions about the complexity of the tax system. Thanks to the survey, I can measure both variables, as well as others capturing potential alternative mechanisms. Arguably, there are at least two other plausible ones that might determine the final effect on compliance. The first one is deterrence and enforcement. It is perfectly plausible to think that just the fact of being in touch with RRA officials, who delivered the program as trainers, acted as a deterrent to non-compliant behaviour. The second one is the set of factors captured by tax morale. The hypothesis here is that the program might have improved taxpayers' perceptions on trust and social duty, amongst others, as they met friendly and helpful RRA officials whose main objective was to provide useful information, rather than aggressively enforce tax laws. Peer effects might have also been at play, as attendees found themselves in the same room as other similar businesspeople who were also trying to navigate the tax system.

Knowledge Comparing the knowledge index before and after the training intuitively suggests that the training has been effective. On average, the knowledge index increases from 3.4 to 4.5. This change is largely driven by attendees, 86% of whom see an increase in their knowledge index.[85] On average, attendees' gain is 91% of their initial knowledge level, while it is only 30% for non-attendees.[86]

Table 2.7 Panel A, in columns 1-3, reports results from a simple OLS regression explaining taxpayer knowledge with attendance. The first row shows that the *Training* variable always displays a 1% level of significance and large coefficient, while coefficients for controls are omitted for brevity. Magnitudes suggest that trainees are 27% more likely to have an ex-post increase in knowledge (column 1), have a 57% higher knowledge gain over the baseline (column 2) and about 1.3 knowledge points (out of 10) than non-attendees (column 3). Considering the baseline value of 3.3, this coefficient suggests a 40% improvement in knowledge as a result of the training. When estimates from propensity scores matching are added in Panel B, results remain the same.

Perceptions As far as perceptions are concerned, the training does not appear to have the same positive impact it has on knowledge. For the sake of brevity, Table 2.7 column 4 reports impacts on the only perception that considerably changes after the training. Using a probit regression in Panel A, it results that the training reduces the perceived complexity by over 9 percentage points, significant at the 5% level, suggesting that attendees are 14% less likely than non-attendees to consider the tax system difficult to deal with. The same result is confirmed by the PSM strategy in Panel B, becoming more statistically significant.

The positive result suggests that the training is successful in correcting perceptions about the complexity of the tax system ("it is difficult to file an income tax declaration") and in bringing the tax administration closer to taxpayers ("it is difficult to to get in touch with RRA"). More specifically, Appendix Table A11 splits the complexity index in the

[85] Reasonably enough, those attendees who do not see an improvement after the training are already scoring the best at the baseline, with an average of 4.9 points out of 10. Instead, those who experience an increase are starting from a lower level of knowledge, 3.2 points on average.

[86] Although I do not expect any direct impact of the training on non-attendees, there are two possible explanations for this increase. First, they may have received some information about the training indirectly (i.e. spillovers). Second, they might have informed themselves about the knowledge questions asked at the baseline, so that they were better prepared at the endline.

two components, showing that most of the reduction is driven by the strong impacts on the perception on how difficult is to file a return.

Table 2.7: Program impact on tax knowledge and complexity perceptions

	(1)	(2)	(3)	(4)
	Increase	Gain	Index	Complexity
	Panel A: Naïve estimation			
Training	0.27***	0.57***	1.28***	-0.09**
	(0.03)	(0.07)	(0.09)	(0.04)
Controls	Yes	Yes	Yes	Yes
Control group Y	0.59	0.30	3.86	0.57
R-sq.	0.211	0.360	0.507	0.098
Observations	820	815	820	815
	Panel B: Kernel Matching			
Training	0.25***	0.58***	1.25***	-0.09***
	(0.03)	(0.07)	(0.10)	(0.03)
Common Support	818	813	818	813
Ps-R^2 matched	0.002	0.002	0.002	0.002
LR chi^2 matched	1.74	1.87	1.74	1.87

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *Increase* is a dummy variable for a *ex-post* increase in knowledge. *Gain* is the ratio of difference in scorer over the 0-10 pre-training score. *Index* is the 0-10 rescaled knowledge index. Controls used are: age, female, large, business life in months, primary education, university education, Kigali, email use, books of accounts, previous business, previous training, tax time in days, respondent manager of taxes. *Complexity* is a 0-1 index from two indicators of perceptions of complexity. In Panel B, standard errors are bootstrapped from 999 repetitions.

Contrarily, there are no significant impacts on any other perception (Table A12). There is some weak evidence on the enforcement mechanism, but the increase is very small in magnitude (0.006), compared to a very high baseline average of 95%. The PSM impact on social duty (and on trust) is negative but still negligible if compared to the 98% average at the baseline. Overall satisfaction with public goods is not affected by the training either, with the only exception of satisfaction with water and sanitation (Table A13). While I cannot fully rule out changes of tax morale due to the training, the data available does not support it as a significant mechanism in this case.

These results, taken together, are highly suggestive that the program's effects on compliance largely occur through improvements in taxpayers' knowledge and perceptions about complexity. Figure 2.1 visually displays the marginal effects on each perception index and a dummy for knowledge increase.[87] By far, the most sizeable effect is found with tax knowledge and perceived complexity. These variables capture taxpayer confusion and compliance costs, intended as the cognitive costs to access information and understand how the tax system works. This mechanism is consistent both with the recent literature highlighting the importance of these aspects of compliance (see section 2.2.3), and with the nature of the program that aims to educate taxpayers about basic elements of the tax system in a context of very low taxpayer knowledge.

2.6.2 Heterogeneity

After having shown the main results in sections 2.5, I study here the heterogeneity of treatment effects along several taxpayer features. As a note of caution, these effects can only be measured for the survey sample and therefore the most robust IV strategy cannot be used. The following results are produced from the naïve specification in section 2.4.3. In order to be conservative, I show heterogeneous results for filing probability only, since the IV strategy showed no significant effects for positive filing and tax remitted.

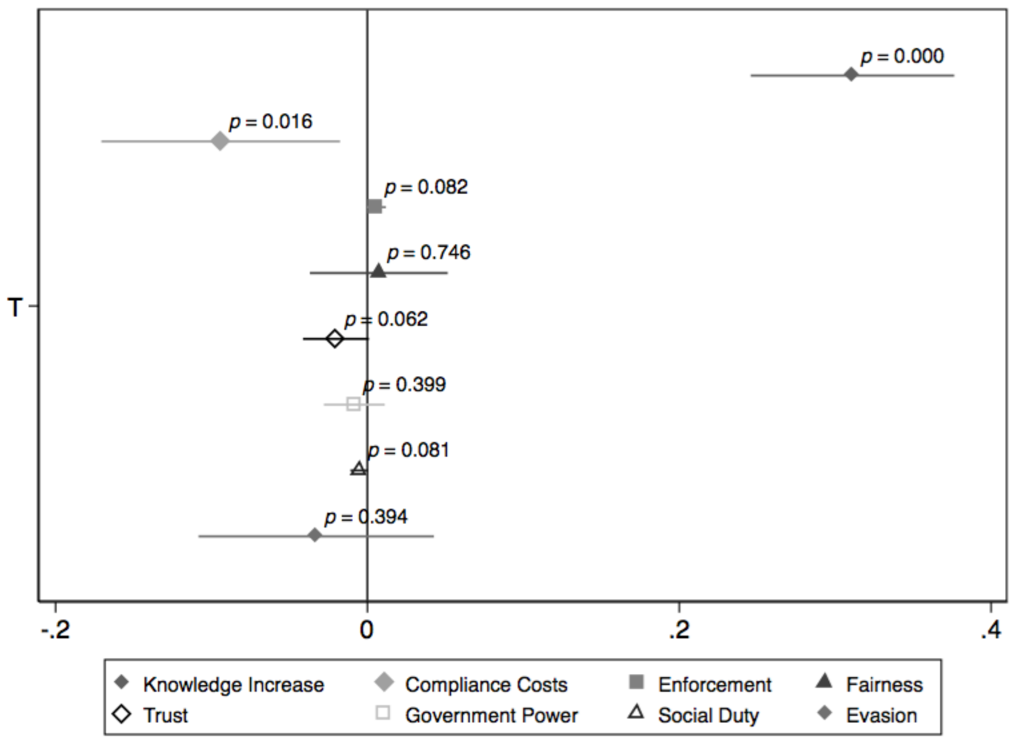
In order to detect any difference in how the training affects the participants, I include an interaction term of *Training* with key background variables, i.e. dummy variables for being a female, being located in Kigali, having completed primary education and running a large company (> 5 employees). For each variable, I partition the sample into two groups: individual belonging to group G and the others, according to the indicator variable g_i. Following the same notation as in section 2.4.3, the new specification is:

$$Y_i = \alpha + \beta \mathbb{1}\{Training\}_i + \theta g_i \mathbb{1}\{Training\}_i + \lambda g_i + \gamma Y_{0i} + X_i \Gamma + \epsilon_i \tag{2.4}$$

With this estimation, the estimated treatment effect of being in *Training* for taxpayers in the group G is $\beta + \theta$, while it is β for taxpayers not in G. θ reflects the difference in

[87]This last outcome is chosen to ease comparability with the other 0/1 variables.

Figure 2.1: Training impact on knowledge and perceptions



Coefficients are marginal effects from probit regressions computed at the mean of the explanatory variables. Perceptions are indexes which range from 0 (disagree) to 1 (agree). All regressions are weighted by sampling weights.

effects across the two groups. The significance of θ implies that there exists heterogeneity across the two groups. The level effect on Y_i of belonging to the group G is λ.

Appendix Table A16 reports the heterogeneity in impacts on the filing probability. An additional indicator variable for the top two deciles of baseline tax knowledge is explored as well. The coefficient of the interaction terms indicate the additional incremental effect of being in group G, on top of the effect of *Training*. Some weak evidence (10% only) on heterogeneous effects is found for female (col. 2) and taxpayers with lower education (col. 4). Interestingly, *Training* estimate is not significant for male only (row 1 column 2), thus indicating that female are more affected from the training when choosing to declare.

In the same fashion, Appendix Table A17 displays the heterogeneity in impacts on tax knowledge. Two outcome variables are used: a dummy for knowledge increase and the percentage gain over the baseline level. The most striking feature of this analysis is the consistency in the estimated treatment effects of *Training* in row 1: for all subgroups, the coefficients are positive and highly statistically significant. Only two interaction terms are statistically significant (5% level): being in Kigali has an additional positive effect on knowledge increase with respect to being in rural areas, while being large has a negative effect on knowledge gain as compared to being small. This is consistent with the fact that large firms have already more resources to navigate the tax system (e.g. dedicated accountants, consultants) as reported in Tanzi (2012). Smaller, unequipped and less educated taxpayers are benefiting more from the training. No heterogeneity is found for gender and education.

2.7 Robustness

2.7.1 Additional evidence from trainings across the country

In this section I address the issue of external validity. In other words, do results hold when looking at the whole population of new taxpayers in Rwanda? For those registered in 2017, both attendance and compliance from administrative data are available, as well as the limited set of administrative covariates (described in section 2.4.1). I can therefore estimate the relationship between program attendance and the three compliance outcomes, for the universe of new taxpayers. However, in this case I am unable to adopt any of the more rigorous estimation methods used elsewhere, such as IV, PSM, or even the inclusion of a large set of controls from the survey (as in section 2.5) as it was not implemented in these trainings.

Table A18 reports these results for the whole population of new taxpayers who registered in 2017. I start by re-estimating the coefficients of interest in the survey sample (column 1) and in the encouragement design group (column 2), but this time using a simple OLS and including only the administrative covariates.[88] The results on the probability

[88] A dummy for being incorporated (CIT), life in months, months since the training, being registered for VAT.

to declare, the strongest result so far, are very similar to the ones reported in Table 2.5, which are estimated with more robust specifications. This gives me confidence on two fronts. First, it confirms that choices over the control variables do not affect results in any substantial way: the coefficients of Table A18 are well within the range of those in Table 2.5. Second, and relatedly, it suggests that the results on the population of new taxpayers may not be overly biased, despite the more limited options I have in terms of methods and data.

Column 3 reports results for the full population of new taxpayers. It shows a much larger effect of almost 30 percentage points, compared to 12 and 14 p.p. in columns 1 and 2. Running separate regressions for rural (Table A19) and urban areas (Table A20) does not help to explain this difference, as the relevant coefficients are always much larger in the broader population than in the other groups. The most plausible explanation for this difference lies in the timing of sessions, which is reported in Table A2. While the three relevant sessions for both the survey sample and the encouragement design group happened in August 2017, the other sessions were organised between late November 2017 and March 2018 – with an almost three months gap in between. The latter timing certainly made the program a lot more salient for taxpayers, who were approaching the end of the year (31 December) and the declaration period (January to March). On the contrary, the August sessions were still quite far from the time when most taxpayers start wrapping up their accounts for the year and think more concretely about their tax affairs. Despite their limitations, these results suggest that, if anything, my main results are a lower bound for potentially much larger effects of this type of intervention.

Panels B and C of Table A18 estimate the program's effect on the other two compliance outcomes, positive filing and the tax amount, for the whole population. As for the declaration probability, the results on the other two compliance outcomes are very similar to the main results reported in section 2.5. However, once I consider the whole population, the coefficient on positive filing and on the tax amount are both highly significant and of similar magnitude to those presented in Table 2.5. Perhaps the most interesting result here is on the tax amount, which seems to sizeably increase for attendees in the broader population. While this result is far from conclusive, it hints at the possibility that the program has also some effect on the intensive margin too – especially when the program is attended closer to the declaration period.

2.7.2 Alternative knowledge outcomes

Table A21 reports estimates on tax knowledge using a couple of alternative outcome variables: i) difference of the post-training minus the pre-training scores, and ii) the Kling index as explained in section 2.4.1. Estimates from column 1 shows that the training causes a difference of 1.26 points, in line with the results from Table 2.7. For what concerns the Kling index in column 2, an effect-size of 0.31 means that the average person in the training group would score higher than 62% of a control group that was initially equivalent.[89] Naïve and PSM estimates are highly consistent.

In the same fashion, Appendix Figure A7 displays the OLS coefficients from regressing training attendance on each of the 19 knowledge items. The items who see a larger impact are those related to knowledge on VAT and PAYE, as well as knowing the deadline and e-filing. Importantly, the latter are directly related to the basic concepts around correct filing, and corroborate the strong impact on the filing probability in Table 2.5. At the same time, somehow counterintuitively, attendees are not learning much about the penalties' structure and the tax rates – a pattern which is confirmed in the Eswatini taxpayer survey of Chapter 3. This might suggest that attendees are not complying out of fear of deterrence, but rather because their compliance costs are decreasing.

2.7.3 Alternative tax outcomes

When considering a multidimensional concept such as tax compliance, other indicators of it can be created, on top of the three main ones in section 2.5.3.

First, I check whether the program had any impact on the probability of making a pre-payment following the declaration. Pre-payments are quarterly payments on account, based on the previous declaration's tax amount and divided in four payments of equal amount (one for each quarter), that need to be performed throughout the year. Compliance with pre-payments is far from being optimal in Rwanda (Mascagni et al., 2016). I find no

[89] If the group consisted of 100 people, this is the same as saying that the average person (i.e. ranked 50^{th} in the group) would now be on a par with the person ranked 38^{th} in the control group.

effect of the training on the probability of making a pre-payment, as shown in column 1 of Appendix Table A23. This is probably due to the fact that data on pre-payments is only available for those who made a positive tax declaration, which is a rather restricted sample in my case, as taxpayers failing to declare or nil-filing would not be required to make a pre-payment.

Second, I also consider the effective tax rate (ETR), usually identified in the literature as a proxy for the tax burden. The ETR is defined as the ratio of tax declared over business income and, in presence of mistakes due to poor tax knowledge, is likely to be larger than the correct level.[90] Focusing on the ETR allows to compare taxpayers from different tax regimes and to understand whether taxpayers benefit in terms of reduced mistakes and lower burden. Results from column 2 of Appendix Table A23 show that the training reduces the ETR by a small fraction (about 1%) but results are not significant at any statistical level. This may be mostly due to the small sample size, since ETRs are calculated only for taxpayers with a positive amount of income declared.

Moreover, for what concerns the amount of tax declared, in addition to the log(tax+1) used in the main specification, Appendix Table A22 reports coefficients for different transformations of the tax amount variable: raw amount in Rwandan Francs, winsorised amount at the 99$^{\text{th}}$ percentile and the inverted hyperbolic sine transformation (IHS).[91] While estimates for both raw and winsorised amounts in RWF are not consistent across the three panels, the more reliable variable IHS does not show any significant change, in line with the findings on log(tax+1) in Table 2.5.

[90] It is important to note that the resulting ETRs are expected to be substantially lower than the statutory tax rates, because the denominator here is overall income instead of taxable income.

[91] IHS is defined as $\log(\text{tax} + (\text{tax}^2+1)^{1/2})$. Except for very small values of y, the inverse sine is approximately equal to log(2y) or log(2)+log(y), and so it can be interpreted in exactly the same way as a standard logarithmic dependent variable. It also accounts for negative values of tax. Relative to the most commonly adopted $\log(x + 1)$ way of dealing with truncated data, the IHS transformation is an increasingly more accepted solution, despite dating back to the late '80s (Burbidge et al., 1988; MacKinnon and Magee, 1990). A nice application is given in Pence (2006).

2.7.4 Alternative matching algorithms

When producing the PSM results in section 2.5, I adopted the Kernel matching as it is usually seen as the most robust algorithms available. To check the robustness of the PSM results, other algorithms have been used, whose description is given in section 2.4.3. Appendix Tables A24 and A25 shows results from the N-N matching and radius caliper techniques, for both knowledge/perceptions and tax outcomes, respectively. Impacts on knowledge and perceptions stay consistently significant, except for the complexity costs when using the N-N matching (Table A24). Likewise, in Table A25 the filing probability is increasing with the training, consistently to what found in Table 2.5 Panel B.

2.7.5 More on the IV strategy

Appendix Table A26 provides some additional results from the IV strategy used in section 2.5. First, in Panel A column 1, the first-stage regression results are shown (as well as the IV estimates from Table 2.5 Panel C). The instrument is highly relevant, i.e. the first condition for a robust instrument is met. Second, Panel B shows estimates from the so-called reduced-form regression of tax outcomes on the IV. It can be seen that the IV is significantly correlated with *Declare* only. This confirms again the significant impact on declaration rates derived in Panel A (or, similarily, in Table 2.5 Panel C).

Moreover, I also attempt to check whether the exclusion restriction is met (section 2.4.3). Appendix Table A27 reports estimates from a regression where both the instrument (IV Surveyed) and the instrumented variable (Training) are included to explain tax outcomes. Ideally, an instrument should not have any additional impact beyond the effect it generates through the instrumented variable. Panel A estimates suggest that IV Surveyed is never significantly explaining tax outcomes. This provides some evidence that the exclusion restriction is not violated.

2.7.6 The effect of participating to the survey

As a final check, I address the concerns surrounding the possible side effects that participation to the survey might produce on final tax outcomes – other than through the mechanisms discussed in section 2.6.1. There is substantial evidence for this concern, notably for developing countries. Zwane et al. (2011) show that completing household surveys has consequences on later behaviour, such as water treatment product use and medical insurance take-up (though not on borrower behavior). If there is such an effect, then IV results may either under- or over-state the consequences of the education program per se.

I test for the null (or negligible) effect of survey participation in two ways. First, I compare tax outcomes of two groups, surveyed non-attendees and non-surveyed non-attendees, which differ only in terms of receiving the survey, i.e. the instrument. This can also be considered as a placebo test on the validity on the instrument (survey participation). If completing the survey has an effect which is independent on attending the training, I would expect to see surveyed non-attendees displaying statistically significantly different tax outcomes than non-surveyed non-attendees. As shown by the t-tests in Appendix Table A28, this is not the case and the two groups file following a similar pattern. This confirms both the lack of any relevant effect from completing the survey and that the instrument respects the condition of exclusion restriction.

Second, I compare the filing behaviour of taxpayers who completed both survey rounds and those who completed only the baseline, i.e. those who attrited. As shown in Zwane et al. (2011), subsequent follow-up surveys can alter later behaviour. Assuming that participating to the followup induces the desired behaviour, if such an effect from the follow-up survey actually exists, I would notice that non-attrited taxpayers are, for example, more likely to be compliant. In Appendix Table A29, I test whether any significant differences exist between the two groups. None of the key tax outcomes result to be different.

In conclusion, participating to the survey does not introduce any significant bias in the estimates of impact. This is likely due to the fact that the surveys took place in August 2017, about 7 months before the filing deadline, and to their shorter duration when compared to the household surveys reviewed in Zwane et al. (2011).[92]

[92] In addition, the surveyors were not and did not present themselves as representatives of the tax authority but independent contractors of a private survey firm.

2.8 Conclusions

Several tax authorities in the African continent are devoting increasing resources to implement educational initiatives to improve taxpayers' knowledge of the tax system (Mascagni and Santoro, 2018). However, empirical evidence on whether tax knowledge can be improved with these strategies and which are the effects in terms of compliance is absent – despite the increasing evidence that taxpayer confusion and compliance costs are key determinants of tax compliance. This paper argues that knowledge of the tax system is crucial in explaining compliance and test the effectiveness of two tax educational inputs. More specifically, I evaluate a taxpayer education program run by the Rwanda Revenue Authority and offered to newly registered taxpayers. At the same time, I measure the impacts of an alternative strategy, a personalised tax coaching. I am particularly interested in capturing the programs' effects on behaviours (tax compliance), as well as on tax knowledge and perceptions.

I investigate my research question by relying on a unique dataset that includes information from phone surveys and from tax records, combined using uniquely identified taxpayer identification numbers. By doing this, I can address some of the common pitfalls both of surveys (e.g. difficulty to measure evasion) and of administrative data (e.g. lack of comprehensive information on taxpayers' characteristics). Although I could not randomise attendance to the program of interest, I am confident of the robustness of my results, based on three methods: a naïve estimation based on two largely comparable groups (see section 2.5.2), propensity-score matching and an unintended encouragement design (section 2.5.3).

My key result is a large and statistically significant effect of the program on declaration rates, which increase by 43% with respect to the control group. Additional compliance outcomes such as positive filing probability and amount of tax declared do not significantly change. The fact that the effect on tax declared is particularly weak is in line with the only other article on this topic (Chetty and Saez, 2013), which shows no effect on reported income for an education program in the United States. These results highlight the importance to look at multiple aspects of compliance, especially in a context such as Rwanda where failure to declare and nil-filing are highly prevalent phenomena (section 2.3.2). Moreover, and contrary to much of the literature on tax compliance, I present suggestive evidence that the program's effect stretch beyond the first year (section 2.5.4). Based on administrative data, I argue that habit plays an important role in tax compliance (Dunning et al., 2017).

Furthermore, thanks to the rich set of survey variables, I test possible alternative mechanisms through which this effect may come about. I argue that the most plausible mechanism is increased knowledge and improved perceptions about the complexity of the tax system (section 2.6.1). I show that training taxpayers about the tax system has a large, 40%, increase on knowledge and a 15% decrease in the probability of perceiving the tax system as complex. This increase happens in a context where I see dramatically low levels of tax knowledge at the baseline, as I show in section 2.5.1. While this is in line with previous evidence on taxpayer knowledge (see section 2.2.3), I use a more accurate and richer measure of knowledge, which is an original contribution to this literature.

Relatedly, the pilot of the RRA-led caching program (section 2.4.2) suggests that a more personalised option is currently ineffective and hardly scalable. The failure to find any significant impact is mostly due to poor implementation in the field. Taking into account the resource constraints within tax authorities in SSA, programs of this type could only be adopted for a limited, targeted group of taxpayers, or based on specific needs, rather than be implemented at scale.

In terms of policy, my results provide a few important insights. One immediate recommendation for the revenue authority would be to provide some basic information to taxpayers at the point of registration, for example stating clearly what taxes they registered for and what are the related obligations.[93] This could be as simple as a brief script for tax officials, and a document to leave with taxpayers containing their main duties, rights and responsibilities. This would help tackle gaps in knowledge, along with other measures.

Secondly, my results also confirm that the existing program is effective to increase compliance, at least on the extensive margin (i.e. declaration rates). From a policy perspective, this means that policymakers may not see immediate revenue gains from such programs. Nonetheless, I argue that bringing new taxpayers into the habit of submitting declarations is crucial to ensure future compliance. As such, investing in taxpayer education seems to

[93] In a parallel study targeting nil-filers, Mascagni et al. (2020) show that registration procedures are often confusing to new taxpayers.

be a sensible strategy – as part of a broader range of measures to improve compliance. For instance, more could be done to increase the program's reach, which is currently quite limited (section 2.5.1). This implies, for example, insuring that the department in charge of taxpayer education initiatives (in RRA's case, Taxpayer Services) has adequate capacity to effectively manage such programs.

Last but not least, this study makes an argument in favor of collecting detailed data for interventions carried by revenue authorities in the field. Data on project implementation is crucial for evaluating the project impacts and using these results to inform evidence-based policy. Carrying out the training evaluation with the data previously available would have been impossible, as I did not know which taxpayers attended and which ones did not. This practical example illustrates the broader issue that good data is an essential foundation to evidence-based policymaking. I would therefore encourage to embed proper data collection in all initiatives where that is possible at a contained cost and effort, as it would be in this case.

Appendices

A1 Background

A1.1 Deviations from Neoclassical Model

The main theoretical formulations which attempted to improve the neoclassical model of tax evasion can be grouped as follows:

- *Intrinsic motivation and psychic costs*: taxpayers experience feelings of pride, positive self-image or warm glow from contributing (Andreoni et al., 1998), or guilt or shame for failing to comply (Gordon, 1989; Erard and Feinstein, 1994; Reckers et al., 1994; Traxler, 2010), according to the formulation of a remorse-based utility function (Eisenhauer, 2006).

- *Reciprocity and fiscal exchange*: taxes are part of social contract, where compliance is affected by the perceived legitimacy of the State and the quantity/quality of public services provided (Cowell and Gordon, 1988; Levi, 1989; Moore, 2013; Falkinger, 1988); the consideration of the fair amount to pay, and whether it is justifiable to evade, takes into account both the provision of public goods and the tax payments made by other taxpayers, following the concept of Kantian morality (Bordignon, 1993).

- *Peer effects and social influences*: the cost from evading is an increasing function of the proportion of taxpayers who comply (Myles and Naylor, 1996; Kim, 2003; Fortin et al., 2007).[94] Theoretical research on herd behavior (Banerjee, 1992) can be applied to tax compliance as well.

- *Cultural factors*: religious beliefs (Torgler, 2003), ethnic-linguistic fractionalisation (Alesina et al., 2003), country-level characteristics (Torgler, 2004a; Cummings et al., 2009) are all taken into account in understanding tax compliance.

- *Deviations from utility maximisation and non-standard preferences*: new formulations include bounded rationality with respect to time discounting, such as hyperbolic discounting (Chorvat, 2007), behavior under uncertainty, in which gains and losses having different utility weights and perceived probabilities, as in the prospect theory (Kahneman and Tversky, 1979; Yaniv, 1999; Dhami and al Nowaihi, 2007), limited information and law complexity (Alm, 1988; Snow and Warrren, 2005) and the framing of the tax (Copeland and Cuccia, 2002; Watrin and Ullmann, 2008).

[94]See Onu and Oats (2016) for a review.

- *Interactions between taxpayer-customers and the State provider of services*: in the so-called "slippery slope" framework, authority and trust dynamically interact and reinforce/undermine each other (Kirchler et al., 2007, 2008). Relatedly, the responsive regulation paradigm with a service-oriented approach has been developed as a new modus operandi of tax administrations (Braithwaite, 2009). Similarly, the crowding out theory formulates that deterrence policy and other external factors (rewards, regulations) negatively affect the intrinsic motivation to pay (Frey and Feld, 2002). In these theories, a psychological contract between citizens and authority is considered (Alm and Torgler, 2011; Alm et al., 2010; Alm and McClellan, 2012).

Table A1: Corporate and Personal Income Tax

	CIT	PIT
Tax rates		
Real regime	30%	0%-20%-30%
Lump-sum regime	3% turnover	3% turnover
Flat-amount regime	pre-defined	pre-defined
Fiscal Year	January to December	
Deadline	March 31	
% total # declarations (2017)	56%	44%
% total tax take (2017)	94.5%	5.5%
% in Kigali (2017)	69.6%	39.7%
Business income USD (2017)	161,645	1,921
Tax declared USD (2017)	2,600	388
Ghosts New (2015)	45%	52%
Nil-filers New (2015)	71%	28%
Ghosts New (2016)	45%	46%
Nil-filers New (2016)	77%	24%
Ghosts New (2017)	45%	58%
Nil-filers New (2017)	74%	30%

Own calculations using RRA administrative data.
Amounts in USD.

Table A2: TPS training plan for 2017/18

	District	Date
1)	Kicukiro	21 August 2017
2)	Musanze	29 August 2017
3)	Rubavu	30 August 2017
4)	Muhanga	22 November 2017
5)	Rusizi	29 November 2017
6)	Huye	6 December 2017
7)	Nyagatare	20 December 2017
8)	Kicukiro	23 January 2018
9)	Gasabo	6 February 2018
10)	Nyarugenge	9 February 2018
11)	Kicukiro	20 February 2018
12)	Gasabo	22 February 2018
13)	Nyarugenge	7 March 2018

Table A3: List of Covariates

Covariate	Description	Mean	St. dev
Age	age in years	33.08	9.37
Female	dummy for female	0.32	0.47
No school or primary	dummy for no school or primary	0.19	0.40
University degree	dummy for university degree	0.45	0.50
Owner	dummy for being the owner of the business	0.78	0.42
CIT	registered for corporate income tax	0.47	0.50
>5 employees	dummy for more than 5 employees	0.09	0.29
Life in months	# months since registration with RRA	4.62	2.01
Kigali	dummy for business based in Kigali	0.49	0.50
Email use	dummy for use email for business communication	0.25	0.44
Bank account	dummy for bank account ownership	0.44	0.50
Books of account	dummy for use of books of accounts	0.47	0.50
Tax manager	dummy for taxpayers managing taxes by himself	0.76	0.42
Had previous business	dummy for having previous business	0.21	0.41
Had previous training	dummy for having attended another training	0.07	0.25
Tax time use	# of days spent in tax per year	1.86	4.99
# RRA visits	# of visits from RRA staff since registration	1.36	2.31

Means and standard deviations measured at the baseline.

Table A4: Tax Knowledge Module

Index	Components	Baseline
E-filing	Do taxpayers always have to visit the local tax office in person to file for income tax?	0.47
Tax center	Do you know where the nearest tax center is?	0.90
Deadline	When is the deadline for your income tax declaration?	0.05
Inactive	Is the following statement true or false? If your business is inactive, you don?t need to file a declaration.	0.70
IQP	Is the following statement true or false? The quarterly prepayments are calculated as a share of the previous tax declaration.	0.36
IQP share	What share of the previous tax declaration is each quarterly prepayment?	0.03
Diff. rate	Do people pay income tax at different rate depending on their income?	0.66
Max rate	What is the maximum rate for CIT/PIT?	0.01
Tax base	What is your tax base? Note: this is the base on which your PIT/CIT tax is calculated.	0.17
PAYE	Do employers generally have an obligation to pay taxes for their employees?	0.69
PAYE exempt	Is the following statement true or false? There is a threshold for wage income under which the tax rate is 0%.	0.34
PAYE threshold	What is that threshold?	0.22
VAT freq.	Take the example of a business with a turnover of 100 million RWF. How often does this business have to file a VAT declaration in a year?	0.07
VAT rate	What is the tax rate for VAT?	0.31
VAT refund	Can all VAT-registered businesses claim VAT refunds for their inputs?	0.28
VAT pay	Are all business required to pay VAT, regardless of their income?	0.43
EBM	Is the following statement true or false? An EBM receipt should not be issued for small sales, even if the business has an EBM.	0.65
Fine failing	How high is the maximum extra penalty for failing to declare income, once you are caught with an audit?	0.01
Fine interest rate	What?s the monthly interest rate for delaying to pay or failing to declare?	0.00

Indexes are 1-0 dummies with the following meaning: 1 - agree or neutral; 0 - disagree.

Table A5: Perceptions and Attitudes before Training

Statement	Totally Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Totally Agree
1. It is difficult to file an income tax declaration	11%	28%	16%	23%	19%
2. It is difficult to get in touch with RRA officials to get information	26%	41%	3%	14%	15%
3. It is right for some people not to pay the taxes they owe on their income	25%	34%	8%	16%	15%
4. If my neighbors do not pay taxes, it is fair for me not to pay them either	55%	40%	2%	1%	2%
5. Most businesses evade their taxes in part or in full	7%	17%	27%	34%	11%
6. The tax system is fair, impartial and uncorrupted	4%	9%	18%	22%	45%
7. Businessmen to make gifts or unofficial payments to get things done	22%	26%	31%	13%	5%
8. The government can make people pay more taxes to increase spending on public health care	5%	15%	15%	27%	36%
9. The tax authorities always have the right to make people pay taxes	3%	11%	5%	25%	55%
10. Citizens must pay their taxes to the government for our country to develop	0%	0%	1%	6%	93%
11. Taxation is a social duty of every citizen	0%	1%	1%	10%	88%
12. RRA does its job professionally and with integrity	1%	5%	11%	28%	54%
13. It is very likely that someone who is evading tax will be caught and sanctioned by the Government	0%	2%	4%	20%	73%

Table A6: Perception Indexes Description

Index	Components	Baseline
Complexity	It is difficult to file an income tax declaration - It is difficult to get in touch with RRA officials to get information	0.704
Enforcement	It is very likely that someone who is evading tax will be caught and sanctioned by the Government	0.980
Evasion	It is right for some people not to pay the taxes they owe on their income - If my neighbors do not pay taxes, it is fair for me not to pay them either - Most businesses evade their taxes in part or in full. - Businessmen are sometimes required to make gifts or unofficial payments to get things done with regard to taxes.	0.582
Fairness	The tax system is fair, impartial and uncorrupted.	0.867
Trust	People in Rwanda generally trust the RRA to do their job professionally and with integrity	0.939
Govt authority	The government can decide to make people pay more taxes or user fees in order to increase spending on public health care. - The tax authorities always have the right to make people pay taxes	0.933
Social duty	Citizens must pay their taxes to the government in order for our country to develop - Taxation is a social duty of every citizen, to contribute to society's welfare	0.981

Indexes are 1-0 dummies with the following meaning: 1 - agree or neutral; 0 - disagree.

A2 Balance Tests

Table A7: Mean differences by reached status

	Unreached		Reached			
	Mean	Obs	Mean	Obs	Diff.	Δ
Kicukiro	0.74	1551	0.48	980	0.27***	-0.56
Musanze	0.10	1551	0.25	980	-0.15***	0.41
Rubavu	0.16	1551	0.28	980	-0.11***	0.27
PIT	0.45	1551	0.52	980	-0.07***	0.14
CIT	0.55	1551	0.48	980	0.07***	-0.14
Jan17	0.25	1551	0.21	980	0.04**	0.09
Feb17	0.17	1551	0.18	980	-0.00	0.01
Mar17	0.20	1551	0.23	980	-0.03*	0.08
Apr17	0.07	1551	0.07	980	-0.00	0.00
May17	0.09	1551	0.09	980	-0.00	0.00
Jun17	0.09	1551	0.09	980	0.00	0.00
Jul17	0.12	1551	0.12	980	-0.00	0.00
Life in months	4.67	1551	4.59	980	0.08	-0.04
Observations	2531					

Observations weighted by sampling weights.

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

Table A8: Covariates mean differences by attrition status

	Non-attrited		Attrited			
	Mean	Obs	Mean	Obs	Difference	Δ
Covariates						
Age	33.24	821	32.82	186	0.42	-0.04
Female	0.30	821	0.41	186	-0.11***	0.22
No school or primary	0.20	821	0.14	186	0.06*	-0.16
University degree	0.43	821	0.54	186	-0.10**	0.20
Owner	0.79	821	0.62	186	0.17***	-0.38*
Individual business	0.52	821	0.45	186	0.07*	-0.13
>5 employees	0.09	821	0.11	186	-0.02	0.07
Email use	0.26	821	0.27	186	-0.01	0.02
Bank account	0.45	821	0.44	186	0.01	-0.03
Books	0.48	821	0.51	186	-0.03	0.06
Tax time use days	2.00	605	1.32	134	0.68	-0.15
Had previous business	0.21	821	0.23	186	-0.03	0.06
Had previous training	0.06	821	0.10	186	-0.03	0.12
Willing to attend	0.99	793	0.98	176	0.01*	-0.11
# RRA visits	1.39	716	1.32	158	0.07	-0.03
Reasons to formalize						
Obeying the Law	0.70	821	0.66	186	0.04	-0.08
Better reputation	0.26	821	0.23	186	0.03	-0.07
Business too small	0.27	821	0.25	186	0.02	-0.05
Constraints to formalize						
High tax rate	0.25	821	0.20	186	0.05	-0.12
No knowledge reg. process	0.13	821	0.12	186	0.01	-0.02
No knowledge reg. benefits	0.15	821	0.12	186	0.03	-0.10
Reasons to attend the training						
I need to know more	0.94	789	0.94	173	0.01	-0.02
Compulsory to attend	0.25	789	0.29	173	-0.04	0.08
I'm curious about training	0.17	789	0.13	173	0.04	-0.12
Observations	1007					

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Δ is the normalized difference from Imbens and Rubin (2015), equal to the difference in means, scaled by the square root of the average of the two within-group variances. Observations weighted by sampling weights.

Table A9: Baseline outcomes mean differences by attrition status

	Non-attrited		Attrited			
	Mean	Obs	Mean	Obs	Difference	Δ
Baseline Knowledge Index	6.37	821	6.42	186	-0.06	0.02
Complexity	0.71	796	0.62	177	0.10**	-0.21
Evasion attitude	0.33	782	0.34	173	-0.01	0.02
Fairness	0.98	788	0.98	176	0.00	-0.02
Govt Authority	0.95	804	0.96	181	-0.00	0.01
Social duty	0.98	819	0.97	185	0.01	-0.06
Enforcement	4.39	821	3.54	186	0.85	-0.10
Health	2.68	821	1.06	186	1.61*	-0.13
Education	2.03	821	2.22	186	-0.19	0.02
WASH	2.87	821	2.15	186	0.73	-0.08
Electricity	3.56	821	2.70	186	0.86	-0.10
Security	4.56	821	3.55	186	1.01**	-0.13
Infrastructure	3.68	821	2.94	186	0.74	-0.08
Observations	1007					

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Baseline Knowledge Index* is the sum of correct answers over 19 questions (range 0-19). Perceptions in row 2-5 are dummy variable indicating whether the taxpayer agrees with a set of statement (section 2.4.1). *Enforcement* is over a 1-5 scale where 5 is totally agree and 1 totally disagree. Last 6 rows show satisfaction variables over a 1-5 scale where 5 is totally satisfied and 1 totally dissatisfied. Δ is the normalized difference from Imbens and Rubin (2015), equal to the difference in means, scaled by the square root of the average of the two within-group variances. Observations weighted by sampling weights.

Table A10: Covariates mean differences by assignment to coaching

	Control		Coaching		
	Obs	Mean	Obs	Mean	Difference
Covariates					
Age	279	32.40	293	32.56	-0.15
Female	279	0.30	293	0.31	-0.01
No school or primary	279	0.25	293	0.18	0.07**
University degree	279	0.45	293	0.44	0.01
Owner	279	0.76	293	0.82	-0.06
CIT	270	0.46	283	0.43	0.03
>5 employees	279	0.06	293	0.08	-0.02
Life in months	270	4.44	283	4.63	-0.18
Kigali	272	0.51	281	0.52	-0.00
Email use	279	0.24	293	0.25	-0.01
Bank account	279	0.45	293	0.40	0.05
Books	279	0.46	293	0.44	0.03
Tax time use days	214	1.73	200	2.24	-0.52
Had previous business	279	0.18	293	0.22	-0.04
Had previous training	279	0.06	293	0.08	-0.02
Knowledge and Perceptions					
Baseline Knowledge Index	279	6.19	293	6.25	-0.05
Complexity	267	0.71	281	0.70	0.01
Evasion attitude	264	0.32	278	0.32	0.00
Fairness	265	0.97	281	0.98	-0.01
Govt Authority	272	0.95	287	0.95	-0.00
Social duty	279	0.96	291	0.99	-0.03**
Enforcement	279	4.67	293	4.63	0.04
Pub. Serv. Satisfaction	273	0.01	283	-0.01	0.02
Observations	572				

Observations weighted by sampling weights. T-test are computed on the *Difference* across the two groups. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A3 Results

Table A11: Program impact on complexity perceptions - PSM

	Difficult to File	Get in Touch
Panel A: Naive Specification		
Training	-0.08**	-0.05*
	(0.04)	(0.03)
Controls	Yes	Yes
Baseline Perception	Yes	Yes
Observations	816	818
Panel B: Kernel Matching		
Training	-0.10***	-0.03
	(0.04)	(0.03)
Common Support	814	814
Ps-R^2 matched	0.02	0.002
LR chi^2 matched	1.79	1.90

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are from 99 bootstrapped repetitions. Note that the Ps-R^2 unmatched is always 0.042 and the LR chi^2 unmatched is always 47.13***. Perceptions outcomes are rescaled variables which range from 0 (disagree) to 1 (agree). *Difficult to File* = difficult to file an income tax return. *Get in touch* = difficult to get in touch with RRA. Controls used are: age, female, large, business life in months, CIT, primary education, university education, Kigali, email use, books of accounts, previous business, previous training.

Table A12: Program impact on attitudes and perceptions

	Evasion	Fairness	Enforcement	Trust	Govt Power	Social duty
			Panel A: Naïve estimation			
Training	-0.003	0.003	0.006*	-0.010*	-0.005	-0.004
	(0.034)	(0.005)	(0.003)	(0.003)	(0.010)	(0.002)
Controls	Yes	Yes	Yes	Yes	Yes	
Observations	811	810	820	817	817	820
			Panel B: Kernel Matching			
Training	-0.012	0.001	0.014	-0.009	-0.006	-0.018**
	(0.03)	(0.001)	(0.32)	(0.01)	(0.01)	(0.01)
Common Support	809	808	818	815	815	818
Ps-R^2 matched	0.002	0.002	0.002	0.002	0.002	0.002
LR chi^2 matched	2.37	2.15	1.74	1.72	1.72	1.74

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are from 999 bootstrapped repetitions. Note that the Ps-R^2 unmatched is always 0.042 and the LR chi^2 unmatched is always 47.13***. Perceptions outcomes are rescaled variables which range from 0 (disagree) to 1 (agree). Controls used are: age, female, large, business life in months, CIT, primary education, university education, Kigali, email use, books of accounts, previous business, previous training, tax time in days, respondent manager of taxes..

A4 Robustness checks

Table A13: Program impact on satisfaction with public goods

	(1) Health	(2) Education	(3) Water	(4) Electricity	(5) Security	(6) Infrastructure
Training	0.020	-0.018	0.066**	-0.006	0.000	0.005
	(0.028)	(0.029)	(0.032)	(0.015)	(0.000)	(0.005)
Baseline Satisfaction	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Observations	814	817	819	817	819	817

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Coefficients are marginal effects from probit regressions computed at the mean of the explanatory variables. The control group are taxpayers who did not attend the training. Satisfaction outcomes are rescaled varianle which range from 0 (unsatisfied) to 1 (satisfied). Controls used are: age, female, large, CIT, business life in months, primary education, university education, Kigali, email use, books of accounts, previous business, previous training.

Table A14: Program impact on probability to declare in year 2

	(1) Declare Y2	(2) Declare Y2
Panel A: Naïve estimation		
T	0.106***	0.070*
	(0.037)	(0.040)
Declared Year 1		0.637***
		(0.043)
Survey covariates	Yes	Yes
Observations	971	969
Panel B: IV Strategy		
T	0.107**	0.014
	(0.049)	(0.038)
Declared Year 1		0.621***
		(0.021)
Admin covariates	Yes	Yes
Observations	2,531	2,531
Control group Y	0.31	0.31

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is measured in year 2. In panel A, coefficients are marginal effects from probit regressions evaluated at mean, all weighted by sampling weights. In Panel B, Training is instrumented by the invite call. Coaching group has been dropped.

Table A15: Mid-term impact on tax outcomes: population of new taxpayers

	(1) Declare Y2	(2) Declare Y2	(3) Positive Y2	(4) Positive Y2	(5) Log Tax Y2	(6) Log Tax Y2
T	0.232***	0.082***	0.147***	0.169***	1.328***	0.548***
	(0.011)	(0.014)	(0.008)	(0.008)	(0.063)	(0.054)
Declare Y1		0.699***				
		(0.010)				
Positive Y1				0.062***		
				(0.008)		
LogTax Y1						0.145***
						(0.005)
Admin covariates	No	Yes	No	Yes	No	Yes
Observaitons	15009	15009	15009	15009	15009	15009

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Coefficients are marginal effects from probit and tobit regression computed at the mean of covariates. The dependent variable is measured in year 2.

Table A16: Program impact on filing probability - Interactions

	(1) No interact.	(2) Female	(3) Kigali	(4) Primary	(5) Large	(6) Top Know
Training	0.12***	0.07	0.10**	0.10**	0.11**	0.12***
	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)
Group		-0.02	0.00	-0.17**	-0.11	0.19**
		(0.06)	(0.05)	(0.06)	(0.09)	(0.09)
Group*Attend		0.15*	0.02	0.17*	0.11	-0.02
		(0.08)	(0.07)	(0.10)	(0.12)	(0.09)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Knowledge Perceptions	Yes	Yes	Yes	Yes	Yes	Yes
Observations	980	980	980	980	980	980

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Interactions used are with dummies for female, Kigali, primary education and large size. Controls used are: age, female, large, CIT, business life in months, primary education, university education, Kigali, email use, books of accounts, previous business, previous training. Coaching group is dropped.

Table A17: Program impact on tax knowledge - Interactions

	(1) F Increase	(2) F Gain	(3) K Increase	(4) K Gain	(5) P Increase	(6) P Gain	(7) L Increase	(8) L Gain
Training	0.245***	0.529***	0.190***	0.602***	0.270***	0.574***	0.280***	0.587***
	(0.036)	(0.079)	(0.042)	(0.086)	(0.033)	(0.079)	(0.030)	(0.074)
Group	-0.066	-0.182**	-0.083*	0.02	-0.080	-0.256***	-0.075	0.062
	(0.054)	(0.083)	(0.047)	(0.064)	(0.061)	(0.085)	(0.118)	(0.153)
Group*Training	0.080	0.092	0.126**	-0.070	-0.004	-0.099	-0.122	-0.346**
	(0.063)	(0.144)	(0.058)	(0.122)	(0.075)	(0.124)	(0.118)	(0.153)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var.	0.714	0.571	0.714	0.571	0.714	0.571	0.714	0.571
R-sq.	0.208	0.360	0.211	0.360	0.206	0.360	0.208	0.361
Observations	820	815	820	815	820	815	820	815

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *Increase* is a dummy variable indicating whether the post-pre knowledge scores difference is positive. *Gain* is the ratio of post-pre scores difference over the pre-training score. Interactions used are with dummies for female, Kigali, primary education and large size. Controls used are: age, female, large, CIT, business life in months, primary education, university education, Kigali, email use, books of accounts, previous business, previous training. Coaching group is dropped.

Table A18: Program impact on tax outcomes: population of new taxpayers

	(1)	(2)	(3)
	Survey Sample	Survey Ref. Pop.	Population
Panel A: Probability to Declare			
T	0.121^{***}	0.140^{***}	0.291^{***}
	(0.040)	(0.024)	(0.012)
Admin covariates	Yes	Yes	Yes
Control group Y	0.333	0.337	0.379
Observations	980	2,543	15,009
Panel B: Probability to Positive file			
T	0.076^{***}	0.080^{***}	0.132^{***}
	(0.026)	(0.014)	(0.007)
Admin covariates	Yes	Yes	Yes
Control group Y	0.121	0.122	0.122
Observations	980	2543	15009
Panel C: Log Tax Declared			
T	0.383^{**}	0.480^{***}	0.994^{***}
	(0.179)	(0.103)	(0.068)
Admin covariates	Yes	Yes	Yes
Control group Y	0.929	1.051	1.051
Observations	980	2543	15009

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Coefficients are marginal effects from probit (panel A-B) and tobit (panel C) regression computed at the mean of covariates.

Table A19: Program impact in rural sample

	(1) Declare	(2) Declare	(3) Positive	(4) Positive	(5) Log Tax	(6) Log Tax
Training	0.42^{***}	0.44^{***}	0.37^{***}	0.33^{***}	2.11^{***}	1.74^{***}
	(0.02)	(0.02)	(0.02)	(0.02)	(0.15)	(0.15)
Controls	No	Yes	No	Yes	No	Yes
Tax Center FE	No	Yes	No	Yes	No	Yes
Control group Y	0.199	0.199	0.080	0.080	0.533	0.533
Observations	2180	2180	2180	2180	2180	2180

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Coefficients are marginal effects from probit (1-4) and tobit (5-6) regressions evaluated at mean. The control group are taxpayers who did not attend the training. *Declare* and *Positive* are dummy variables with value 1 if the taxpayers declared and positively filed, respectively. *Log Tax* is the log(tax+1) transformation of the raw tax amount variable. Covariates are a dummy for Corporate type of business and a continuous variable for life in month. Tax centers are: Huye, Muhanga, Nyagatare and Rusizi.

Table A20: Program impact in urban sample

	(1) Declare	(2) Declare	(3) Positive	(4) Positive	(5) Log Tax	(6) Log Tax
Training	0.32^{***}	0.31^{***}	0.05^{***}	0.07^{***}	0.44^{***}	0.48^{***}
	(0.02)	(0.02)	(0.02)	(0.01)	(0.16)	(0.15)
Controls	No	Yes	No	Yes	No	Yes
Tax Center FE	No	Yes	No	Yes	No	Yes
Control group Y	0.461	0.461	0.120	0.120	1.078	1.078
Observations	2441	2441	2441	2441	2441	2441

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Coefficients are marginal effects from probit (1-4) and tobit (5-6) regressions evaluated at mean. The control group are taxpayers who did not attend the training. *Declare* and *Positive* are dummy variables with value 1 if the taxpayers declared and positively filed, respectively. *Log Tax* is the log(tax+1) transformation of the raw tax amount variable. Covariates are a dummy for CIT type of business and a continuous variable for life in month. Tax centers are: Kicukiro, Gasabo and Nyarugenge (see section 2.4.1). Sampling weights are used.

Table A21: Program impact on tax knowledge - alternative outcomes

	(1) Difference	(2) Kling Index
Panel A: Naïve estimation		
Training	1.28***	0.32***
	(0.09)	(0.02)
Controls	Yes	Yes
Control group Y	0.55	-0.01
R-sq.	0.377	0.492
Observations	818	818
Panel B: Kernel Matching		
Training	1.26***	0.31***
	(0.09)	(0.03)
Common Support	816	816
Ps-R^2 matched	0.002	0.002
LR chi^2 matched	1.74	1.87

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *Difference* is the *ex-post* minus *ex-ante* difference in knowledge points. *Kling Index* is the standardized Kling index (Kling, 2007). Controls used are: age, female, large, business life in months, primary education, university education, Kigali, email use, books of accounts, previous business, previous training, tax time in days, respondent manager of taxes. In Panel B, standard errors are bootstrapped from 999 repetitions.

Table A22: Program impact on tax amounts

	(1)	(2)	(3)
	Raw	Wins.	IHS
	Panel A: Naïve estimation		
Training	4,888.42**	3,279.88***	0.96***
	(2,358.82)	(1,252.32)	(0.14)
Controls	Yes	Yes	Yes
Observations	971	971	971
	Panel B: Kernel Matching		
Training	-17,348.36	4,149.42	0.64**
	(22,893)	(4,899)	(0.32)
Common Support	827	827	827
Ps-R^2 matched	0.002	0.046	0.056
LR chi^2 matched	2.14	25.78	29.18
	Panel C: IV Strategy		
Training	-1,796.7	-7,123.31	0.29
	(8,517)	(21,123)	(1.06)
Controls	Yes	Yes	Yes
Observations	2378	2378	2378

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Coefficients are marginal effects from tobit regressions evaluated at mean. The control group are taxpayers who did not receive any educational input. *Raw* is the tax declared in RWF. *Wins.* is a winsorized transformation of the raw tax amount at the 99$^{\text{th}}$ percentile. *IHS* is the inverted hyperbolic sine transformation, defined as log(tax + $(\text{tax}^2+1)^{1/2}$). All regressions weighted by sampling weights.

Table A23: Program impact on prepayments and ETR

	(1) Prepayment	(2) ETR
Panel A: Naïve estimation		
Training	0.022	-0.012
	(0.110)	(0.011)
Controls	Yes	Yes
Control group Y	0.545	0.033
R-sq.	0.223	0.215
Observations	151	119
Panel B: Kernel Matching		
Training	0.037	-0.008
	(0.090)	(0.010)
Common Support	149	117
Ps-R^2 matched	0.004	0.036
LR chi^2 matched	2.76	21.43
Panel C: IV Strategy		
Training	-0.151	-0.008
	(0.151)	(0.007)
Controls	Yes	Yes
R^2	0.02	0.03
Observations	331	279

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *Prepayment* is a dummy variable with value 1 if the taxpayer made a prepayment *ETR* is the ratio of tax payable over business incomes. In Panel A, coefficients are marginal effects from probit (1) and OLS (2) regressions. Covariates and Knowledge and Perceptions used are those indicated in section 2.4.3. All regressions weighted by sampling weights. In Panel B, standard errors are bootstrapped from 999 repetitions. In Panel C, Training is instrumented by the invite call. Refer to Table A26 for first-stage results.

Table A24: Program impact on knowledge and perceptions - PSM

	(1) Increase	(2) Gain	(3) Index	(4) Complexity
Panel A: N-N Matching with Replacement				
Training	0.28***	0.62***	1.21***	-0.05
	(0.05)	(0.08)	(0.16)	(0.05)
Common Support	818	813	818	813
Ps-R^2 matched	0.032	0.031	0.032	0.031
LR chi^2 matched	34.98	34.54	34.98	34.06
Panel B: Radius Matching with Caliper 5%				
Training	0.25***	0.57***	1.25***	-0.10***
	(0.03)	(0.07)	(0.12)	(0.04)
Common Support	818	813	818	813
Ps-R^2 matched	0.032	0.002	0.002	0.002
LR chi^2 matched	1.76	1.91	1.76	1.91

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses, bootstrapped from 999 repetitions.*Increase* is a dummy variable for a *ex-post* increase in knowledge. *Gain* is the ratio of difference in scorer over the 0-10 pre-training score. *Index* is the 0-10 rescaled knowledge index. Controls used are: age, female, large, business life in months, primary education, university education, Kigali, email use, books of accounts, previous business, previous training, tax time in days, respondent manager of taxes. *Complexity* is a 0-1 index from two indicators of perceptions of complexity.

Table A25: Program impact on tax outcomes - PSM

	(1)	(2)	(3)
	Declare	Positive	Log Tax
Panel A: N-N Matching with Replacement			
Training	0.13^{***}	0.15^{*}	0.33
	(0.04)	(0.05)	(0.31)
Common Support	807	807	807
Ps-R^2 matched	0.022	0.099	0.105
LR chi^2 matched	26.07	54.84^{***}	55.19^{***}
Panel B: Radius Matching with Caliper 5%			
Training	0.08^{**}	0.12^{**}	0.29
	(0.03)	(0.04)	(0.38)
Common Support	807	807	807
Ps-R^2 matched	0.002	0.047	0.056
LR chi^2 matched	2.36	26.02	29.37

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses, bootstrapped from 999 repetitions. *Declare* and *Positive* are dummy variables with value 1 if the tax-payers declared and positively filed, respectively. *Log Tax* is the log(tax+1) transformation of the raw tax amount variable. Coaching group has been dropped.

Table A26: Program impact on tax outcomes - IV strategy

	(1) Training	(2) Declare	(3) Positive	(4) Log Tax
Panel A: Two-Stages Least Squares				
IV Surveyed	0.37***			
	(0.02)			
Training		0.15**	-0.01	0.55
		(0.05)	(0.07)	(0.68)
Covariates	Yes	Yes	Yes	Yes
Control group Y	0.19	0.38	0.14	1.23
R^2	0.19	0.22	0.03	0.01
F-stat		349.41	124.4	119.51
Observations	2378	2378	2378	2378
Panel B: Reduced Forms				
IV Surveyed		0.10***	-0.02	0.10
		(0.02)	(0.03)	(0.38)
Covariates	Yes	Yes	Yes	Yes
R^2	0.18	0.07	0.06	0.02
Observations	2378	2378	2378	2378

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *IV Surveyed* is the instrument, referring to a dummy for whether the taxpayer is surveyed. Coaching group has been dropped.

Table A27: Tests for exclusion restriction

	(1)	(2)	(3)
	Declare	Positive	Log Tax
[1em] Attended	0.13^{***}	0.11^{***}	0.43^{**}
	(0.03)	(0.04)	(0.21)
IV	0.01	0.03	-0.29
	(0.02)	(0.04)	(0.41)
Covariates	Yes	Yes	Yes
Control group Y	0.33	0.14	0.760
R^2	0.217	0.069	0.025
Observations	2378	2378	2378

Standard errors in parentheses.

Coaching group has been dropped.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A28: Placebo test on non-attendees by survey participation

	Non-surveyed		Surveyed		
	Mean	Obs	Mean	Obs	Difference
Declared	0.34	1332	0.33	390	0.00
Nil-filed	0.67	448	0.68	130	-0.01
Log Tax	2.91	425	2.86	121	0.05
Log Tax>0	9.82	126	10.50	33	-0.68
Tax RWF	109737.24	425	64824.10	121	44913.13
Log income	4.65	448	4.15	130	0.50
Observations	1722				

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

Table A29: Filing outcomes mean differences by attrition status

	Non-attrited		Attrited		
	Mean	Obs	Mean	Obs	Difference
Declared	0.39	820	0.46	125	-0.07
Nil-filed	0.61	323	0.64	58	-0.02
Log Tax	3.43	300	2.40	52	1.03
Tax RWF	44500.24	300	25814.23	52	18686.02
Log income	5.28	323	4.79	58	0.49
Income RWF	11053526.42	323	5179536.07	58	5873990.35
N	945				

Observations weighted by sampling weights.

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

A5 Figures

Figure A1: Tax to GDP ratios - World Bank data

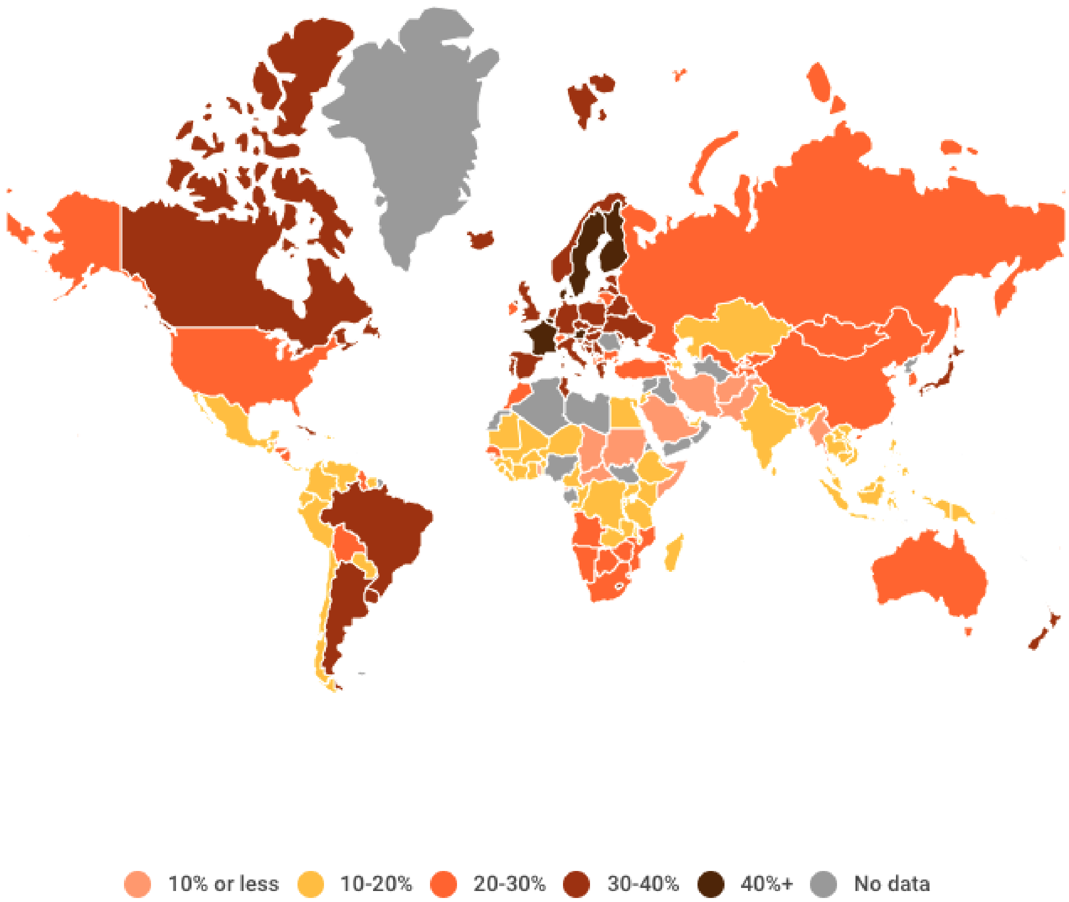


Figure A2: Tax revenues and Aid in Rwanda - World Bank data

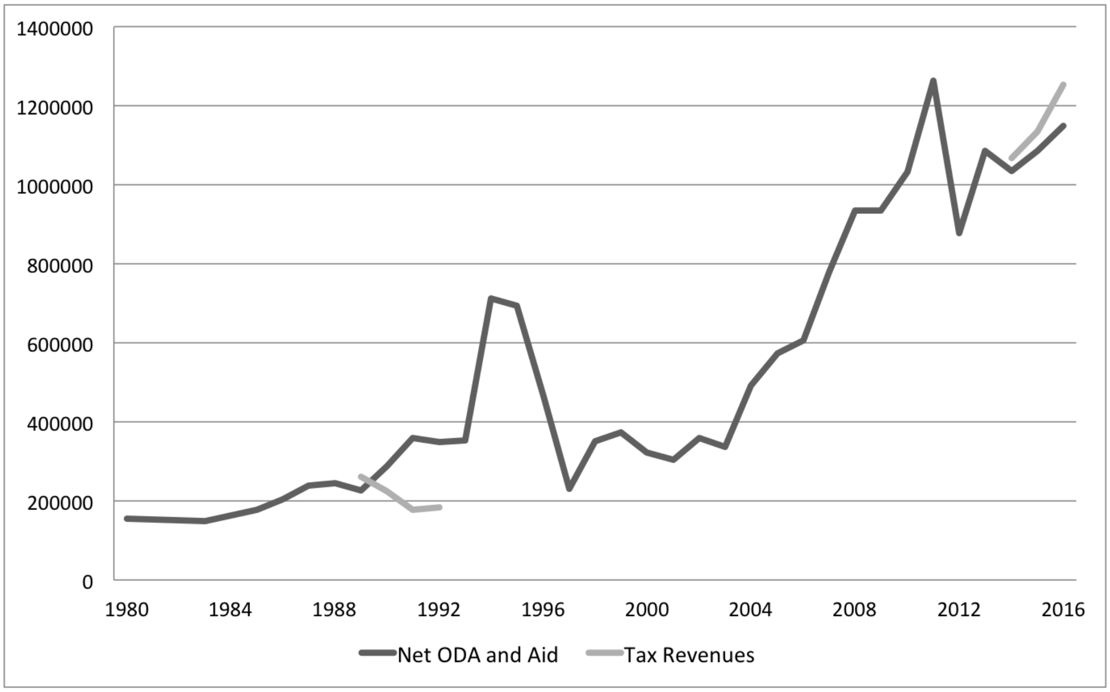
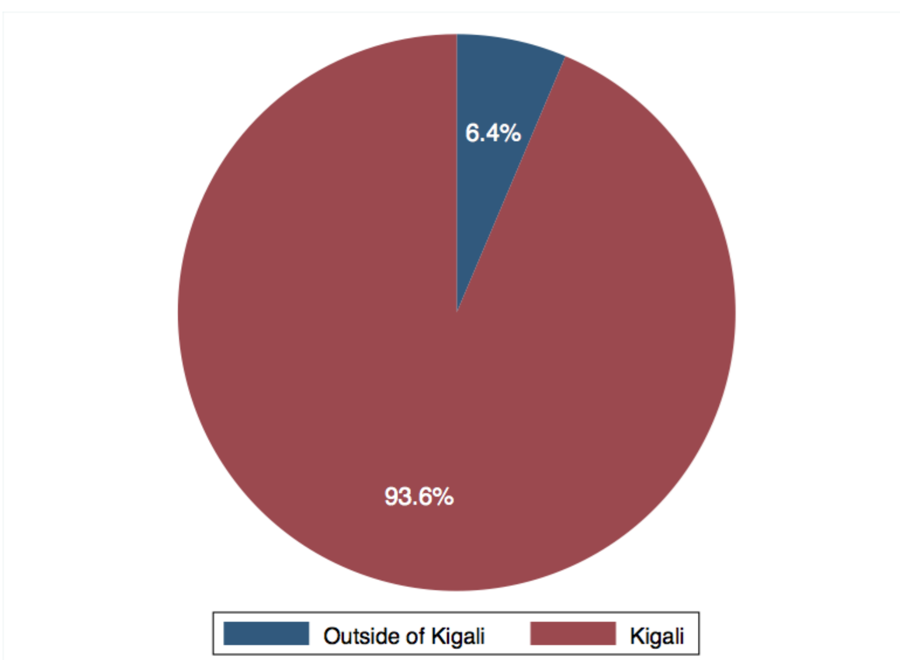
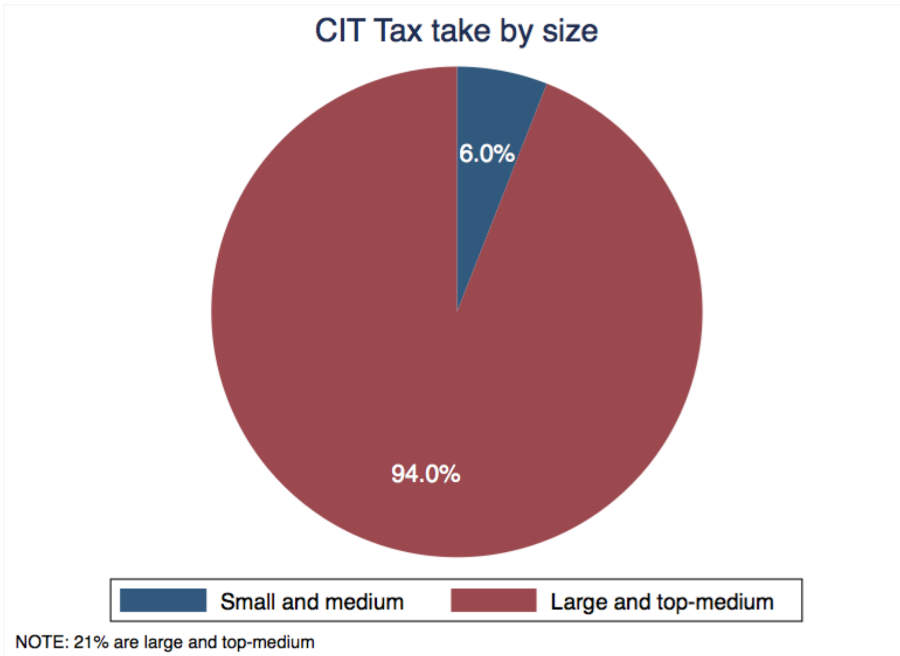


Figure A3: Tax Take by Location





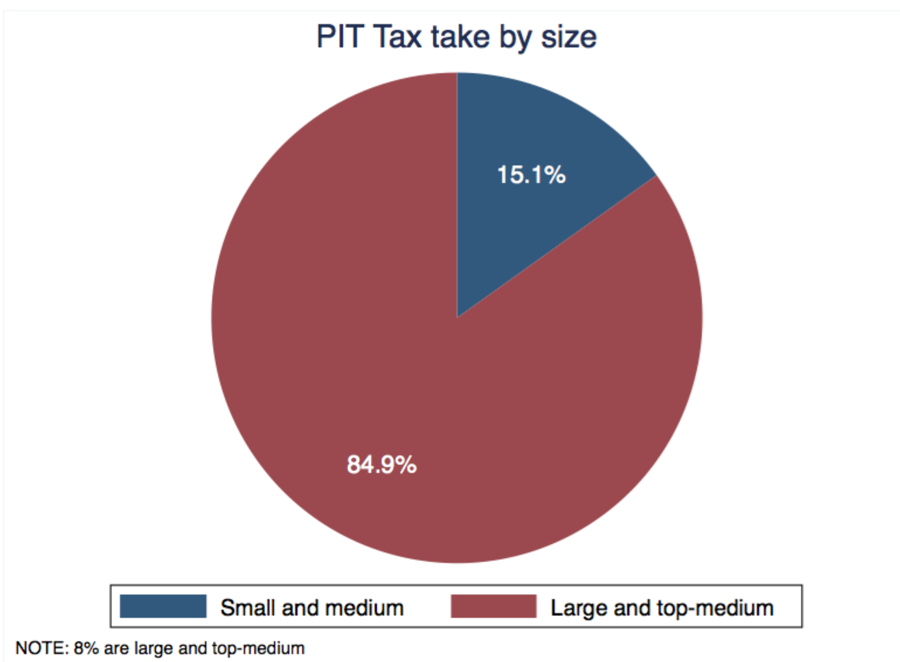


Figure A4: Tax Take by Taxpayer Size

Figure A5: New Registrations over Time

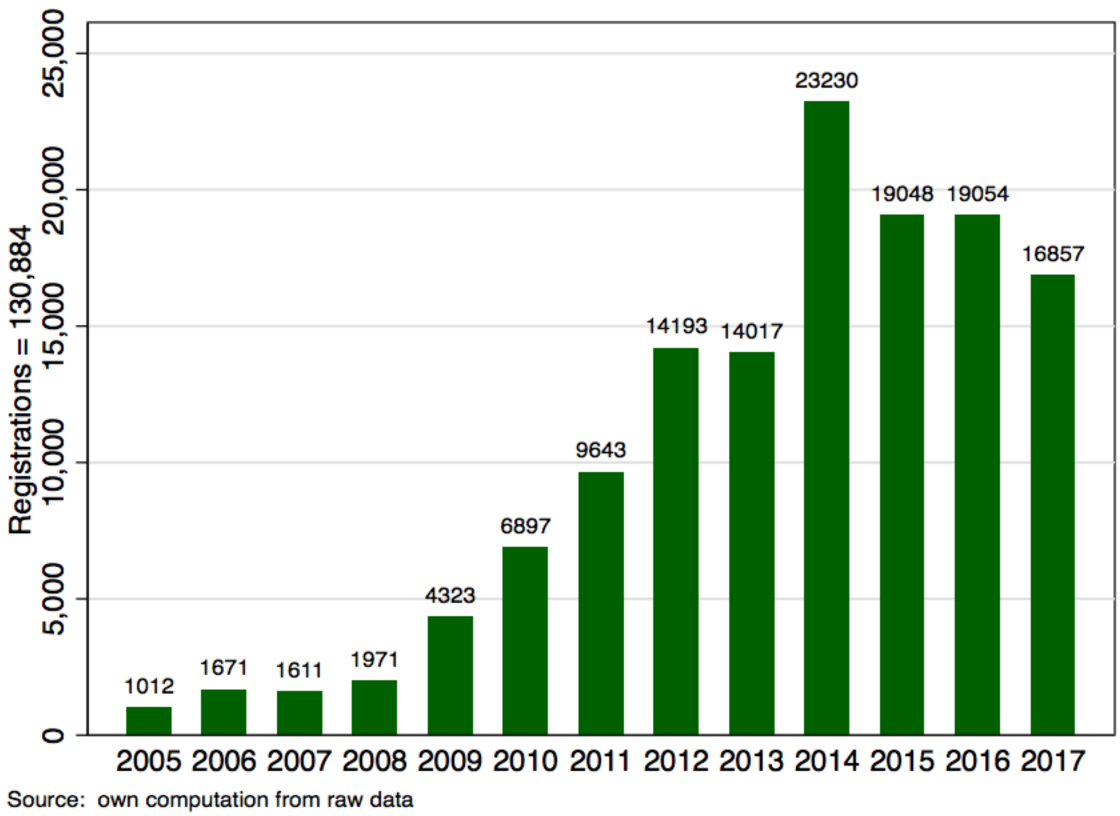


Figure A6: Propensity Score Distribution by Training groups

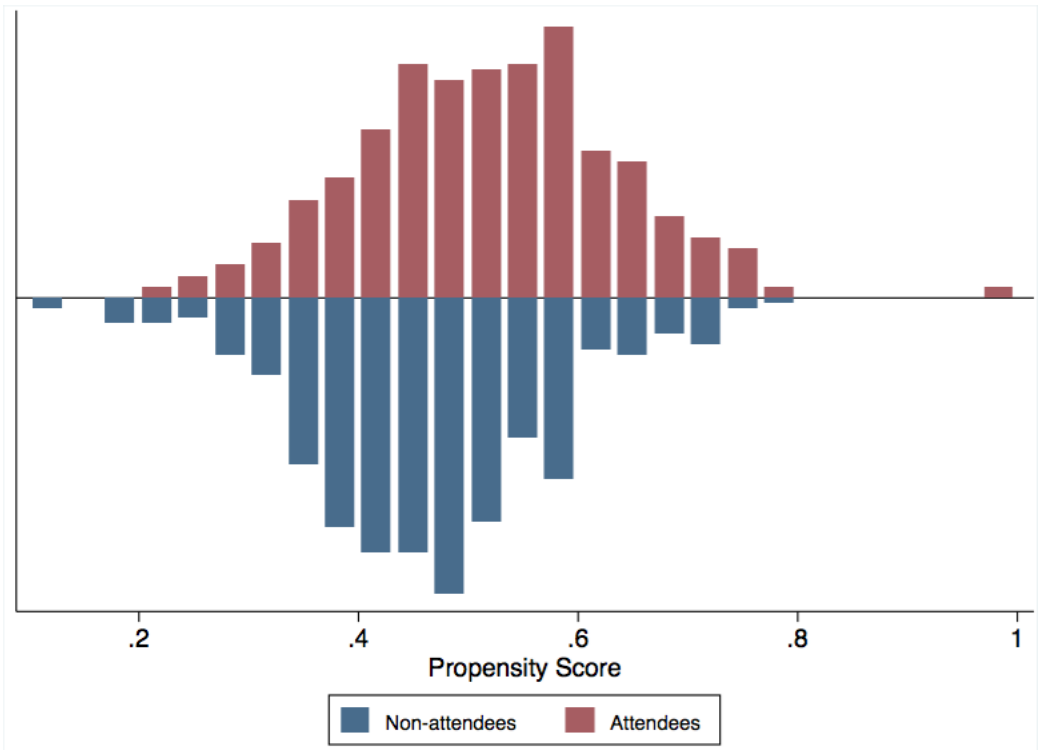
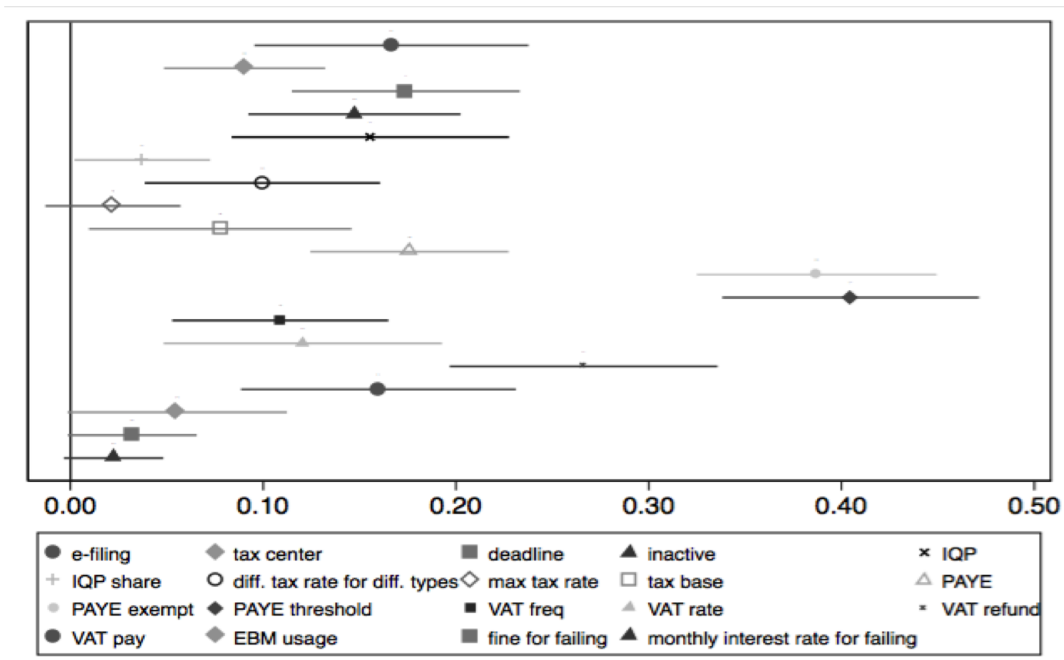


Figure A7: OLS estimates on single Knowledge items



Chapter 3

To File or Not To File? Another Dimension of Compliance - the Eswatini Taxpayers' Survey

Abstract
Non-filing refers to taxpayers who fail to submit a tax declaration, thus becoming ghosts in the eyes of tax authorities. It is a widespread phenomenon in sub-Saharan Africa, and has a number of detrimental fiscal effects. Non-filing has been largely unexplored in the literature, which focusses more on active filers. The overall aim of this paper is to shed light on the determinants of non-filing, building on neoclassical and behavioural theories, as well as to contribute to the methodological discussion on how to measure tax compliance. Focusing on Eswatini, the analysis combines survey data from a thousand entrepreneurs with their tax returns and filing history 2013-2018. I show that economic deterrence, compliance costs and moral factors, such as intrinsic motivation and peer pressure, are strongly correlated with actual filing. I also study how my key factors change when controlling for the persistence of filing behaviour in past years, or using a self-reported measure of compliance. I argue that tax knowledge plays a major role in understanding the decision to file. In terms of policy, results show that the tax authority could improve filing rates by adopting both a deterrent and an assistance-related approach, and also by triggering the role of social norms.

3.1 Introduction

Tax compliance can be considered as the sequential achievement of three main actions: filing a return, accurately reporting and paying the tax owed (Slemrod et al., 2001). In contrast with tax filers who may decide to under declare their income, non-filers choose the extreme compliance shortcut of not filing a return at all. This study looks at the first step of compliance, which is a necessary condition for the other two actions to take place. Non-filing has a number of important implications. First, especially in low-income countries, non-filing significantly erodes the tax base of already budget-constrained economies, with detrimental fiscal effects.[1] Second, a non-filer eventually becomes a *ghost* in the eyes of the tax agency, as he is missing from the tax records and fails to share valuable information with the authority. Third, non-filing creates economic inefficiencies and horizontal inequalities, since the effective tax rates faced by filers and non-filers of the same business size evidently differ. Fourth, non-filing goes against the law, generates unfairness, lowers the moral fibre of a society and ultimately delegitimises the government.

A growing descriptive evidence has been produced on the extent of non-filers in sub-Saharan Africa (SSA). In Rwanda, over three-quarters of individuals supposed to file a personal income tax (PIT) and half the companies liable to file a corporate income tax (CIT) failed to do so. Figures from Uganda are even higher, with the average rate of PIT non-filing being 86% over the period 2014-2018.[2] In Malawi, almost 50% of income taxpayers have filed no tax returns and/or made no tax payments over the period 2014-2016 (Ligomeka, 2019a). Moore (2020) notes that in 2016 the Nigerian federal revenue

[1] In the last decades, revenue authorities in sub-Saharan Africa (SSA) have made impressive progress increasing tax collections with respect to other low income countries (Moore et al., 2018). Yet, mobilised domestic revenue is clearly not sufficient to finance development (Bird and Gendron, 2007). According to the International Monetary Fund, on average SSA will need additional resources amounting to 19% of GDP to finance the Sustainable Development Goals (SDGs) in education, health, roads, electricity, and water by 2030 (IMF, 2019). The IMF target does not seem to be realistically achievable: on average, the tax-to-GDP ratio in SSA has risen by 2 to 3 percentage points of GDP in the past two decades and there are still 10 SSA countries with tax-to-GDP ratio still below 15%, most of them fragile states (Akitoby et al., 2019).

[2] Figures from Rwanda and Uganda have been computed by the author in parallel studies on tax compliance, drawing on detailed tax returns data.

authorities declared the following proportions of non-filing taxpayers: 98% for PIT; 94% for CIT; and 95% for VAT. Additional descriptive evidence from Kenya shows that only 3.5 million of the more than 9 million registered taxpayers filed their 2018 returns.[3] Eswatini, the country under study, is no exception: more than half (57%) of PIT returns are missing each year on average over the period 2013-2018; figures for CIT are lower (43%) but still alarming.

Despite the relevance of non-filers in SSA, non-filing is a neglected area of research with most of the tax literature being cast on tax filers and their reporting behaviour (Erard et al., 2018). Addressing tax non-filing requires some understanding of the factors underlying the taxpayer's decision whether to file a return or not. It is fair to assume that decisions related to file a return are substantially different from those based on how much to declare, conditional on filing. In this paper, I focus on non-filing of PIT, a progressive tax on income generated by non-incorporated traders, and seek to answer two interrelated questions: (i) Which economic and behavioural factors explain the decision to file a return in a given year?; and (ii) Do these factors differently impact the persistence of non-filing behaviour over time? A third crucial question naturally arises: Are the same factors explaining *self-reported* compliance? Note however that this study is of a descriptive nature. Due to the non-experimental set-up, my results point to strong correlations, and cannot be interpreted in a causal way.

To shed light on these issues, I combine a detailed taxpayer's perception survey of over a thousand taxpayers – the first data collection effort of this type ever carried out in the country – with rich administrative data provided by the Eswatini Revenue Authority. To the best of my knowledge, tax return data from Eswatini has not been studied yet, while most of tax research from Southern Africa comes from the recently expanding work on tax administrative data from South Africa (Ebrahim et al., 2019). The merging of survey and tax data is achieved through the use of uniquely identifying Taxpayer Identification Numbers (TINs). Thanks to the TINs I am able to link, for each taxpayer in the sample, tax attitudes and perceptions with their filing history in a quite novel way.

With the administrative data available I identify two main categories: active taxpayers

[3]See `https://www.businessdailyafrica.com/lifestyle/profiles/what-to-expect-file-nil-return/4258438-5232858-oiqmom/index.html`, accessed on June 10, 2020

who file their returns for the 2018 tax year, and non-filers who fail to do so. I am able to track the filing behavior of PIT payers and define *perpetual* (or *persistent*) active/non-filers as those taxpayers who consistently file/fail to file every year in the period 2013-2018. I consider it to be important to account for past filing behaviour, since it is commonly believed that once a filer enters the income tax system, he is likely to remain in the system (Erard et al., 2018).

In addition, survey data enables me to capture the key factors likely to correlate with compliance, as derived from the relevant literature on the behavioural drivers of tax compliance. These factors are organised in six groups and follow the recently published World Bank framework for understanding compliance (Prichard et al., 2019), which suggests three main paradigms or approaches to revenue authorities in low-income countries in order to encourage compliance - enforcement, facilitation and trust. The groupings are: (i) economic deterrence and pecuniary incentives (enforcement paradigm in Prichard et al. (2019)), (ii) compliance costs when filing a return (facilitation), (iii) trust in the authority and political legitimacy (trust), (iv) fiscal exchange, or the idea that taxpayers contribute to the public purse to get in return public services of adequate quantity and quality (trust), (v) social norms against or in favour of tax evasion (trust), and (vi) intrinsic motivation to comply (main outcome in the framework of Prichard et al. (2019)). Furthermore, I am able to test the relevance of a set of ancillary factors, such as risk aversion, tax knowledge, demographics and business-related characteristics. These factors are used to dig deeper into the main results and explore the possible mechanisms in place.

My findings suggest that some key factors are highly correlated with compliance, while others are not. More specifically, four out of the six theoretical motivations are able to discriminate between active and non-filers, and statistically significantly so. First, the perception of audit risk is positively related to active filing. Taxpayers who are above the median of the perceived audit risk distribution are 12% and 11% more likely to file last year and being perpetually active, respectively. Second, compliance costs are also important and account for a reduction of 16% in the probability of being active, and of 15% in the probability of being perpetually so. Third, social norms also seem to affect compliance: adhesion to a social norm seems to explain a fifth to a third of last year's and perpetual compliance, respectively. Lastly, having a high tax morale implies an increase of 21% and

12.5% of the probability to file last year or persistently doing so, respectively. In contrast with conventional wisdom, I find that neither trust nor reciprocity motivations covary with compliance. All results are robust to dimension reduction through principal component analysis and best subset selection with statistical learning methods, as well as the choice of the econometric model and the inclusion of context-specific fixed effects.

Also, as a second set of results, I compare the evidence above with the results from a regression in which the self-reported willingness to comply is the dependent variable – as in Prichard et al. (2019). This is to show how this self-reported measure, often used as a proxy for compliance in similar survey studies (more on this in section 3.2.2), is in reality driven by different factors. My results mean that, while compliance costs and fiscal exchange correlate with actual and self-reported compliance in the same way, other key factors, such as social norms and deterrence show different, if not opposite, patterns.

Lastly, I find that lack of tax knowledge, as a component of compliance costs, is strongly correlated with filing compliance. One extra question answered correctly in the tax quiz is associated with an increase of 14% and 9% in the probability of filing last year's tax return and being persistently active, respectively. Linked to that, background characteristics, such as employing a tax accountant and having a more mature business, are also crucial in understanding compliance.

This paper aims to contribute to three main strands of literature, which are discussed in detail in the next section. First, it tests whether the theory-based formulations (see section 3.2.1) on the drivers of tax compliance are practically relevant in Eswatini. In doing so, it produces a clearer understanding of the behavioural forces motivating filing behaviour in a country not studied before. Second, it contributes to the ongoing debate surrounding the measurement of tax compliance (see section 3.2.2). By merging survey and administrative data, this study adds an element of novelty to existing evidence, and promotes a third way of collecting data to measure tax evasion. Third, this study adds to the specific literature on non-filers. A common starting point is provided by Erard and Ho (2001), who extend the neoclassical model of tax evasion to account for the existence of non-filers. The authors then test their model on tax audit data of both filers and non-filers of US federal income tax returns. Consistent with what found in this paper, the main determinants of non-filing in the US are the probability of getting caught and the tax burden or compliance costs to file.

Several papers have studied the drivers of filing experimentally, through tax nudges, both in high- (Guyton et al., 2017; Meiselman, 2018; De Neve et al., 2019) and middle-income countries (Kettle et al., 2016; Brockmeyer et al., 2019). Apart from some considerations in Mascagni et al. (2019) in the context of Rwanda, no robust study on non-filing has been produced in the African continent.

These results are not only of academic interest but directly relevant to policy debates within tax authorities on the effective strategies to address poor tax compliance in low- and middle-income countries. Especially in SSA, the reality seems to be that tax policy, as written in books, is often very different from tax administration on the ground. While it is true that international advice is gearing towards simplification of tax systems in SSA, still very little is known about the practical considerations of African taxpayers. Therefore, it is of paramount importance for revenue authorities in SSA to extract knowledge on how taxpayers perceive the tax system, and how perceptions ultimately influence compliance, in order to implement more successful and realistic policies. In this way, informed, evidence-based, tax policies are more likely to be compatible both with what taxpayers really believe and the actual capacity and resources available within revenue authorities. This survey study attempts to stress that more systematic and robust collection of primary data on SSA taxpayers constitutes an important direction for research on tax issues in the region.

The rest of this paper proceeds as follows. Section 2 presents the theoretical framework for analysing tax compliance and reviews the methods to measure it. Section 3 describes the country under study. The methodological approach is addressed in Section 4. Section 5 presents the results, the robustness of which is dealt with in Section 6. Section 7 summarises and concludes the paper.

3.2 Literature Review

3.2.1 Understanding Tax Compliance: Theoretical Foundations

The theoretical framework through which tax compliance should be conceived and explained is composed by a full-house of motivations, largely referring to two main branches:

neoclassical and behavioral. These theories are (non-exhaustively) summarised here below:[4] they are highly interconnected and, in most cases, more recent formulations build on pre-existing ones. Testable hypotheses are formulated for each theoretical category, and are summarised in Table 3.1.

Deterrence As already discussed in section 2.2 of Chapter 2, economic deterrence stems from the neoclassical standard theory of utility maximisation, as developed by Allingham and Sandmo (1972). According to this theory, the key drivers of taxpayers' decisions are two pecuniary factors which increase the cost of evade: the size of the penalty and the probability of getting caught. Despite having been criticised of limiting the attention on monetary factors only and not considering the consensual side of compliance (Andreoni et al., 1998; Sandmo, 2005), economic deterrence results to be effective in motivating compliance, both in developed and developing countries.[5] A large number of theoretical deviations from the neoclassical model have been proposed (see section 2.2 of Chapter 2). The key elements of this model (e.g. audits and sanctions) are a cornerstone of tax administration across the world. However, overtly deterrent strategies, often blindly implemented in low-income countries, may be unsustainable and expensive over the long term. The enforcement paradigm requires a certain level of credibility in the authority's action, which is not necessarily guaranteed in low-income countries (McKerchar and Evans, 2009; Prichard et al., 2019). In a context of limited budget as in most African countries, including Eswatini, the key question is then how can compliance be encouraged without recurring to extensive and costly enforcement strategies, an aspect that will be further tested in

[4]The tax compliance literature is extensive and we do not attempt to review it all in detail here. For more extensive reviews, see Alm (2012); Mascagni (2018); Slemrod (2019).

[5]Exogenously manipulating penalty rates is challenging in real-world settings. Therefore, most evidence on the impact of fines comes from laboratory experiments Friedland et al. (1978); Beck et al. (1991); Alm et al. (1992). Relatedly, there is now growing evidence that non-financial penalties (e.g. public disclosure of one's tax returns) may also act as a deterrent (Bosco and Mittone, 1997; Fortin et al., 2007; Lefebvre et al., 2015; Perez-Truglia and Troiano, 2016). Evidence from developing countries reinforces the assumption that fines matter in some cases (Ortega and Scartascini (2016a) in Venezuela; Shimeles et al. (2017) in Ethiopia, Brockmeyer et al. (2019) in Costa Rica, Bergolo et al. (2019) in Uruguay; McCulloch et al. (2020) in Nigeria; Fjeldstad et al. (2020) in Tanzania) and not in others (Del Carpio (2014) in Peru; Carrillo et al. (2017) in Ecuador; Mascagni et al. (2017) in Rwanda).

For the purpose of this study, it is assumed that both a higher perception of the risk of being caught if evading and having been audited or fined in the past are related to better compliance outcomes.[6]

Compliance costs Complying with the tax law can be extremely challenging. It involves a high degree of technical knowledge on which rules and regulations apply, as well as on how to practically fill a tax return. The more complex a tax system is, the more difficult will be for taxpayers to comply (Richardson, 2006). Hiring a tax accountant often helps,[7] but the help of more sophisticated agents can be too expensive for small taxpayers or not enough in too complex tax systems (Chetty and Saez, 2013). Therefore, compliance costs are believed to be highly regressive, especially in low-income countries (Coolidge and Ilic, 2009). A growing descriptive evidence reports very low levels of tax knowledge in Africa: using Afrobarometer data on 36 African countries, Isbell (2017) reports that the majority of respondents have difficulty figuring out what taxes they owe to the government, and the same confusion is reported when studying a previous round of Afrobarometer data (Aiko and Logan, 2014). Fjeldstad et al. (2012) show that taxpayers in South Africa and Tanzania who perceive it as difficult to find out what taxes they are supposed to pay are less likely to have a tax compliant attitude.

Despite the increasing interest in tax policy – most African revenue authorities have started implementing strategies to improve tax knowledge (Mascagni and Santoro, 2018) –, tax knowledge is usually considered only as a background variable by tax researchers, proxied by the taxpayer's level of education. Only recently tax awareness has been measured in surveys, even if inadequately. In most cases, tax knowledge and overall compliance costs are operationalised with single survey items, such as the ease of finding out what taxes and

[6] A key departure from the neoclassical model is that I focus on the individual perceptions of audit rates, rather than actual audit rates. The neoclassical model assumes that taxpayers know with what probability they will be caught by the tax authority. On the contrary, I expect that taxpayers largely misperceive actual audit rates and substantially overweight a (actually low) probability of audit, as already theoretically formulated (Webley, 1991; Kirchler et al., 2007).

[7] It is also true that, although the availability of tax practitioners reduces informational and computational barriers to tax compliance, their use can also contribute to increase non-compliance by helping taxpayers take advantage of legal ambiguity Klepper and Nagin (1989); Erard and Feinstein (1994).

fees should be paid to the government (Fjeldstad et al., 2012; McCulloch et al., 2020). The investigation discussed in Chapter 2 was the first of its type in measuring tax knowledge in a more structured way, through a tax quiz (consistently replicated in this study, see section 3.4.1), and in capturing the causal link between tax knowledge and actual compliance. In the same fashion, I expect that taxpayers with higher tax knowledge and better tax-related practices, such as having a tax accountant, keeping books of accounts and spending more time on tax matters, are also more likely to actively file a tax return.

Fiscal exchange As already mentioned in section 2.2.2 of Chapter 2, the fiscal exchange theory builds on the concept of reciprocity between citizens and the State. In this setting, tax compliance is encouraged if the State is perceived to be using taxpayer's money in a transparent and just way, providing public services in sufficient quantity and quality (Cowell and Gordon, 1988; Falkinger, 1988; Levi, 1989; Moore, 2013). If taxpayers do not see the private benefits from contributing, they are more likely to engage in evasion and adjust their terms of trade with the government. The fiscal exchange theory has been abundantly developed theoretically and recently tested in low-income countries (Fjeldstad and Semboja, 2001; D'Arcy, 2011; Bodea and Lebas, 2016; Blimpo et al., 2018). Empirical evidence to support the theory is, however, ambiguous.[8]

With the survey data available, I am able to test the hypothesis that taxpayers who are more satisfied with the quality of public services and feel they are getting something in return from paying taxes are more likely to comply.

Trust and political legitimacy Closely related to the fiscal exchange theory, the quality and role of the institutions matters for compliance (Levi, 1989). More specifically, individuals' attitudes towards these institutions are crucial in deciding whether to comply or not (Torgler, 2003). Individuals who have a negative perception of the quality of

[8] Fjeldstad (2004) finds not evidence in favour of the fiscal exchange theory in his analysis of survey data in South Africa. Likewise, D'Arcy (2011) provides limited support for fiscal exchange using cross-country Afrobarometer data. Investigating a different wave of the same Afrobarometer data, Sacks (2012) finds that citizens who are satisfied with their government's provision of services and goods are more likely to defer to the tax authority. In a similar fashion, the nudging exercise of Mascagni et al. (2017) proved to be effective in raising revenues in Rwanda by stressing the link between tax compliance and better public services.

the government and the level of fairness in the tax system tend to comply less, both in the laboratory (Webley, 1991) and in the real world (Pommerehne and Weck-Hannemann, 1996). At the same time, if citizens judge the government as administratively competent and transparent in its actions, enforcement tends to be more effective in curbing evasion, because the State is considered more credible in its fight against it. In the slippery-slope framework, for example, trust and enforcement coexist and both contribute to create an environment more conducive to consenting compliance (Kirchler et al., 2008). In this setting, the combination of fairness, equity, reciprocity and accountability produces trust in the government and foster quasi-voluntary compliance, as recently summarised in Prichard et al. (2019). The aspect of taxpayers' trust is of paramount importance in the context of SSA, where citizens tend to trust more informal institutions such as religious and traditional leaders than tax officials (Bratton and Gyimah-Boadi, 2016) and perceive corruption in tax agencies as rampant (Isbell, 2017). Levels of trust in SSA are incredibly low when compared to, say, Nordic countries (Pirttilä, 2017).

In this study, I measure the extent of trust towards the tax authority, as well as perceptions on corruption, fairness and (lack of) transparency. As shown in section 3.3.2, Eswatini scores fairly poorly in terms of the quality of governance. Therefore, I hypothesise that taxpayers with worse perceptions on corruption, fairness and transparency are more likely to be non-filers.

Peer pressure It is also documented that one's behaviour in society is likely to be shaped by their peers' behaviour and social norms (Elster, 1989). It derives that tax evasion can become the norm within a society if it is socially approved by its members, thus becoming an informal and self-reinforcing rule of behaviour. Theoretical elaborations have stated that the cost of evading is an increasing function of the proportion of taxpayers who comply (Myles and Naylor, 1996; Kim, 2003; Fortin et al., 2007). Therefore, it is relevant to consider a given taxpayer's decision as not happening in a vacuum, but rather as something taking place in reference to the group in which the taxpayer lives in. Peer pressure and social comparison have also been observed empirically, with interesting results.[9]

[9] In Peru, disclosing information on the level of compliance in the subjects' reference group had a large positive impact on compliance (Del Carpio, 2014). In Guatemala, Kettle et al. (2016) shows that nudging taxpayers with a social norms message successfully impacted compliance with profit tax. The message

Against this framework, I test the hypothesis according to which a taxpayer is less likely to file the more he feels that is accepted to imitate his peers engaging in non-filing.

Individual morality In contrast to the high levels of evasion predicted by the neoclassical model, it has been established that some people never evade, even when the evasion gamble is better than fair (Baldry, 1986). These taxpayers can be categorised as honest or intrinsically motivated, since they always believe that evading taxes is the wrong thing to do. For example, Dwenger et al. (2015) show that compliance is observed even in setting, such as the local church tax in Bavaria, where tax enforcement is non-existent and private pecuniary benefits of compliance are likely to be minimal. Individual morality has been included in the taxpayer's decision problem either through a component of warm glow from contributing (Andreoni et al., 1998), or guilt or shame for failing to comply (Gordon, 1989; Erard and Feinstein, 1994; Reckers et al., 1994; Traxler, 2010). The notion of morality in compliance is also loosely captured by the term tax morale (Luttmer and Singhal, 2014). Importantly, the extrinsic incentives of a too harsh deterrence approach risk to crowd-out one's intrinsic motivation to comply (Frey and Torgler, 2007). A plethora of empirical studies have attempted to influence tax compliance through appeals to morality. However, most of these studies have failed to find any significant results.[10]

This research captures the individual morality of taxpayers in Eswatini and tests whether it is linked to compliance behaviour. A positive relationship is expected between the two.

Risk aversion It is well established that risk attitudes are one of the key explanatory factors of a wide range of individual economic behaviours. Relevant decisions concerning business investment, technology adoption ad growth (Banerjee and Newman, 1994), taking

referred to the (rather low share of) 64.5 percent of taxpayers that had already paid this tax and invited non-compliers to join the status quo. In a more developed context such as the UK, Hallsworth et al. (2017) significantly enhanced compliance through sending a letter informs the taxpayer that "nine out of ten people in the United Kingdom remit their tax on time. You are currently in the very small minority of people who have not paid us yet."

[10] Among the many: Blumenthal et al. (2001) in the USA; Torgler (2004b) in Switzerland; Fellner et al. (2013) in Austria; Pomeranz (2015) and Bergolo et al. (2019) in Latin America. Few exceptions can be found in Bott et al. (2014) for Norway and Mascagni et al. (2017) for Rwanda.

a loan (Banerjee et al., 2015), insurance (Friedman, 1974; Karlan et al., 2014; Cole et al., 2017), human capital investment and career decisions (Weiss, 1972) as well as migration (Bryan et al., 2014) are highly affected by subjective risk preferences. On top of that, risk aversion declines with income and is significant among the poor (Andrisani, 1978; Cicchetti and Dubin, 1994; Shaw, 1996).

A number of empirical studies have shown the role played by risk aversion in many different fields, especially so in developing countries.[11] When it comes to tax compliance, different theoretical considerations have been made concerning risk attitudes (Segal, 1990; Bernasconi, 1998), in the attempt of deviating from the model of Allingham and Sadmo (1972), which assumes that all taxpayers have the same level of risk preference. However, very little, if anything, is known about the relation between risk preferences and tax compliance in the field, even less so in developing countries. In some instances, risk aversion is questionably proxied by the taxpayer's perceived probability of being caught evading (Yücedoğru and Hasseldine, 2016).

Within the risk aversion framework, this paper contributes in two ways. First, it measures risk aversion more accurately, through an experimental technique widely adopted in the literature: the Multiple Price List (Andersen et al., 2006). Second, it attempts to link the experimental measure of risk aversion with actual compliance, deriving the proposition that more risk-loving taxpayers are more likely to engage themselves in riskier activities, thus less likely to abide to the law and file their returns.

Demographics Lastly, this study also focuses on the relation between compliance and a number of demographic variables. A long-standing, mostly US-based, tradition of tax research suggests that younger, single and self-employed individuals tend to comply less (Clotfelter, 1983; Erard and Feinstein, 1994; Erard and Ho, 2001). Younger and male individuals also appear to contribute less in laboratory experiments (Baldry, 1986; Alm and McKee, 2003; Kastlunger et al., 2010). At the same time, some growing evidence

[11] Pioneering in this field, (Binswanger, 1980) conducted experiments eliciting measures of risk aversion from farmers in rural India. More recent examples include Attanasio et al. (2012) in Colombia, Voors et al. (2012) in Burundi, Gilligan et al. (2014) in Nepal, Maertens et al. (2014) and Ward and Singh (2015) in India, Jakiela and Ozier (2016) in Kenya. See Cardenas and Carpenter (2008) for a recent review of the literature in developing countries.

from low-income countries shows that tax systems in Africa can be severely biased against women, who in turn may have worse feelings about taxes (van den Boogaard et al., 2018; Siebert and Mbise, 2018; Ligomeka, 2019b; Akpan and Sempere, 2019). In much the same way as tax knowledge, also the level of education is positively correlated with compliance (Spicer and Lundstedt, 1976; Song and Yarbrough, 1978; Kinsey and Gramsick, 1993). These results indicate clearly that demographic characteristics play a non-negligible role in explaining compliance (Hofmann et al. (2017) for a meta-analysis). Also, they are important in the way they interact with the perceptions discussed above, as abundantly documented in the literature (Richardson, 2006; Cyan et al., 2016).

3.2.2 Measuring Tax Compliance: Empirical Approaches

The literature so far has produced different definitions of tax compliance, which can be considered a spectrum of often hard-to-measure actions, especially when it comes to the grey area between tax evasion and avoidance (Slemrod, 2007). For the sake of this study, a relevant dichotomy arises when considering the extensive (failure to file) or the intensive (income understatement) margin of evasion. Following the categorisation in Halla (2010), I summarise here below the methods of measurement of (any definition of) tax compliance as direct or indirect. I also discuss a third approach, which is believed to be more robust and is the one adopted in this paper.

Direct approaches Direct approaches of measuring (non-)compliance are manifold. A first example is provided by administrative data, such as data on audits. Assuming that the agency is capable of unveiling all hidden income through an audit, such an approach would directly capture the extent of evasion. The most reliable source of data from tax audits is given by the US Taxpayer Compliance Measurement Program (TCMP). Importantly, the TCMP implemented *random* in-depth audits from 1963 to 1988. Despite the robustness of this method (Advani et al., 2019), it is inapplicable to contexts of limited investment in fiscal capacity such as in low-income countries. Apart from data on audits, tax returns data is used more and more as a direct approach of measuring evasion. The main advantages and disadvantages of this approach are addressed at the end of this section.

A second direct approach consists of measuring individual-level tax compliance in a lab-

Table 3.1: Theoretical Background

Explanatory category	Hypotheses tested
Deterrence	1. Taxpayers who perceive a higher probability of getting caught, are more likely to file 2. Taxpayers who have been audited or fined, are more likely to file 3. Taxpayers who have had more interactions with the authority, are more likely to file
Compliance costs	1. Taxpayers with more tax knowledge, are more likely to file 2. Taxpayers who perceive it as easier to file, are more likely to file 3. Taxpayers with a tax accountant/bookkeeping/more time on tax, are more likely to file
Risk aversion	1. Taxpayers who self-report to be more risk averse, are more likely to file 2. Taxpayers with a higher CRRA measure, are more likely to file
Fiscal exchange	1. Taxpayers who are more satisfied of the quality of public services, are more likely to file 2. Taxpayers who think they are getting something in return, are more likely to file 3. Taxpayers who think taxes can be raised to fund better healthcare, are more likely to file
Trust and political legitimacy	1. Taxpayers who think bribing is less common, are more likely to file 2. Taxpayers who think the tax system is fair, are more likely to file 3. Taxpayers who think the tax system is transparent, are more likely to file
Social norms	1. Taxpayers who would not imitate their peers' evasion decision, are more likely to file
Intrinsic motivation	1. Taxpayers who believe that evading is always wrong, are more likely to file
Demographics	1. Female taxpayers are more likely to file 2. Older taxpayers are more likely to file 3. More educated taxpayers can be more or less likely to file 4. Swazi-national taxpayers can be more or less likely to file

oratory, with early applications dating back to Friedland et al. (1978); Spicer and Thomas (1982); Alm et al. (1992). Pre- and post- survey data are usually collected to enrich the analysis (Bosco and Mittone, 1997; Torgler et al., 2010). Lab experiments have been criticised for their lack of external validity (Levitt and List, 2007). The debate is ongoing and results supporting the comparability of lab and real subjects have recently been produced (Alm et al., 2015).

The third, and most widely adopted, direct approach refers to capturing compliance through survey techniques. With this method, researchers directly ask the respondent whether they fail to comply with taxes. In a more preferable scenario, they find reasonable approximations of non-compliance through less direct questions. Tax surveys have gained relevance in low-income countries (see Fjeldstad et al. (2012) for a SSA-based review), also given the challenges in following the two other methods described above. Relevant examples of tax surveys are grouped in two categories: (i) cross-country international business[12] or citizen-level[13] surveys, and (ii) ad-hoc surveys implemented by researchers in a single country.[14]

Despite being expensive, surveys still represent the most powerful tool to capture relevant information, such as tax attitudes and perceptions, which cannot be extracted otherwise. Further, survey data allows for in-depth descriptive analysis, which often sheds light on new behavioural patterns and provides the basis for more experimental studies. Lastly, policymakers are interested in understanding the views of citizens and embedding survey evidence in policy decisions and future strategies.

At the same time, tax surveys present weaknesses and inconsistencies. The first point of criticism states that it is difficult to get honest answers about dishonest behaviour

[12]The most comprehensive of these is the Doing Business (DB) survey conducted by the International Finance Corporation of the World Bank. DB surveys are run every year world-wide and in most African countries. DB produces world rankings on the ease of doing business and a number of different sub-areas. Importantly, a specific module of DB focuses on the ease of paying taxes. Another example is given by the World Bank Enterprise Surveys, firm-level surveys of a representative sample of a country's private sector.

[13]The main examples are provided by Afrobarometer and World Values Surveys.

[14]Notable examples from low and middle income countries are provided by Gauthier and Reinikka (2001) in Uganda; Fjeldstad and Semboja (2001); Fjeldstad et al. (2020) in Tanzania; Fjeldstad (2004); Coolidge and Ilic (2009) in South Africa; Bodea and Lebas (2016); McCulloch et al. (2020) in Nigeria.

when respondents are motivated to present themselves in a positive light (Ajzen, 1991). Andreoni et al. (1998) suggest that taxpayers might overstate their degree of compliance in self-reports and those who have evaded might want to excuse their behavior by declaring a higher tax attitude. Relatedly, response rates can diverge by income groups and undermine the sample representativeness: it is more difficult to survey wealthy people and detect their levels of evasion (Alvaredo and Atkinson, 2010; Higgins and Lustig, 2013; Pirttilä and Tarp, 2019).

The second main critique refers to the operationalisation of the key dependent variable – tax compliance. It is true that asking for the willingness to pay taxes (often labeled as tax morale, as in section 3.2.1) is less blunt than enquiring about an illicit behaviour, and researchers follow this strategy to get higher degrees of honesty. At the same time, scholars often claim to be measuring tax compliance when they are just capturing an attitude. Also, it is unclear whether all respondents perceive the concept of *compliance* in an unequivocal way and this can undermine the internal validity of the survey instruments. The relationship between attitudes towards compliance and actual behaviour has been abundantly questioned in the literature, as reviewed by Onu and Oats (2016). For example, Elffers et al. (1987) find that there are significant differences between actual tax evasion, as derived from tax audits of 700 Dutch taxpayers, and survey responses. Likewise, Hessing et al. (1988) find no correlation at all between self-reports and documented compliance status with the Dutch tax authorities.

In addition, even if pretending that attitudes are consistent with behaviour, the way in which tax compliance is usually defined in surveys is not necessarily specific to the behaviour under study. Many examples can be provided in this regard. D'Arcy (2011) uses as dependent variable answers to the Afrobarometer question: "For each of the following statements, please tell me whether you disagree or agree: The tax department always has the right to make people pay taxes." This does not necessarily mean that a taxpayer is compliant, rather whether she believes that the State has the authority to collect taxes. Using another rounds of Afrobarometer data, both Levi et al. (2009) and Sacks (2012) adopt the same dependent variable to study the willingness to comply. Blimpo et al. (2018) create an index of tax morale to proxy tax compliance, in which the same question on government authority is included, together with one on trust in tax officials. In contrast, McCulloch et al. (2020) prefer to use the question "Which of the following options is closest to what you think about people not paying taxes on income?", where the options are: not wrong at all; wrong but understandable; and wrong and punishable. In conclusion, some confusion exists in the operationalisation of survey items and greater consensus is needed in order to improve the reliability and comparability of empirical tax research (Fjeldstad et al., 2012).

Indirect approaches Indirect approaches aim to provide macro-level estimates on tax evasion by inference from key observable indicators, such as currency demand or national income and product accounts. These observable indicators are what Slemrod (2019) defines *traces-of-income*. If performed correctly, indirect approaches can provide approximation of tax evasion cross-country and for a reasonably long period of time. The pioneering work of Pissarides and Weber (1989) uses food consumption as a proxy for income, and ends up inferring that self-employed individuals understate their income more than employees. Other examples of indicators are given by hoarding of high-value currency (Feige, 1990), the ratio of currency to money (Tanzi, 1980) and electricity consumption.

As explained in Slemrod and Yitzhaki (2002), indirect approaches are questionable in their methodology. For what concerns currency demand, it is usually difficult to estimate it accurately. When comparing national accounts measures of income and income reported to the tax authority, it is often the case that national income estimates of several key forms of income are based on tax return data itself. In addition, there are often inconsistent definitions of income for tax purposes and for national accounts.

The third way: surveys and administrative data The third solution, which aims at overcoming some of the weaknesses mentioned above, consists of merging survey and tax return data. While tax surveys have been implemented for decades, tax authorities of low-income countries have only recently inaugurated a collaboration with researchers in which a wealth of administrative data is shared and analysed (Mascagni et al., 2016). This collaboration has been fuelled by the impressive evolution of IT within revenue authorities, which produces a massive amount of tax data every day. Gathering and understanding such data has become a priority for making informed tax policy decisions.

There is a lot that can be learned from administrative data (Mascagni et al., 2016).

First, it captures actual filing behaviour, as opposed to biased survey self-reports. Second, the availability of tax returns across many years offers the opportunity to study trends in compliance over time, and to have a more comprehensive view of compliance patterns. Third, collaboration with international researchers builds technical capacity within the tax agencies themselves with the ultimate goal of improving internal processes.

Based on tax data, rigorous experiments and impact studies from SSA and the developing world have recently been published.[15] It is also fair to stress that administrative data comes with its own drawbacks. First – linking back to the introduction to this section and Chapter 1 –, income data from tax returns only captures the information that taxpayers decide to disclose to the revenue authority. All income derived from informal activity is therefore excluded. Second, administrative data, despite being anonymised before being shared, is highly confidential and often accessed by a small group of academics only, so reducing possibility of replication. Third, administrative data is available only for those who are registered in the first place, thus does not cover the informal sector. Survey data can address this concern by framing the sample to include non-registered taxpayers. Lastly, as Slemrod (2019) points out, results that provides an unfavourable picture of the way in which a tax authority operates are more likely to encounter resistance from senior management and eventually not be published.

This study represents one of the few examples of tax research which combines alternative data sources.[16] The merging takes place based on the taxpayer identification number (TIN), a unique identifier assigned to each taxpayer at the time of registration.

[15] See Mascagni (2018) for a comprehensive review. Relevant studies from Africa include Eissa and Zeitlin (2014); Mascagni and Mengistu (2016); Mascagni et al. (2017); Almunia et al. (2017); Mascagni et al. (2019); Santoro and Mdluli (2019); Mascagni et al. (2020). Field experiments from other developing context include: VAT payments in Chile (Pomeranz, 2015), individual municipal taxes in Argentina (Castro and Scartascini, 2013), firm taxes in Ecuador (Carrillo et al., 2017) and corporate income tax in Uruguay (Bergolo et al., 2019).

[16] Other studies include Mascagni et al. (2019) for Rwanda; Del Carpio (2014) in Peru; Bergolo et al. (2019) in Uruguay. When considering high-income countries, it is worth mentioning Lefebvre et al. (2015) in France, Belgium and the Netherlands; Fellner et al. (2013) in Austria; De Neve et al. (2019) in Belgium.

3.3 Institutional context

3.3.1 Country overview

The Kingdom of Eswatini[17] is a landlocked country in Southern Africa, bordered by Mozambique to the north-east and South Africa to the north, west and south. Eswatini is classified as a lower-middle income country with a GDP per capita in 2017 of $3,243 in PPP, according to the WB Development Indicators. Its main local trading partner is South Africa, and the countries' currency, the Lilangeni (SZL), is pegged to the South African Rand.[18] Economic growth is estimated to have slightly risen to 2.4 per cent in 2018 from 2 per cent in 2017, although growing fiscal challenges resulted in a projected growth rate of just 1.3 per cent for 2019 (World Bank, 2018). However, the country faces major development challenges. Based on the international poverty line of $1.90 a day, and the lower-middle income poverty line of $3.20 a day, it is estimated that 38 per cent of the Swazi population live in extreme poverty, and a total of 60.4 per cent are poor overall. This is accompanied by an unemployment rate of 23 per cent 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 per cent, with a total population of 1.2 million in 2018 (World Bank, 2018).

3.3.2 Tax system in Eswatini

The Eswatini Revenue Authority (SRA) is a semi-autonomous institution established by the Revenue Authority Act in 2008, as part of the government's reform strategy for revenue administration. SRA officially took over the function of revenue collection on 1 January 2011. The SRA collects both direct taxes, representing about 57 per cent of tax revenues

[17] Formerly known as Swaziland. The name change took place in April 2018. While in most places the paper reflects this change, several documents and reports issued prior to this change still make reference to Swaziland. The revenue authority is called Eswatini Revenue Authority but SRA is still its acronym.

[18] The country is highly dependent on South Africa, which provides around 85% of its imports and a market for about 60% of exports.

in 2017/18, and indirect taxes, amounting to 43 per cent of revenues (SRA, 2018). The main direct income taxes are taxes on companies (16% of total revenues) and taxes on individuals (36%), which is labelled here as personal income tax (PIT). The main indirect taxes are VAT (30%) and fuel taxes (12%). Concerning the focus of this study, PIT is a tax on income generated by individuals, and has a progressive structure – a maximum marginal rate of 33 per cent and exemptions for income below SZL41,000 (US$2,848). Three main categories of individuals are targeted by PIT: non-business employees taxed at source (PAYE), directors of companies and sole traders, with the latter being the focus of this study. From the analysis of PIT returns in 2012-2017, the relevance of the three categories in terms of number of returns lodged is as follows: PAYE (41%), sole traders (37%) and director of companies (21.5%).

In terms of filing obligations and deadlines, income tax returns must be submitted according to a staggered timeline. Non-VAT-registered small and medium enterprises are expected to furnish their CIT returns by 31 October each year, PIT payers have to file by 30 November, and large companies and CIT/PIT entities who are registered for VAT must submit their returns by 31 December. The tax year ends on 30 June. Importantly for this study, the law mandates that every registered taxpayer is required to file his return regardless of whether he is operative during the year. Strict sanctions are imposed by law for non-filing and for false assessment. Anyone who fails to furnish a return within the stipulated period may be liable on conviction to a fine of SZL10,000 (US$719) and/or imprisonment for a period of up to one year. Those making false assessments with an intention to evade are liable to a fine of SZL50,000 (US$3,591) or imprisonment up to five years. These amounts are discouraging, given that the average monthly turnover (total sales) of the taxpayers in the sample is of about SZL32,500. However, lack of human resources means that audit probability is likely to be low for small taxpayers and high for the most profitable cases. According to (ATAF, 2017), in 2017 auditors accounted for 6.5% of total tax administration staff, well below the SSA average of 12% and the 30% international benchmark (Gallagher, 2004).

3.3.3 Tax performance and tax compliance

Revenue collection has continued to show a steady increase year on year since the inception of the SRA. A growth of 8 per cent was recorded in 2017/18, compared to an average of 13 per cent over the past five years, as indicated in Appendix Figure A1 (SRA, 2018).[19] In terms of tax-to-GDP ratio, the country registered a positive trend from 12.3 per cent in 2011/12 to 14.7 per cent in 2017/18, but this is still far from OECD's 25 per cent. As shown in section 3.3.2, Eswatini – like Rwanda – collects proportionally more income tax than other African countries.

Table 1.1 from the introductory chapter reports key fiscal and governance indicators for Eswatini and, at the same time, reproduces the indicators for Rwanda in order to allow for cross-country comparisons. According to ATAF (2017), the 2015 tax-to-GDP ratio in Eswatini is about half that of Southern Africa and lower than Rwanda's. Eswatini scores worse in terms of governance outcomes, both when considering the Corruption Perception Index and the World Bank Governance indicators. For the latter, Eswatini underperforms in terms of voice and accountability, political stability, government effectiveness, regulatory quality and corruption – all institutional factors presumably linked to voluntary tax compliance. The quality of governance in Eswatini seems to be poorer both when compared to Rwanda, which may be seen as an outlier in Sub-Saharan Africa, but also, more tellingly, when benchmarking against the rest of Southern Africa. In much the same fashion, scores for government integrity and judicial effectiveness are worse than the regional average, while the tax burden is higher. The World Bank Doing Business indicators depicts a context more in line with Southern Africa, but still far from Rwanda: while Rwanda ranks 38^{th} in the world for the ease of doing business, Eswatini is 122^{nd}; when considering the ease of paying taxes, however, Eswatini performs slightly better, being 77^{th} in the world (Rwanda ranking 38^{th}). Less than a third of adults in labor force have a bank account, compared to 42% in the region.

SRA (2018) reports 53,208 registered taxpayers in 2017/2018. Taxpayers registered for income tax account for 83 per cent of the total. The positive trend in registrations

[19]Total revenue collection amounted to SZL8.453 billion ($617 million) in 2017/18. The reduced growth rate in 2017/2018 reflects the slow economic growth observed in the domestic economy in recent years and the fiscal challenges faced by the government, which is a major driver of the economy (SRA, 2018).

reflects the efforts of SRA to foster formality, as well as other service-oriented initiatives.[20] However, the informal sector still represents about 41 per cent of Eswatini national income, compared to 32% in Southern Africa (Table 1.1).

When it comes to revenues from income taxes, Appendix Figure A2 shows the trend of PIT and CIT collection over time. While CIT collection reported a 14 per cent below-target gap in 2017/2018, individual income tax performed fairly well, being 13 per cent above target. However, this performance was underpinned by higher PAYE collections mainly due to an increase in employee numbers in the Public administration and Manufacturing sectors (SRA, 2018). It is fair to assume that compliance of individual businessmen was not the key driver of this positive trend.[21] Hence, studying the drivers of individual businesses' compliance assumes an important value for tax policy as well.

Initial evidence on personal income tax compliance gaps can be gathered from SRA administrative data. I have access to the universe of 31,414 PIT payers registered up to December 2017 and the PIT returns for the period 2013-2018, lodged by a total of about 24,000 individuals. As explained in the introduction, I focus my attention on two main filing categories: active taxpayers and non-filers. The data shows:

- *Active taxpayers*: conditional on filing, the 6-year average of active (non-nil) PIT returns is 74 per cent.[22] Perpetually active taxpayers amount to 61 per cent of the filing population (or 14,637 units). Relevant to this study, sole traders as a subgroup of PIT payers are below the average rate of PIT filing at 70 per cent.

- *Non-filers*: the 6-year average of missing PIT returns is 57 per cent – 54 per cent for sole traders. This implies that more than half of all PIT payers (or 24,386) who were supposed to file a return in a given year failed to do so. In the last tax year, 72%

[20] A noteworthy example of this approach is given by *Operation Bakhumbute*. This was a door-to-door compliance campaign, which aimed to increase the taxpayer base and remind taxpayers of their tax obligations. The operation was carried out on 733 businesses in the Lubombo, Shiselweni and Manzini districts. About 20% of the businesses visited were found not to have registered with SRA for tax purposes. These businesses were educated on their compliance obligations, furnished with registration forms and advised on the registration process. Following initial engagement with these businesses, follow-up visits ensured that they actually registered.

[21] This view is also shared by SRA, as emerged from preliminary discussions with the senior management.

[22] The remaining returns are nil, lodged by the so-called nil-filers (Santoro and Mdluli, 2019).

of the taxpayers expected to file failed to do so. Considering the persistence of this behaviour over time, as many as 10,035 are persistent non-filers, meaning that they never filed a return after registration. In other words, about a third of all registered PIT payers never lodged a tax return.

In conclusion, taxpayers' compliance with personal income tax is far from optimal: every year, about half the returns are missing. Remarkably, non-filers outnumber active taxpayers every single tax year. This study attempts to explain why.

3.4 Research Design

3.4.1 Data

Administrative data Access to administrative data has been granted by SRA, with whom I signed a confidentiality agreement (see Chapter 1). More specifically, I have access to the taxpayers' registry, which contains information on the universe of taxpayers registered with the SRA for any tax type, and the PIT returns for the period 2013-2018, which provide information on the filing behaviour of the study population. Each taxpayers is assigned a taxpayer identification number (TIN), which is consistent across all SRA datasets and used to merge the registry and tax returns.

Administrative data serves two main purposes. First, it is needed to identify and locate the taxpayers to be targeted. Second, it assists in the unequivocal categorisation of taxpayers into the two main mutually-exclusive categories, active vs non-filers. The filing behaviour is classified by looking at the most recent tax year, 2018. This means that an active taxpayer positively filed the 2018 tax return, while a non-filer failed to do so in the same year. More specifically, the population of non-filers for a given year is a moving target, as non-filers are potential filers who have not yet filed. Therefore, the categorisation into non-filing depends on the specific time at which the data is observed, in this case the end of July 2019, or nine months after the filing deadline of 30 November 2018. Relatedly, I am able to observe the filing behaviour over a 6-year period and create the *perpetual* sub-category – taxpayers who keep filing in the same way every year.

As discussed in section 3.2.2, administrative data is more effective than survey data in capturing tax compliance accurately. In this case, the filing behaviour is measured exclusively from tax returns. It is also true that tax data has limitations (see section 2.4.1 of Chapter 2 and section 3.2.2), but in this case the quality of the SRA taxpayers' registry is higher than in other contexts, especially when it comes to identifying information such as phone numbers and location.

Survey data First-hand survey data has been collected by the author over a one-month period, with fieldwork starting on 7 November and ending on 8 December 2019. The survey has been programmed to be run on tablets through SurveyCTO software. The survey team consisted of ten enumerators and one team leader. Interviews were administered in person, with the enumerators first contacting the potential respondents on phone numbers extracted from SRA administrative records. The survey protocol was strictly followed, and taxpayers had to provide an informed consent before starting the interview and were free to quit at any time. The average duration of the questionnaire was about 40 minutes. Data collection and entry were accompanied by back-checks and other validation processes, consistently with academic best practice.

The content of the questionnaire was produced by the author with the support of SRA. The final questionnaire consisted of 9 modules, as summarised in Appendix Table A1 and described more in detail in section 4.3.2. After the pre-interview module 1 and the consent form in module 2, relevant background information was collected both at the taxpayer (module 3) and business level (module 4). Module 5 focused on attitudes towards risk, both through a self-reported measure of riskiness and an experimental measure for risk aversion. After that, module 6 collected other important information on key factors linked to compliance: tax knowledge (through a mini-quiz on tax of 5 questions) and compliance costs, enforcement likelihood, perceived corruption, moral attitudes towards compliance and perceptions on fairness, fiscal exchange and peers' behaviour. Module 7 explored satisfaction with public services while module 8 captured past interactions with the revenue authority.[23]

In some instances, the survey script replicated standard questions from well-established

[23]To reduce errors of recall, questions on business activity and interactions with SRA refers to the last 12 months only, i.e. from October 2018 up to the time of the survey.

international surveys, such as the Afrobarometer, the World Values Survey and the International Social Survey Programme.[24] At the same time, questions on tax knowledge, an aspect which is usually neglected in existing international surveys, have been mostly derived from the tax quiz used in the tax training study in Rwanda (section 2.4.1 of Chapter 2), where it proved to effectively capture (lack of) tax literacy. Overall the quality of the data is good: 95% of interviews are classified by the enumerator as having gone 'somewhat well' (26%) or 'very well' (69%).

3.4.2 Sample selection

The final sample consists of 1,009 PIT-registered taxpayers. The sample was supposed to equally represent active and non-filers, even if non-filers in 2018 amounted to the 70% of the population. In order to increase the power of within-category analysis, active have been overrepresented. The target of equal split has been successfully reached in the field: the final sample contains 513 active (51%) and 491 non-filers (49%). About 60% (613 taxpayers) are persistent in their behaviour: 76% of active (395) and 44% of non-filers (218) have been filing in the same way every time.

The sample has been randomly extracted from the taxpayer registry as at July 2019. Inclusion criteria include: (i) phone number is available so that the respondent could be contacted by the survey team and a meeting could be arranged,[25] (ii) to be registered anytime before January 2018, so to be liable to file a tax return for the tax year 2018 and therefore be categorised as active or non-filer,[26] (iii) to be located in Eswatini,[27] (iv) to be required to file for income tax,[28] and (v) the type of business. In relation to the latter point, all taxpayers in the sample fall in the category of *sole traders*, meaning that they

[24]For instance, questions on trust towards the authority and transparency in government spending are derived from the Afrobarometer series.

[25]Less than 3% do not have any phone number available and therefore have been excluded.

[26]About 4.8% of the population registered after December 2017 and have therefore dropped from the sampling.

[27]For less than 1% of the taxpayers in the registry, the location is not available or is from outside the country, mostly South Africa.

[28]Exempted entities are very few in Eswatini, 57 only. All of them are corporate taxpayers (mostly churches, NGOs, and the like).

are entrepreneurs running a business. Other categories, such as non-business employees, high-net worth individuals and directors of companies, even if liable to remit PIT, have been disregarded as it can be assumed that their tax compliance decision is affected by different motivations and constraints. Instead, sole traders are fully responsible for their own compliance behaviour and are the ones who decide whether to declare or not and, if yes, how much. Moreover, it is fair to believe that their own perceptions and attitudes towards taxation have an immediate effect on their compliance behaviour. Therefore, studying sole traders is more interesting both from a research and a policy perspective.

The sample contains both urban and rural taxpayers. All four districts in Eswatini have been covered and the sample is geographically representative – at least at the district level.[29] Appendix Figure A3 reports the location of each respondent, using different colours for the two compliance types. The main agglomeration of respondents refers to the main cities, Mbabane, Manzini and Lobamba, while more rural areas spread across the four corners of the country.

3.4.3 Estimation strategy

Main specification

Results are estimated through a linear probability model, according to the following OLS specification:[30]

$$Y_i = \alpha_i + \beta_i Z_i + X_i \Gamma + \epsilon_i \tag{3.1}$$

Where the outcome Y is the compliance behaviour of taxpayer i, i.e. a dummy for active

[29] The coverage of each district is as follows: (i) Hhohho 37%, (ii) Lubombo 16%, (iii) Manzini 38%, and (iv) Shiselweni 9%. These shares are very similar to those of the overall population of PIT payers: 38%, 15%, 39%, 8%, respectively. Therefore, no sampling weights will be used throughout the analysis.

[30] The linear probability model provides easier interpretations for the marginal effects on the probability of actively filing, compared to probit and logit. While the assumption of homoskedascity does not hold in a LPM, calculating *robust* standard errors controls for that (Angrist and Pischke, 2009). Moreover, LPM does not restrict predicted values within the 0-1 interval, but the share of such values is not high, ranging from a minimum of 0% to a maximum of 10% of the sample. The section on robustness show that, as a matter of fact, my results do not change (quantitatively or qualitatively) if I use a probit model.

filing status. As already stated in section 3.4.1, two compliance outcomes are considered: (i) whether the taxpayers actively filed in the most recent year, and (ii) whether he is a perpetually active or a perpetually non-filing taxpayer. With outcome (ii), I intend to run a robustness check in order to control for endogeneity issues such as reverse causality: it could well be that the fact of being active last year has affected the explanatory variables (even if the survey took place about 10 months after the most recent filing deadline) and, by focusing on the perpetual sample only, it is fair to assume that the outcome variable is constant over time. Additionally, I aim to compare the determinants of the extensive margin of compliance with those of self-reported compliance. The latter is built from answers to the question on whether tax evasion is justifiable (see section 3.4.3).

The vector Z_i refers to the set of key explanatory factors under study. These factors are grouped following the theoretical formulations on tax evasion (section 3.2.1): (i) deterrence, (ii) compliance costs, (iii) trust and political legitimacy, (iv) fiscal exchange and reciprocity, (v) social norms and (vi) intrinsic motivation. The control vector X_i includes both taxpayer-level and business-level characteristics. The operationalisation these factors is explained more in detail in section 3.4.3. For the sake of this study, the coefficients of interest are given by the β_i. Each explanatory factor Z_i will be used with and without controls.

Each factor is regressed either alone, in a bivariate regression setting, or together with all the other factors, in a multivariate regression setting. In this way, I control for the bias caused by the potential interactions between the right-hand-side (RHS) variables. The lack of significance that I could find for a factor, say trust, may be driven either by the fact that there is truly no relationship taking place, or by the fact that trust is also correlated and explained by a number of other RHS variables, such as reciprocity, accountability, fairness and social norms. This would be a case of bad controls (Angrist and Pischke, 2009). If bivariate and multivariate coefficients do not differ much, as happens in this case, it is a sign that such a bias is not undermining my results. Finally, the option of robust standard errors is used to control for heteroscedasticity.

In the same fashion, a probit specification will be run and marginal effects computed. As shown in section 3.5.3, results do not change when a probit model is used.

Independent variables

Deterrence Deterrence is captured in multiple ways. First, the perceived risk of being audited is measured. I have information on both an individual's likelihood of being audited and the likelihood of a peer, or a business like the respondent's one. Dummy variables indicating a perceived risk audit higher than the median are created to ease the interpretation of results.[31] Second, survey module 8 enquires about interactions with the SRA. Indicators such as distance from SRA, ever having been audited (and number of audits), ever having been fined (and number of fines), ever having interacted with the authority (number of interactions) will be used as alternative predictors.

Compliance costs As a measure of compliance costs, I adopt two survey items. First, perceptions on compliance costs are gathered through answers to questions on how difficult it is to file and how difficult it is to get in touch with SRA to get tax-related information. Second, in order to further probe the role of complexity perceptions, I consider tax knowledge as a specific proxy for compliance costs – as in Chapter 2. While perceptions of complexity are somehow subjective, the answers to a tax quiz can provide a more objective measure of tax ignorance and compliance costs. In order to capture the quality of tax knowledge, both a raw index and a standardised index (Kling et al., 2007) are created from the five-item quiz on tax. Additionally, background characteristics on taxpayers' practice, such as having a tax accountant and the time spent on tax in a month, are used as further indicators of compliance costs in the mechanisms section 3.6.

Fiscal exchange A specific survey module captures the respondents' satisfaction with the government's provision of six public services.[32] The first component from a principal component analysis is gathered as an overall satisfaction index. Further, (lack of) fiscal exchange is also captured by two other survey items: (i) disagreement with the fact that the government can decide to make people pay more taxes in order to increase spending

[31] The median, rather than the average, is usually chosen as threshold so to create two groups of similar sizes and control for skewed distributions. However, the data at hand are not extremely skewed and median and average are often very similar. In the case of audit likelihood, the median is 60% while the average is 62%.

[32] Primary schools, tertiary education, infrastructure, electricity, healthcare and security.

on public health care, and (ii) feeling of not getting anything in return from paying taxes.

Trust and political legitimacy I use a number of variables in order to capture political legitimacy. First, I measure trust in the revenue authority as a rank response for the extent of mistrust towards the SRA. Second, perceptions on corruption are captured by individuals' agreement with the fact that businessmen are sometimes required to make gifts or unofficial payments to get things done with regard to taxes. I also collect a more quantitative variable, as the share of total annual sales that businesses pay in informal payments or gifts to public officials for tax purposes. Third, I ask how fair the respondent feels the amount of income taxes she remits is. Fourth, to proxy for transparency in the governance, I measure how easy it is for the respondent to find out how the government uses the revenues from people's taxes.

Peer pressure I use perception of other people's tax compliance as proxy to measure the influence of other people's behavior on the respondent's tax compliance. Specifically, I use two measures: (i) a more quantitative one, asking for the perceived share of businesses in the respondent's area understating their income, and (ii) a more qualitative one, enquiring about the level of agreement with the statement: "If my neighbours do not pay taxes, it is fair for me not to pay them either."

Individual morality As a measure of the intrinsic motivation to comply, I capture the level of disagreement with the following statement: "It is right for some people not to pay the taxes they owe on their income." This variable is often used as a proxy for compliance. For this reason, I also use it as a dependent variable to test whether the factors impacting actual compliance differ when it comes to self-reports.

Risk aversion Survey module 5 captures risks aversion both through a self-reported attitude towards risky situation on a 1-10 scale and a more experimental measure.[33] The experimental exercise is also known as multiple price list (MPL), previously used by Holt and Laury (2002) and Harrison et al. (2007), among others. This measure has rarely been

[33]The 1-10 quantitative variable is transformed in a dummy for self-reported riskiness above median. Again, median (5) and average (5.1) are very similar in magnitude.

used in relation to tax compliance (see section 2.3) and is discussed more in detail in the Appendix section A1.1.

Demographics and business characteristics A large number of background characteristics are collected, which serve as controls: (i) demographics, such as gender, age, education and country of origin; (ii) business background information on being currently in operation, having run a previous business, location, sector, level of competition (both with formal and informal businesses), change in the size of the business in the last year, and total sales in a given period[34]; (iii) taxpayers' related practices on bookkeeping, having a bank account, using emails to communicate with clients and suppliers. While these factors are used as controls in the main specifications (section 4.4.3), they are also explicitly studied in section 3.6.

3.5 Results

3.5.1 Anatomy of Survey Sample

Response rates and attrition Implementing face-to-face interviews with small entrepreneurs, whose opportunity cost of giving up their time for a 40-minute survey is presumably high, is challenging. A large group of replacements is allocated to each enumerator in order to swiftly address non-responses. Importantly, the replacement order is randomised by the survey software. This means that taxpayers in the replacement group are comparable to those in the main sample. In many instances, enumerators had to replace hard-to-reach respondents. At the end of data collection, about a third (31%) of taxpayers in the main sample were successfully reached, while the remaining two-thirds (69%) were randomly picked from replacements. Appendix Table A3 shows that the group of taxpayers consenting to be surveyed is comparable to those who refused, except for minor deviations: consenting taxpayers show slightly fewer years of filing, look smaller in terms of log tax declared and are 4 percentage points less likely to be perpetual non-filers. On the other

[34]From Anderson et al. (2019), I first ask the respondent to choose a reference period: week, month or year. Then I enquire about the total sales in an typical period, meaning not the best and not the worst.

hand, the geographical distribution is nicely balanced. This evidence is supportive of the fact that heterogenous attrition is not a main threat to the analysis.

Summary statistics Before enquiring the regression tables, it is worth exploring the sample descriptively. Appendix Tables A4 to A7 report the summary statistics for the survey items observed in this study. As derived from Table A4, the average taxpayer in the sample is a married male (60%), aged between 41-50, a Swazi national, without higher education.[35] While the majority in the sample employ tax accountants (57%), keep at least some form of records (65%) and have a business-related bank account (58%), only a fifth of the sample use emails to communicate with clients or suppliers (20%).

When it comes to business-level characteristics, it is striking to realise that the vast majority of the sample reports to be in operation (74%). This is true for non-filers as well, 56% of whom state they have had at least one business transaction in the last year (compared to 90% of active). This may already hint to the fact that non-filers are indeed operative but not declaring to the fiscus as requested by the tax code (section 3.3). This is also suggested by data on reported monthly sales, with non-filers reporting an average of $1,028 USD ($1,180 USD excluding zeros). While this amount is significantly lower than actives' ($3,130 USD or $3,474 USD excluding zeros), the fact that non-filers are openly disclosing this information raises the question of whether they are just unaware of their filing responsibility. Appendix Figure A4 reports the distribution of reported sales by group, indicating the group averages with vertical lines. Also, the average business is about 6 years old, working in the wholesale/retail trading sector (56%) and, in most cases (80%), competing with informal businesses.[36]

Relatedly, a striking majority (74%) suffer from high competition, both from formal and informal activities, and half of the sample saw a reduction in business profits in the last year. The geographical distribution is the one already shown in Figure A3 and reflects the higher economic relevance of the two main business districts in the country, Hhohho and Manzini. Lastly, it is interesting to see which reasons push our traders to register with the authority and spot any difference between active and non-filers. Appendix Figure

[35] About two thirds of the sample have, at most, a high-school qualification.

[36] When fielding the questionnaire, enumerators made sure to explicitly refer to those businesses who are not registered with the SRA and therefore remitting no formal taxes.

A5 displays the results. While it seems that actives perceive a stronger sense of State legitimacy and feel more scared of breaking the law, non-filers seems to have registered out of a hope of growth opportunities, as well as to access government services and attract more clients.

For risk preferences (Table A5), the negligible levels of indifference and inconsistency represent a positive assessment of the quality of the lottery data. The sample is on average slightly risk averse, with a coefficient of relative risk aversion (CRRA) of 0.17.[37] About 40% of the sample still opts for the safest choice A in the last lottery round, in which option B is far more convenient (see Table A2). Furthermore, Table A6 reports the level of interactions with the authority. About 15% had been audited since registration, while a higher share received a fine (25%). At the same time, as many as 40% of the sample have had other types of interactions in the last year, with more than two interactions on average.

In the same fashion, Table A7 summarises the key explanatory factors of the study. It is worth noting that the perceived audit probability is high in general, and much higher when referring to other businesses (83%) than when referring to the taxpayer himself (62%), as also displayed in Figure A6. This is likely to suggest the existence of computational biases from bounded rationality, as framed in detail by the prospect theory (Kahneman and Tversky, 1979; Dhami and al Nowaihi, 2007). Also, compliance costs seem to be large, with the average number of correct answers in the tax quiz being 1.6 out of a maximum of 5, and as many as two thirds of the sample reporting difficulty to file. As another dimension of the compliance costs, for a sizeable 40% of the sample it is difficult to get in touch with the authority to get assistance.

At the same time, indicators of trust and political legitimacy depict a situation in which the majority (55%) think that businesses are bribing tax officials, about half the sample sees the tax system as unfair and 74% believe that the government processes are not transparent. This evidence somehow confirms the poor scoring at the international level from Table 1.1. This is also reflected in the relatively high level of mistrust towards the revenue authority, with an average score of 2.5 out of 4. Relatedly, about 58% of the sample disagree with the fact that taxes can be increased to finance health care. In much

[37]See Appendix section A1.1 for more information on the CRRA.

the same vein, half the sample believes to get nothing in return from contributing to the public purse. Social norms also seem to point towards a context in which tax evasion is present, with the perceived average frequency of businesses evading being 40%. However, when it comes to measuring peer pressure and neighbour effect, just 12% believes that it is fair to emulate a neighbour who is cheating on taxes.

Finally, despite this initial evidence on deteriorated tax attitudes and perceptions, the vast majority of the sample (82%) reports a surprisingly high intrinsic motivation, believing that tax evasion is not the right thing to do. This factor will also be used as a dependent variable when exploring how the behavioural determinants differ across actual and self-reported compliance.

3.5.2 Regression results

In this section, I report the results from the model discussed in section 4.4.3. In the regression tables below, the columns *All* show results for last year's filing behaviour (all taxpayers in the sample), while the columns *Perpetual* consider the persistent filing behaviour (the subset of perpetual taxpayers).

Which factors explain actual tax compliance? Table 3.2 below shows the results from my main specification. Columns 1 and 2 do not control for any background variables, while columns 3 and 4 control for both demographic and business level features. Standard errors are shown in parentheses. Also, Appendix Figure A7 plots the coefficients from columns 5-6 to ease comparisons across groups.

From the table, some statistically significant patterns emerge. First, perceptions on the general audit probability are negatively related to the probability to file, both with and without controls: being below the median of the perceived audit probability distribution is associated with a reduction in filing probability of 6-7 percentage points when all controls are added (col. 3-4), significantly so at the 5% level. This translates in a reduction of the probability of being active last year of 12% and of being perpetually active of 11%.

As a second set of results, compliance costs play a major role. Taxpayers who think that filing a tax return is somewhat or very difficult are 8-9 percentage points less likely to file. The coefficients are always significant at the 1% level and are meaningful in magnitude:

compliance costs account for a reduction of 16% in the probability of being active and of 15% in the probability of being perpetually so. Relatedly, the difficulty in getting assistance from the SRA is weakly associated with the failure to file for last year's return while it turns insignificant for the persistent behaviour.

Third, social norms seem to covary with compliance: while the perceived share of evaders in the community is not significant, the consideration of peers' behaviour seems to be positively correlated with filing, with coefficients highly statistically significant and doubling for persistent filers. Therefore, it seems that it is not so much the perceived incidence of evasion in the community that matters for compliance, but rather the moral adherence given to the existing norms surrounding compliance. The magnitude is also strikingly sizeable: adhesion to a social norm seems to represent a fifth (col. 3) to a third (col. 4) of last year's and perpetual compliance, respectively.

Finally, the intrinsic motivation to comply also explains a non-negligible share of the filing probability. Having a high tax morale implies an increase of 11 percentage points in filing probability, falling to 8 percentage points for persistent taxpayers.

Besides these significant results, the importance of factors related to trust and political legitimacy, as well as to fiscal exchange, is not confirmed in this exercise. Perceived corruption, lack of transparency and mistrust towards the agency do not play any role in explaining compliance. The only exception is with the perceived unfairness, which implies a noticeable reduction in filing probability (6-8 percentage points, or about 12-13%) for last year's compliance only (col. 1-3), but is never significant for perpetuals. More in general, perpetuals are arguably not moved by any consideration on trust and legitimacy when complying with the law. Also, fiscal exchange motivations perform poorly. Neither satisfaction with public services nor feelings of reciprocity show a significant coefficient. If anything, expecting nothing in return from paying taxes is positively correlated with being a persistently active taxpayers. The immediate consideration would be that perpetuals are so intrinsically motivated up to the point in which they: (i) are not affected by corruption, unfairness, lack of trust/transparency in the system, and (ii) do not expect to receive anything back from their contributions.

Table 3.2: Determinants of Active Filing Behaviour

	(1) All	(2) Perpetuals	(3) All	(4) Perpetuals
		Deterrence		
Risk audit below median	-0.11***	-0.14***	-0.06**	-0.07**
	(0.03)	(0.04)	(0.03)	(0.03)
		Compliance costs		
Difficult to file	-0.10***	-0.10**	-0.08***	-0.09***
	(0.03)	(0.04)	(0.03)	(0.03)
Difficult to get in touch	-0.07**	-0.06	-0.05*	-0.02
	(0.03)	(0.04)	(0.03)	(0.03)
		Trust and political legitimacy		
Bribing above median	0.02	-0.03	0.01	-0.05
	(0.05)	(0.07)	(0.05)	(0.05)
Unfairness	-0.08**	-0.06	-0.06**	-0.04
	(0.03)	(0.04)	(0.03)	(0.03)
No transparency	0.02	0.04	-0.01	0.02
	(0.04)	(0.05)	(0.03)	(0.04)
No trust above median	0.01	-0.00	0.01	-0.02
	(0.03)	(0.04)	(0.03)	(0.03)
		Fiscal exchange and reciprocity		
Services	-0.02	-0.02	-0.02	-0.02
	(0.02)	(0.02)	(0.01)	(0.02)
No fiscal exchange	0.00	-0.01	-0.02	-0.05
	(0.03)	(0.04)	(0.03)	(0.03)
Nothing in return	0.01	-0.00	0.04	0.06*
	(0.03)	(0.04)	(0.03)	(0.03)
		Social norms		
Evaders % above median	0.01	0.06	-0.01	0.04
	(0.05)	(0.06)	(0.05)	(0.05)
Peer pressure	0.11**	0.23***	0.10**	0.20***
	(0.05)	(0.06)	(0.04)	(0.04)
		Intrinsic motivation		
High tax morale	0.15***	0.12**	0.11***	0.08*
	(0.04)	(0.06)	(0.04)	(0.05)
Demographics	No	No	Yes	Yes
Business Char.	No	No	Yes	Yes
Mean of Y	0.513	0.644	0.513	0.644
R-sq.	0.057	0.077	0.326	0.486
Observations	1009	613	1009	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

What about the drivers of self-reported compliance? After the analysis on actual compliance, it is prudent to see if the key factors motivating it are also explaining self-reported compliance. In the first instance, it is worth mentioning that 82% of the sample never or almost never justify evasion. In most survey studies studying tax behaviour, this subsample is mistakenly considered as compliant. However, also non-filers, who are non-compliant in practice and may want to excuse their behaviour by declaring a high tax attitude (Andreoni et al., 1998), seem to never justify evasion: 77% of them think so, compared to 86% of filers.

Table 3.3 studies the correlations with the high tax morale dummy – the indicator used as a proxy for self-reported compliance (see section 4.4.3) – of the same factors explored in Table 3.2.[38] Column 1 reports LPM coefficients without controls, while column 2 adds taxpayer-level and business-level background characteristics. Interestingly, some of the explanatory factors from Table 3.2 remain statistically significant: compliance costs and peer pressure strongly influence the probability of having a higher tax morale. Those who believe that it is difficult to file are 7 percentage points less likely to have a high tax morale (col. 2), hinting at the fact that a complex tax system often frustrates taxpayers, discouraging them from complying.[39] In addition, difficulty in getting in touch with SRA contributes to hamper compliance. While communication issues only weakly correlate with actual filing (Table 3.2), probably due to other major constraints with compliance, they strongly covary with the (un-)willingness to contribute and add up to the negative relation expressed by the difficulty to file. At the same time, peer pressure is strongly negatively correlated with self-reported compliance: those who feel the pressure of their peers are 35 percentage points (or 43%) less likely to be compliant. It results that, while peer pressure is pushing taxpayers to file their return (see Table 3.2), it produces a totally opposite impact on intrinsic motivations. The reason could be that tax morale and peer pressure are substitutes and while they both explain actual compliance, they offset each other when it comes to self-reports.

Consistently, variables on trust still remain not significant, with the exception of the lack of trust which appears to have a weak negative relation. Relatedly, the absence of

[38] Given that there is no trend over time for tax morale, no *perpetual* category exists in this case.

[39] Similar results have been produced in developed economies, mostly in lab settings (Roberts et al., 1994; Eriksen and Fallan, 1996).

reciprocity mechanism positively moves with tax morale by 6 percentage points (col. 2), in much the same vein as the results for actual compliance (Table 3.2). This evidence suggests that the fiscal exchange theory does not find a confirmation in the Eswatini context. Maybe due to cultural or historical reasons, the reciprocal link between contributions and public services does not seem to hold in this setting.

As a last consideration, the deterrence indicator shows no association with tax morale, in line with the evidence on the crowding out effect of pecuniary incentives, such as penalties and fines, on intrinsic motivations (Frey and Feld, 2002).

Results remain consistent when considering as *self-reportedly compliant* only the 69% of the sample who has a very strong willingness to comply, as shown in Table A8.[40] Likewise, results do not change if I remove the share (77%) of non-filers who self-report positive attitudes towards compliance, since this mismatching could probably bias the direction of impacts – as shown in Table A9. If anything, the significant coefficients from Table 3.3 get even larger, while the non-significant factors remain so.

This exercise confirms that considering survey-based measures as a proxy for compliance can often be misleading. In the case of Eswatini, while compliance costs and reciprocity motives impact actual and reported compliance in the same way, other key factors such as peer pressure and perceptions on deterrence have different, if not opposite, effects on the two outcomes.

[40] In the words of Onu (2016): if a taxpayers feels very strongly that being fully compliant is the right thing to do, then it is likely that her attitudes will predict behaviour more than someone who feels equally favourable towards compliance, but does not have an equally strong attitude. To corroborate this line of reasoning, I focus on taxpayers with a very strong attitude. However, despite the evidence that the strength of attitudes is a valuable information to consider (Sparks et al., 1992), coefficients from the new regression remain highly consistent with those in Table 3.3. The only change is that the (lack of) fiscal exchange mechanism loses significance, while the perception of bribing turns to be significant at the 5% level and reduces the tax attitude by 10 percentage points. This could suggest a slight change in motivations when the strength of the self-reported attitude increases.

Table 3.3: Determinants of Self-reported Compliance

	(1)	(2)
Deterrence		
Risk audit below median	0.01	0.01
	(0.02)	(0.02)
Compliance costs		
Difficult to file	-0.07***	-0.07***
	(0.02)	(0.02)
Difficult to get in touch	-0.10***	-0.08***
	(0.03)	(0.02)
Trust and political legitimacy		
Bribing above median	-0.05	-0.03
	(0.03)	(0.03)
Unfairness	0.00	0.02
	(0.02)	(0.02)
No transparency	0.00	0.02
	(0.03)	(0.03)
No trust above median	-0.07***	-0.04*
	(0.02)	(0.02)
Fiscal exchange and reciprocity		
Services	0.00	0.01
	(0.01)	(0.01)
No fiscal exchange	0.07***	0.06***
	(0.02)	(0.02)
Nothing in return	-0.04	-0.04
	(0.02)	(0.03)
Social norms		
Evaders % above median	-0.06*	-0.05
	(0.03)	(0.03)
Peer pressure	-0.41***	-0.35***
	(0.05)	(0.05)
Demographics	No	Yes
Business Char.	No	Yes
Mean of Y	0.817	0.817
R-sq.	0.199	0.295
Observations	1009	1009

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

3.5.3 Robustness checks

Dimension reduction and best subset selection In an attempt to check for the robustness of the main results, Table 3.4 reports coefficients when all above factors are included in the same specifications, grouped by conceptual indexes through Principal Component Analysis (PCA). The first component is retained as it explains most of the variance in the model (see Appendix Table A10). In this model, demographics and business-related characteristics are explicitly presented, thus adding a new layer of information to what shown in Table 3.2. Also, the role of such background characteristics will be explored more in detail in section 3.6.3. From Table 3.4 some clear patterns emerge, which closely confirm those in Table 3.2. While demographics play a role (more on this in section 3.6.3), the high deterrence index is again strongly significant, together with the index representing less compliance costs. Consistently, both the strong social norm and high tax morale indexes are again significantly covarying with filing behaviour. In line with Table 3.2, fiscal exchange and trust indexes do not show any significant relation.

As an additional robustness exercise, I recur to statistical learning methods for selecting the best subset of predictors (James et al., 2013). The exercise consists in running a number of different subset selection methods on a randomly selected training set (half of the sample) and validating the results on a test set (the remaining half of the sample). Models with lower test mean squared errors (MSE) are preferred (Friedman et al., 2001). Once the best model is chosen, the most relevant predictors are retained and then applied in the original linear probability model (see section 4.4.3). Appendix Table A11 shows the result of this exercise as applied to the probability of being an active filer in the last return. Column 1 reports the original model from Table 3.2. Column 2 shows the results from running a linear probability model on the training set. Column 3 to 8 report the estimation from a number of statistical learning methods, in which only the relevant predictors are kept, while the others are dropped. These methods are backward stepwise selection (James et al., 2013) in col. 3 and lasso (Tibshirani, 1996) in col. 4 to 8. The different lasso models differ by the way in which the optimal penalisation term (lambda) is chosen.[41] All such methods are run

[41] In col. 4, lambda optimal is derived from cross-validation over the whole sample; the Akaike Information Criteria in col. 5 (Akaike, 1974), the AICc (Sugiura, 1978; Hurvich and Tsai, 1989) in col. 6; the Bayesian Information Criteria (Schwarz, 1978) in col. 7; and the EBIC (Chen and Chen, 2008) in col. 8.

on the same training set and validated over the same test set. The lowest value of the test MSE is reached with the cross-validation lasso, which therefore is the preferred method to select the best subset of predictors. Using this subset, the specification in column 9 shows the impact on the probability to file.

The main result of this exercise is that all the significant factors from the original model are retained: deterrence, compliance costs, unfairness, peer pressure and tax morale. Two out of three factors related to fiscal exchange are kept but they do not exert any significance impact. Other factors related to corruption, transparency and distrust are all dropped.

I repeat the same exercise for being a perpetual active (tables omitted for brevity). In this case, the same factors as in the previous exercise are selected, with the main difference that now unfairness loses its significance, in much the same vein of what seen in Table 3.2.

Alternatives to LPM As an additional robustness check, I re-run the main specification using alternative econometric models. Appendix Table A12 reports the coefficients from a probit regression. The specifications in each column have the same structure as those in Table 3.2. For the sake of better interpretation, I report marginal effects evaluated at the mean of the regressors. In this fashion, coefficients can be seen as percentage change in the outcome variable. As shown in the table, results remain consistent both in the level of significance and magnitude. Again, factors such as perceived risk, difficulty in filing, peer pressure and tax morale are significantly affecting the probability to comply. Remaining factors such as trust and reciprocity, on the other hand, show no significant impact.

The same specifications have been run using a logit model. Results remain consistent and the table is omitted for brevity.

Controlling for enumerators' ability and day of the week The data collection has been carried out by a team of ten enumerators. Despite being adequately trained, the survey team may differ in intrinsic motivations and skills. Heterogeneity in enumerators' performance can have an impact on the estimates discussed above, biasing them upwards or downwards. For this reason, I re-run the specifications from Table 3.2 by including enumerators' fixed effects, thus controlling for such heterogeneity. Results are shown in Appendix Table A13. When enumerators' effects are kept constant, coefficients are largely consistent, both in terms of significance and magnitude. This finding confirms that differ-

Table 3.4: Active vs Non-filers - Indexes

	(1)	(2)	(3)	(4)	(5)	(6)
	All	Perpetuals	All	Perpetuals	All	Perpetuals
Demographics	0.03*	0.05***			0.02	0.05**
	(0.02)	(0.02)			(0.02)	(0.02)
Business Char.	0.04***	0.04***			0.02*	0.02
	(0.01)	(0.02)			(0.01)	(0.02)
Profitable business	-0.02	-0.01			-0.03*	-0.02
	(0.02)	(0.02)			(0.02)	(0.02)
High deterrence			0.12***	0.10***	0.12***	0.10***
			(0.02)	(0.02)	(0.02)	(0.02)
Less compliance costs			0.07***	0.09***	0.06***	0.09***
			(0.02)	(0.02)	(0.02)	(0.02)
Fiscal Exchange			0.02	0.04*	0.02	0.04**
			(0.02)	(0.02)	(0.02)	(0.02)
Less Trust			0.01	-0.02	0.01	-0.02
			(0.01)	(0.02)	(0.01)	(0.02)
Strong social norm			0.04**	0.04**	0.03*	0.04*
			(0.02)	(0.02)	(0.02)	(0.02)
High tax morale			0.15***	0.08	0.15***	0.09*
			(0.04)	(0.05)	(0.04)	(0.05)
Mean of Y	0.513	0.644	0.513	0.644	0.513	0.644
R-sq.	0.016	0.025	0.091	0.097	0.098	0.111
Observations	1009	613	1009	613	1009	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The construction of PCA component is presented in Appendix Table A10.

ences in enumerators' ability are not likely to be driving the results.

A similar check has been run by controlling for the day of the week on which the survey took place. This exercise is run in order to address the fact that, as shown in Kahnemann et al. (2004), the context of the interview itself can cause a bias. Results do not change and are reported in Table A14.

Bivariate vs multivariate analysis As mentioned in section 4.4.3, the key explanatory factors are included both separately and jointly so to partially address concerns on bad controls (Angrist and Pischke, 2009). While the main results in Table 3.2 refer to a multivariate regression analysis, it is useful to consider also the stand-alone correlation of factors taken separately. Figures 3.1 and 3.2 compare the coefficients of regressors in bivariate and multivariate specifications, for both last year's filings and perpetuals, respectively. It is reassuring to see that coefficients from multivariate analysis do no differ much from those in bivariate regressions. If anything, the multivariate coefficients are slightly reduced in size with respect to their stand-alone counterparts. The only factor which exhibits a significant jump in magnitude is peer pressure. A future avenue of research could focus on why the relevance of social norm is enhanced when considered together with alternative behavioral factors.

External validity of survey results As a final robustness check, I test for the external validity of my survey instrument and results. I compare the results from this study with those I get from running the same analysis on PIT payers from Rwanda. In a parallel and ongoing study fielded in January 2020, me and ICTD co-authors collected survey data from a random and nationally representative sample of 1,172 sole traders in Rwanda, 63% of which are active - i.e. they filed a positive tax return in 2018 - and the remaining 37% non-filers. The survey tool has been kept as close as possible to the one in this study in order to enhance comparabiliy. Few exceptions refer to missing information on *Nothing in return* and *No transparency* variables.[42] All other key determinants are asked and operationalised in the same way as those in this study.

Table 3.5 reports the comparison of LPM estimates between Eswatini in col 1-2 - the

[42] The questionnaire in Rwanda has been carefully agreed upon after consultations with the Rwanda Revenue Authority. Hence, it is has been difficult to include every single question from the Eswatini survey.

Figure 3.1: Last year's filing - bivariate vs multivariate analysis

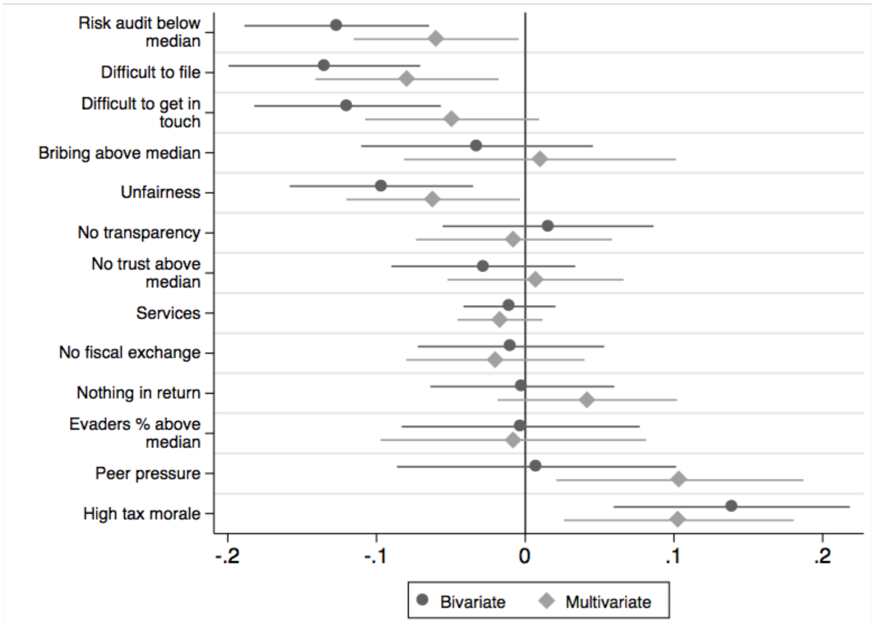
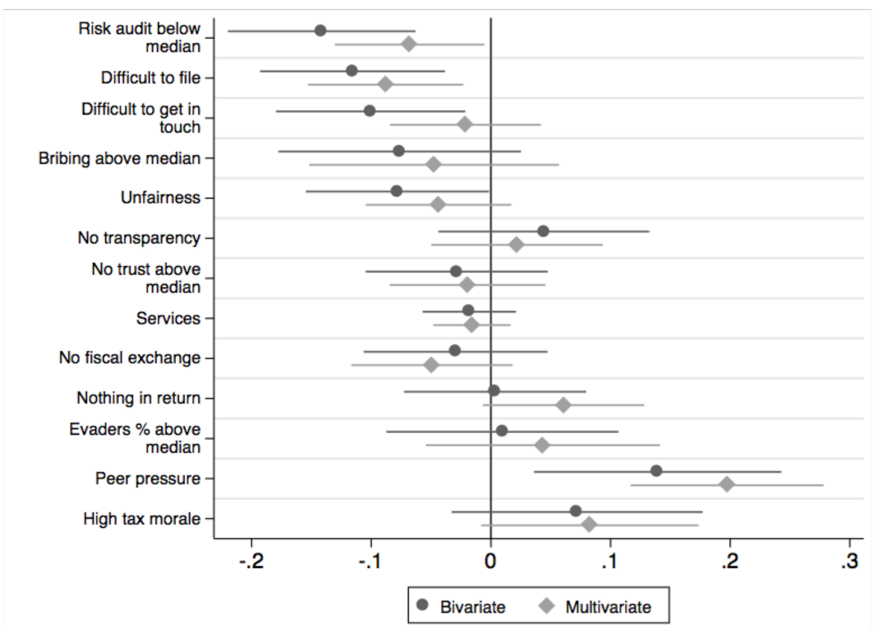


Figure 3.2: Persistent active filing - bivariate vs multivariate analysis



same estimates from Table 3.2 – and Rwanda (col. 3-4). Unfortunately, the analysis could not be run on perpetuals, given the lack of adequate information on taxpayers' filing history in Rwanda. In some cases, results are impressively similar, while in other cases small differences arise. First, deterrence-related factors are strongly significant in Rwanda as well and have the same positive effective on filing as in Eswatini. Second, compliance costs also play a role, but whether in Eswatini the difficulty in practically filing a return is more prominent, in Rwanda it is the difficulty to get in touch with the authority to ask for assistance what matters - as it does in Eswatini as well. Third, trust and reciprocity motives are negligible in the Rwandan context, while perception of unfairness was a key driver of filing in Eswatini.[43] Fourth, consistently with Eswatini, neighbour effects are present in Rwanda as well. Lastly, the intrinsic motivation to comply significantly correlates with filing in Rwanda, with a surprisingly similar coefficient magnitude to the one derived in Eswatini.

In conclusion, this exercise confirmed some key patterns in the determinants of compliance across two quite different countries. Rwanda, despite being an equally small and landlocked country, has a ten times larger population size than Eswatini. The open question still remains, however, on whether results will be consistent in even larger African countries and I consider this as a future research avenue to follow.

[43]This can be due to the higher difficulty in getting truthful responses on trust-related questions in Rwanda, as discussed in Chapter 2.

Table 3.5: Determinants of Active Filing Behaviour - External Validity

	(1) Eswatini	(2) Eswatini	(3) Rwanda	(4) Rwanda
	Deterrence			
Risk audit below median	-0.11***	-0.06**	-0.08**	-0.08**
	(0.03)	(0.03)	(0.03)	(0.03)
	Compliance costs			
Difficult to file	-0.10***	-0.08***	0.02	0.02
	(0.03)	(0.03)	(0.03)	(0.03)
Difficult to get in touch	-0.07**	-0.05*	-0.08***	-0.06**
	(0.03)	(0.03)	(0.03)	(0.03)
	Trust and political legitimacy			
Bribing above median	0.02	0.01	0.05	0.04
	(0.05)	(0.05)	(0.05)	(0.03)
Unfairness	-0.08**	-0.06**	-0.01	0.01
	(0.03)	(0.03)	(0.04)	(0.04)
No transparency	0.02	-0.01	-	-
	(0.04)	(0.03)	-	-
No trust above median	0.01	0.01	0.01	0.00
	(0.03)	(0.03)	(0.04)	(0.04)
	Fiscal exchange and reciprocity			
Services	-0.02	-0.02	0.00	0.00
	(0.02)	(0.01)	(0.01)	(0.01)
No fiscal exchange	0.00	-0.02	0.03	0.04
	(0.03)	(0.03)	(0.03)	(0.03)
Nothing in return	0.01	0.04	-	-
	(0.03)	(0.03)	-	-
	Social norms			
Evaders % above median	0.01	-0.01	-0.09**	-0.09**
	(0.05)	(0.05)	(0.04)	(0.04)
Peer pressure	0.11**	0.10**	-0.05	-0.07
	(0.05)	(0.04)	(0.11)	(0.10)
	Intrinsic motivation			
High tax morale	0.15***	0.11***	0.12**	0.15**
	(0.04)	(0.04)	(0.06)	(0.06)
Demographics	No	Yes	No	Yes
Business Char.	No	Yes	No	Yes
Mean of Y	0.513	0.513	0.634	0.634
R-sq.	0.057	0.326	0.043	0.152
Observations	1009	1009	1172	1172

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

3.6 Mechanisms

3.6.1 More on deterrence and risk preferences

One of the main findings from Table 3.2 refers to the role played by deterrence. In the attempt to further explore this result, I consider here two sets of additional variables. I first focus on the level of taxpayers' interactions with the revenue authority and, second, look at risk preferences. In Appendix Table A15, I regress the active filing dummy over a set of variables indicating the extent and intensity of taxpayers' interactions with the SRA. Results are not always consistent across specifications but some considerations can be derived. First, the fact of filing a return increases the likelihood of doing it again, in line with the recent evidence on paying taxes as a habit (Dunning et al., 2017; Mascagni et al., 2019).[44] Also, deterrence factors such as the fact of being audited and the number of audits positively explain active filing, even if only for last year's returns. Given the low number of auditors in the authority (see section 3.3.2), this could suggest that more resources could be channeled to audits. The experience of being fined has also a positive impact, while it seems that the frequency of pecuniary sanctions eventually backfires. Lastly, it seems that taxpayers are not benefitting much from interacting with the authority. Interestingly, those taxpayers that receive information on tax matters from SRA officials are less likely to file.[45] This adds up to the negative impact on filing of the difficulty of getting in touch with the authority, as displayed in Table 3.2 and calls for improvements in the way the SRA communicates with its clients.

When it comes to risk preferences, Appendix Table A16 shows that neither the experimental risk measure or the self-reported risk attitude seem to play any role in motivating compliance. It is true that, descriptively, CRRA risk averse taxpayers are more likely to perceive a higher probability of audit, 65%, than risk loving taxpayers, 58%. However, the

[44] This finding is confirmed further by looking at filing behaviour for tax year 2019, after the survey. 90% of taxpayers filing in 2018 filed again in 2019, while just 20% of non-filers did the same. When considering perpetual taxpayers, this difference is exacerbated even more, with 91% of persistent active and 12% of persistent non-filers filing in 2019.

[45] About 28% of the sample report to get tax-related information either formally or informally from SRA officials.

model in Table A16 rejects any significant correlation. Results are similar when I substitute the CRRA with a dummy for risk aversion.

3.6.2 The role of tax knowledge

One of the main findings of this study is that compliance costs matter. Ease of filing a return is a key predictor of the probability of actually doing so. It is interesting to dig deeper to understand the role of such costs and consider one important component of them – tax knowledge. Tax knowledge is included in the (highly significant) compliance costs principal component used in Table 3.4 and it is therefore important to consider its stand-alone impact. Higher tax knowledge seems also to be correlated with better perceptions on the ease to file, as depicted in Figure A9: taxpayers who believe it is somewhat or very easy to file report a 20% higher knowledge score than those who think it is somewhat or very difficult, with the difference being significant at the 1% level. As explained in section 3.4.1, the survey data contains answers to a quiz on tax, from which a tax knowledge index is formed. In Table 3.6, I study the impact on filing of either the raw knowledge score, ranging from 0 to 5 (col. 1-4) or the standardised score (Kling et al. 2001), expressed in terms of standard deviations (col. 5-8).

Most notably, Table 3.6 shows that the indicator of tax knowledge is always statistically significant in explaining compliance, from the least to the most complete specification (col. 1-4). One extra question answered correctly in the tax quiz is associated with an increase in the probability of filing last year's tax return of 7 percentage points (or 14%) when all controls are added (col. 3). The same figure is of 6 percentage points (or 9%) for perpetuals (col. 4). Consistently, a standard deviation increase in the Kling index implies an increase in filing of a similar magnitude (col. 7-8).

In the same fashion, Appendix Figure A10 plots the predictive marginal increase in the probability to file by each index score, ranging from 0 to 5. The increasing pattern is consistent for both last year's filing and perpetual filing. It results that tax knowledge plays a greater role in explaining persistent compliance over time, with taxpayers scoring the maximum being almost 100% likely to be perpetual filers.

As an additional investigation, I also run the main specification using as regressors each of the five tax knowledge questions composing the score. This can help shed light on the

Table 3.6: Active vs Non-filers - Tax Knowledge

	(1) All	(2) Perpetuals	(3) All	(4) Perpetuals	(5) All	(6) Perpetuals	(7) All	(8) Perpetuals
Knowledge score 0-5	0.13***	0.15***	0.07***	0.06***				
	(0.02)	(0.02)	(0.02)	(0.02)				
Standardised score					0.19***	0.22***	0.09***	0.07**
					(0.03)	(0.04)	(0.03)	(0.04)
Demographics	No	No	Yes	Yes	No	No	Yes	Yes
Business Char.	No	No	Yes	Yes	No	No	Yes	Yes
Mean of Y	0.513	0.644	0.513	0.644	0.513	0.644	0.513	0.644
R-sq.	0.059	0.090	0.313	0.456	0.042	0.061	0.309	0.452
Observations	1009	613	1009	613	1009	613	1009	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The construction of the knowledge scores is discussed in section 3.4.3.

awareness of which aspect of the system is more critical in order to comply. Results are displayed in Appendix Table A17, both without (Panel A) and with controls (Panel B). Results are quite informative, as answering correctly to some specific questions strongly predicts filing behaviour, while it does not for others. Namely, question 1 (inactive should file anyway), 4 (filing deadline) and 5 (tax type the taxpayer is registered for) are always statistically significant. On the other hand, questions on the size of the penalty for missing a declaration (Q2) and income understatement (Q3) are never significant.

This evidence is illuminating in the sense that it seems that non-filers are not aware of very basic tax-related rules, such as the requirement to always file or even their own tax type, while possibly more complex concepts such as the penalty amounts do not discriminate between them and active.[46] The fact that knowledge of the penalty structure does not affect compliance is due to the extremely low level of knowledge of the main penalty amounts across all taxpayers. Stunningly, only 2.6% and 1.5% of the sample are aware of the penalty amounts for failing to file and false declarations, respectively. Evidence of under or overestimation of penalty amounts is almost inexistent, since the vast majority just answer that they do not know (88% and 94%) rather than providing an estimate. This

[46]Very much consistently, recent evidence collected in a nation-wide survey of 2,000 taxpayers in Rwanda shows exactly the same pattern (Mascagni and Santoro, forthcoming)

means that deterrence is shaping compliance more through the perceived audit probability (as in Table 3.2) than through the penalty structure. In turn, audit probability is probably overestimated given the limited resources available to the authority (section 3.3).

Lastly, a note of concern consists in the possibility of reverse causality being at play here. For example, if intrinsic motivation in paying taxes is a significant predictor of compliance – as is the case here – then this might lead to acquiring more knowledge about the tax system. However, the data contradicts this hypothesis. The 83% of the sample with a high intrinsic motivation show, if anything, a lower tax knowledge score (1.54) than those with a low motivation (1.62). Another concern could come from the fact of being in the system for a longer period, which could in turn affect tax knowledge. Also this concern is not reflected in the data. First, the coefficients of tax knowledge in Table 3.6 are not consistently higher when restricting the analysis to perpetuals. If anything, magnitudes get smaller when controls are added. Second, the 15% of the sample who had a previous business before the current one report slightly lower scores (1.53) than those who are at their first experience (1.56). Third, even if it is true that years since registration and tax knowledge have a positive correlation, the magnitude of it is quite small, 0.11, and in any case a variable indicating years since registration is included as a control in my model.

3.6.3 The relevance of background characteristics

As a last set of supplementary results, I study the importance of two sets of background characteristics: (i) demographic and individual tax practice related covariates, and (ii) business-level covariates. As already discussed in section 3.2.1, abundant empirical evidence has been produced stressing the relevance of such variables, who are often considered just as background controls. Consistently, based on the R^2 of the regression tables above, demographics and business characteristics appear to be explaining a lot of the variability in the model, even more than the variables of interest. For this reason, it is important to understand how they correlate with active filing as this would be valuable information for risk management and audit strategies.

Table A18 reports the coefficients of taxpayer-level features in columns 1-2, business-level features in columns 3-4 and both groups together in columns 5-6.[47] Among the

[47] In order to avoid multicollinearity, some variables such as using e-mails, book-keeping and suffering from

first set of variables, demographics in columns 1-2 seem to play a key role in affecting compliance. As expected, age is positively related to filing behaviour, as well as having higher education, even if they lose significance in the more complete models of columns 5-6. In line with the existing evidence, being married seems to positively affect filing behaviour, but for last year's declaration only. Gender and nationality do not have any strong impact, even if the sign of the coefficient seem to suggest that female and Swazi national are less likely to comply, unexpectedly.

Furthermore, individual tax-related practices decisively influence the outcome: the factor *has tax accountant* remains highly significant across all different specifications. It results that having a tax accountant increases the probability of being active by 16 percentage points when considering last year's filing (col. 5) and 21 percentage points when considering the persistent behaviour (col. 6). These coefficients are sizeable and among the largest found across all sets of results: having a tax accountant translates into being 31% more likely to file in a given year and 33% more likely to be persistently active. This finding is in line with the strong impact of the perception on the difficulty to file from Table 3.2, as well as with the corresponding principal component index from Table 3.4. Unsurprisingly, also the time spent on tax matters is positively correlated with actively filing, even if weakly so. The same is true for having a bank account for the business.

Additionally, when considering business-level characteristics, being operative has a sizeable impact on filing of about 31-32 percentage points (col. 5-6), or 61% for last year's filing and 50% for persistent filing. The coefficient of being operative is the largest in magnitude among all those observed in this study and points to the fact that taxpayers file their taxes only when operative, contrarily to what the law prescribes. From the role of the corresponding tax knowledge question in Table A17, it can be assumed that inoperative businesses do not file partly because they are not aware that it is required by the law.

Also, it can be noticed that older firms are more likely to file, probably due to the fact that they are more likely to be operative and run by older taxpayers, who are more compliant. Trading sector is also presumably a sector with more frequent operations taking place, hence its significant role in explaining compliance. Despite being smaller in size,[48]

competition with informal businesses have been removed since highly correlated with the variables used in this exercise.

[48] The average turnover for trading activities is of USD 1,383, somehow smaller than that of other service

taxpayers in the trading sector are more likely to be operative than the others, 75.5% versus 72%. It is worth specifying that this sector encompasses a wide range of activities, such as transport, storage, accommodation and food services, who serve the open public and are therefore more exposed to inspection (Monteiro and Assuncao, 2012). While the probability of being actually audited is lower for traders (13%) than taxpayers in all other sectors (17.5%), traders are much more likely to have been fined (28%) than the others (21.5%), statistically significantly so at the 5% level. Relatedly, 59% of traders have an audit risk perception above median compared to 54% of taxpayers in other sectors.

3.7 Conclusions and policy recommendations

In this paper, I have explored the factors that correlate with taxpayers' compliance in Eswatini building on rich attitude and perception data from a nationally representative sample of one thousand sole traders. The data collection represents the first exercise of this type ever carried out in the country. Since self-reported compliance is likely to be inaccurate, I link the survey data to tax returns data from the Eswatini Revenue Authority, which enables me to identify compliant (active/filing) and non-compliant (non-filing) taxpayers. To the best of my knowledge, tax data from Eswatini has never been explored in the literature. Also, I compare the relevance of theoretically founded motivations to actually file a return with those explaining self-reported compliance and the intensive margin of compliance. As a robustness check, I employ dimension reduction and best subset selection methods and a discrete probit model, as well as control for enumerators' ability.

The results provide a complex and nuanced picture of tax compliance in Eswatini that can be summarised in the following points. First, standard deterrence motives are at work, with higher perceptions of the audit probability being strongly associated with active filing. Interestingly, a stronger sense of state legitimacy and fear of getting caught is more prevalent in active vis-a-vis non-filers when it comes to the reasons they registered with the

activities (the second most common sector, 26% of the sample), with about USD 2,400 of turnover, and many times smaller than real estate (USD 15,000), construction (USD 8,971), agriculture (USD 6,184), manufacturing (USD 5,769).

authority in the first place (Figure A5). Second, other non-standard determinants are also crucial in shaping taxpayers' compliance. In particular, compliance costs, social norms and intrinsic tax morale positively covary with filing behaviour. Third, some important non-pecuniary factors, such as trust in the authority, political legitimacy of the State and fiscal exchange, which have proved to be essential in similar studies, do not seem to be important in Eswatini. These three sets of results also confirm – at least in part – the validity of the 3-tiered World Bank framework for understanding tax compliance, as formulated in Prichard et al. (2019). Consistently with the framework, factors related to *enforcement* (audit risk) and *facilitation* (ease to file, tax knowledge) paradigms are at work. For what concerns the third paradigm in the framework, the one based on *trust*, results are more inconclusive, with only a strong influence of peers taking place. As a last result, self-reported tax compliance is driven by partly different factors than actual tax compliance. While this is true for Eswatini, this finding may also suggest that, more in general, researchers should be cautious in using self-reports as a proxy for actual compliance.

The limitations of this study need to be acknowledged. First, I cannot be completely confident that I am capturing causal relationships. Despite the extensive use of fine controls, the robustness to alternative estimation methods, and the restriction on persistent filing behaviour to address unobserved variability over time, I cannot rule out the possibility that other unobservables may be linked to both the explanatory factors and the outcome variable.

Second, mostly due to time constraint in implementing the survey, some additional background information is missing, such as the extent to which taxpayers in my sample also pay local fees or informal contributions to non-state actors, information on political engagement and pro-social behaviour in the community, or a more refined measure of risk preferences. Related to the latter point, the fact that the coefficient of risk aversion as derived from the lottery is not significantly explaining compliance may be linked, for example, to the absence of real-stake lottery decisions.

Third, mostly due to budget constraints, the main focus of this study is on the extensive margin of compliance only, i.e. the probability to file a return, while I cannot explore the drivers of the intensive margin of compliance, i.e. income underreporting. Extensive and intensive compliance are likely to be explained by different set of motivations and

I leave this to future research. In the attempt to explore the intensive margin at least descriptively, I am able to link the survey data with tax returns for the tax year 2019, submitted after the survey. I then compare the self-reported business income as extracted with the questionnaire (see Figure A4) with what is actually declared in the tax return. Surprisingly, (i) the vast majority of taxpayers in the sample, 79%, reports a lower income than what was declared in the survey; (ii) a minority of 7% and 14% declare the same or a higher income, respectively; (iii) non-filers are more likely (85%) to under-declare than active (74%) but both figures remain high; (iv) the gap is increased when comparing persistent non-filers (88%) with persistent active (73%); (v) the average underreporting is higher for active (USD 31,000) than non-filers (USD 9,500). This initial evidence calls for further research on such discrepancies: while evasion can surely be part of the story, additional explanations such as poor record-keeping and computational constraints might affect these results.

On a different note, my study focusses on registered taxpayers only, and I do not study informal traders. Determinants motivating the compliance of registered taxpayers may be different than, for example, those pushing informal traders to register. At the same time, it could be argued that non-filers in my sample resemble informal traders in the fact that the majority of them (56%) report being in operation despite not sharing any information with the authority. Future research could be devoted to study how registered and non-registered entities differ (see Chapter 5).

Despite the weaknesses, this study points to some important policy recommendations. First, it shows the revenue authority that enforcement is important. The SRA should continue stressing its role as a monitoring agency. A wiser use of the limited resources would imply that increased auditing efforts can be directed towards non-filers, who can be automatically detected on the database and contacted by tax officials. The system could automatically trigger follow-up messages or reminders to non-filers, signalling that the authority is aware of their failure to file and has the technical resources to track their behaviour.

Second, the authority should focus more on improving taxpayers' awareness and knowledge. Educational initiatives could also be tailored to non-filers more specifically, given that they lack the knowledge of very basic concepts. Survey data show that lack of knowledge is an important obstacle to filing for 89 per cent of non-filers, while the same figure for active payers is 72%. Consistently, knowledge of the tax system (46%) and how to file a return (30%) are the most urgent aspects for which taxpayers would like to receive assistance from the authority. This can happen through a variety of options. In the sample under study, while fewer taxpayers use online tools (12%), more rely on direct relations with tax officials (29%) and the majority use more traditional methods such as radio/TV (57%) as the main channel of getting tax-related information. While this evidence highlights the importance of radio/TV and direct interactions with taxpayers, it is important to target resources carefully towards channels that are likely to have the biggest impact – especially in the context of typically under-resourced taxpayer education departments. More experimental studies, such as randomised controlled trials, can better test the effectiveness of alternative strategies, such as one-to-one coaching vs radio programs (Chapter 2). This shift towards a service-based paradigm should also affect the way the SRA provides information to taxpayers, as it seems that currently getting information from the authority is not correlated with active filing. This is also linked to the fact that non-filers, when they interact with SRA officials, are less likely to discuss about filing a return (25%) than active taxpayers (35%). Overall, 40% of the sample see it difficult to get in touch with the authority to receive assistance. Similarly, communication with taxpayers (46%) is the most frequently mentioned area in which taxpayers believe the authority is underperforming. This calls for an improvement in the communication strategy: as it results from Table 3.2, non-persistent taxpayers are negatively affected by communication issues, meaning that they may easily turn to non-filing if not promptly reached by the authority.

Third, another possible avenue of intervention is provided by a major focus on the social norms of compliance. The SRA could exploit the fact that filing taxes seems to be motivated by adhesion to a social norm and could adopt a new way of communication which stresses this aspect.

Fourth, while it seems that trust, transparency and reciprocity motives are not important, the authority should not neglect them and possibly find better ways of emphasising these concepts in its communication strategy.

Fifth, the tax administration itself could adapt its strategies to the fact that a gap exists between self-reported intentions and actual filing behaviour. Knowing what drives willingness to comply is as important as knowing what motivates the decision to file. From the results in Table 3.3 it seems that, once again, assisting taxpayers with effective educational and communication strategies might increase their intrinsic motivation to comply.

In conclusion, this paper will hopefully encourage more researchers to engage in primary data collection in relation to tax and development. Eswatini is a small country and it is an open question whether these lessons can be applied to other contexts. Cross-country comparisons will surely be beneficial in gaining a better understanding of what drives compliance in Africa. At the same time, this paper makes the point for a stronger reliance on tax administrative data, which revenue authorities in SSA produce every day. An important future direction for research consists in exploiting the combined potential of survey and administrative data, so to gain direct knowledge of the practical *life* of taxation in low and middle income countries and eventually inform more realistic and successful tax policies.

Appendices

A1 Research background

Table A1: Taxpayer's Perception Survey

No.	Module	# Questions
1	Pre-interview identifying information	5
2	Consent form	4
3	Respondent's demographics	4
4	Business' characteristics	23
5	Risk preferences	9
6	Tax Attitudes and perceptions	28
7	Satisfaction with public services	6
8	Interactions with revenue authority	12
9	Post-interview quality assessment	4

Module 1 and 9 were filled by the enumerator herself.

A1.1 Risk Preferences

In the Multiple Price List experiment, each respondent is presented with a choice between two lotteries, A or B. Appendix Table A2 shows the payoffs structure implied in the experiment.[49] At the beginning of the experiment, the two lotteries have a relatively large difference in expected values, such as SLZ 3,000 in lottery 1. As one proceeds down the matrix, the expected value of lottery A stays the same, while that of B increases, so that the difference in payoff is now in favour of B. The logic behind the test is that only risk-loving subjects would take lottery B in the first and second row, and only risk-averse subjects would take lottery A in the last three row.[50] In line with the relevant literature, risk attitude is operationalised with the coefficient of relative risk aversion (CRRA), which is calculated for each lottery choice, as shown in Table A2.[51]

[49]It is worth stressing that the last four columns of the table were not shown to the respondent.

[50]A risk-neutral respondent should switch from choosing A to B when the difference between the two payoffs is about zero, so she would choose A for the first two/three rows and B thereafter.

[51]The CRRA utility is defined as $U(y) = (y^{1-r})(1-r)$, where r is the CRRA coefficient. With this parameterization, $r = 0$ denotes risk-neutral behavior, $r > 0$ denotes risk aversion, and $r < 0$ denotes risk loving. When $r = 1$, $U(m) = ln(m)$. More details in Harrison et al. (2005).

Table A2: Lottery Choices and Risk Aversion Classification

Lottery A		Lottery B				Risk profile			
Prob.	Payoff	Prob.	Payoff	Prob.	Payoff	Diff.	CRRA interval	CRRA est.	Risk category
1	11000	0.5	16000	0.5	0	3000	$r <$ -0.85	-1.23	very risk loving
1	9500	0.5	16000	0.5	0	1500	-0.85 $< r <$ -0.33	-0.59	risk loving
1	8000	0.5	16000	0.5	0	0	-0.33 $< r <$ 0	-0.16	slightly loving to neutral
1	6500	0.5	16000	0.5	0	-1500	0 $< r <$ 0.23	0.16	neutral to slightly averse
1	5000	0.5	16000	0.5	0	-3000	0.23 $< r <$ 0.40	0.31	risk averse
1	3500	0.5	16000	0.5	0	-4500	0.40 $< r <$ 0.54	0.47	very risk averse
1	2000	0.5	16000	0.5	0	-6000	0.54 $< r <$ 0.67	0.60	highly risk averse

All currency units are Swazi Lilangeni (SZL). At the time of the experiment 1 USD=15.02 SZL. The last three columns in this table, showing the difference in expected values of the lotteries and the implied CRRA intervals, were not shown to subjects. Based on expected utility theory and assuming constant relative risk aversion, the CRRA parameter r refers to a utility function $U(x) = x^{1-r}(1-r)^{-1}$. The CRRA intervals refer to the choice of switching to lottery B. In case the subject never switches to lottery B, his CRRA interval is 0.67 to infinity. CRRA estimates are approximated as midpoints of the closed CRRA intervals.

A2 Results

Table A3: Mean differences by consent to the survey

	Refuse		Consent		
	Mean	Obs.	Mean	Obs.	Difference
Perpetual Active	0.74	984	0.76	518	-0.02
Perpetual Non-filer	0.49	1481	0.44	491	0.04*
Hhohho	0.34	2465	0.37	1009	-0.02
Lubombo	0.14	2465	0.16	1009	-0.02
Manzini	0.41	2465	0.38	1009	0.03
Shiselweni	0.11	2465	0.09	1009	0.02
# Years filing 2014-2018	4.39	2465	4.30	1009	0.09**
VAT registered	0.01	2465	0.02	1009	-0.01
Log Tax declared	5.00	984	3.63	518	1.36***
N	3,474				

Table A4: Summary Statistics - Background Variables

	N	Mean	SD	Min	Max
		Taxpayer-level			
Female	1009	0.40	0.49	0.00	1.00
Age group max=7	1009	4.27	1.24	0.00	7.00
Higher education	1009	0.35	0.48	0.00	1.00
Swazi national	1006	0.95	0.21	0.00	1.00
Married	576	0.63	0.48	0.00	1.00
Has tax accountant	1009	0.57	0.49	0.00	1.00
Days spent on tax	1009	4.72	6.27	0.00	31.00
Time on tax > median	815	0.39	0.49	0.00	1.00
Book-keeping	1009	0.65	0.48	0.00	1.00
Email for business	1009	0.20	0.40	0.00	1.00
Bank account	1009	0.58	0.49	0.00	1.00
		Business-level			
Operative	1009	0.74	0.44	0.00	1.00
Years since registration	1009	6.33	3.31	0.00	12.00
Had a previous business	1009	0.15	0.35	0.00	1.00
Log(USD turnover)	1009	1.77	1.36	0.00	5.12
Hhohho	1009	0.37	0.48	0.00	1.00
Lubombo	1009	0.16	0.37	0.00	1.00
Manzini	1009	0.38	0.49	0.00	1.00
Shiselweni	1009	0.09	0.29	0.00	1.00
Wholsale/retail trade	1009	0.56	0.50	0.00	1.00
Compete with informals	871	0.80	0.40	0.00	1.00
High Competition	1009	0.74	0.44	0.00	1.00
Less business	958	0.52	0.50	0.00	1.00

Table A5: Summary Statistics - Risk Preferences

	N	Mean	SD	Min	Max
Self-reported riskiness	1009	5.12	2.66	1.00	10.00
CRRA	903	0.17	0.84	-1.23	0.98
Risk averse	990	0.70	0.46	0.00	1.00
Switch-point to risky lottery	903	4.02	2.78	0.00	7.00
% A for choice 1	920	0.79	0.41	0.00	1.00
% A for choice 2	915	0.74	0.44	0.00	1.00
% A for choice 3	913	0.67	0.47	0.00	1.00
% A for choice 4	907	0.57	0.50	0.00	1.00
% A for choice 5	904	0.47	0.50	0.00	1.00
% A for choice 6	899	0.42	0.49	0.00	1.00
% A for choice 7	900	0.39	0.49	0.00	1.00
Indifference	1009	0.09	0.28	0.00	1.00
Inconsistency	1009	0.02	0.14	0.00	1.00

Table A6: Summary Statistics - Interactions with SRA

	N	Mean	SD	Min	Max
# Years filing 2014-2018	1009	4.30	1.22	1.00	5.00
Distance to SRA (km)	1009	40.41	34.27	0.43	129.45
Ever audited	957	0.15	0.36	0.00	1.00
# audits	78	1.50	0.66	1.00	3.00
Ever fined	960	0.25	0.43	0.00	1.00
# fines	171	1.70	2.05	1.00	16.00
Interacted with SRA	968	0.40	0.49	0.00	1.00
# interactions	326	2.28	2.94	1.00	30.00
N	1009				

Table A7: Summary Statistics - Key Factors

	N	Mean	SD	Min	Max
	Deterrence				
Risk audit below median	988	0.43	0.50	0.00	1.00
Audit % general	988	83.05	25.03	0.00	100.00
Own audit %	988	61.94	33.58	0.00	100.00
# fined businesses	986	1.43	2.83	0.00	20.00
	Compliance costs				
Knowledge score max=5	1009	1.56	0.95	0.00	5.00
Standardized score	1009	-0.00	0.54	-0.75	3.52
Difficult to file	1009	0.65	0.47	0.00	1.00
Difficult to get in touch	1009	0.40	0.49	0.00	1.00
	Trust and political legitimacy				
Bribing	649	0.55	0.49	0.00	1.00
% bribe over sales	413	13.43	19.85	0.00	100.00
Unfairness	1009	0.47	0.50	0.00	1.00
No transparency	1009	0.74	0.44	0.00	1.00
Poor SRA performance	933	2.51	1.18	1.00	5.00
No trust max=4	961	2.51	1.07	1.00	4.00
	Fiscal exchange and reciprocity				
Primary schools	1009	3.07	1.34	1.00	5.00
Tertiary Education	1009	2.57	1.28	1.00	5.00
Roads/bridges	1009	2.35	1.33	1.00	5.00
Electricity	1009	2.54	1.39	1.00	5.00
Healthcare	1009	2.77	1.34	1.00	5.00
Security/police	1009	2.79	1.34	1.00	5.00
No fiscal exchange	1009	0.58	0.49	0.00	1.00
Nothing in return	1009	0.48	0.50	0.00	1.00
	Social norms				
% neighbours evading	401	39.73	31.98	0.00	100.00
Peer pressure	1009	0.12	0.33	0.00	1.00
	Intrinsic motivation				
High tax morale	1009	0.82	0.39	0.00	1.00

A3 Robustness checks

Table A8: Determinants of very strong Self-reported Compliance

	(1)	(2)
Deterrence		
Risk audit below median	-0.01	-0.00
	(0.03)	(0.03)
Compliance costs		
Difficult to file	-0.06**	-0.09***
	(0.03)	(0.03)
Difficult to get in touch 1-4	-0.04***	-0.04***
	(0.01)	(0.01)
Trust and political legitimacy		
Bribing above median	-0.10**	-0.05
	(0.04)	(0.04)
Unfairness	-0.02	0.02
	(0.03)	(0.03)
No transparency	-0.00	0.04
	(0.03)	(0.03)
No trust above median	-0.06**	-0.03*
	(0.03)	(0.03)
Fiscal exchange and reciprocity		
Services	0.01	0.03**
	(0.01)	(0.01)
No fiscal exchange	0.03	0.00
	(0.03)	(0.03)
Nothing in return	-0.02	-0.01
	(0.03)	(0.03)
Social norms		
Evaders % above median	-0.01	0.00
	(0.04)	(0.04)
Peer pressure	-0.40***	-0.32***
	(0.05)	(0.05)
Demographics	No	Yes
Business Char.	No	Yes
Mean of Y	0.699	0.699
R-sq.	0.145	0.265
Observations	1009	1009

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

Table A9: High Tax Morale - Tax Perceptions - No Mismatchers

	(1)	(2)
Deterrence		
Risk audit below median	-0.04	-0.03
	(0.03)	(0.03)
Compliance costs		
Difficult to file	-0.13***	-0.13***
	(0.03)	(0.03)
Difficult to get in touch 1-4	0.05***	0.04***
	(0.02)	(0.01)
Trust and political legitimacy		
Bribing above median	-0.06	-0.04
	(0.05)	(0.05)
Unfairness	-0.03	-0.01
	(0.04)	(0.03)
No transparency	-0.00	-0.00
	(0.04)	(0.04)
No trust above median	-0.08**	-0.02
	(0.03)	(0.03)
Fiscal exchange and reciprocity		
Services	-0.01	0.01
	(0.02)	(0.02)
No fiscal exchange	0.10***	0.07**
	(0.03)	(0.03)
Nothing in return	-0.04	-0.02
	(0.04)	(0.04)
Social norms		
Evaders % above median	-0.09*	-0.07
	(0.06)	(0.05)
Peer pressure	-0.41***	-0.29***
	(0.05)	(0.05)
Demographics	No	Yes
Business Char.	No	Yes
Mean of Y	0.706	0.706
R-sq.	0.239	0.399
Observations	629	629

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

Table A10: PCA Indexes - First Components

Demo	Business fixed	Doing business	Deterrence	Compliance costs	FE	Trust	Social norms	Morale
				Variables and Coefficients				
Siswati 0.56	Business age -0.13	Operative 0.58	Audit Y/N 0.43	Knowledge 0.48	No FE 0.57	Bribing Y/N 0.31	Nonfilers 0.69	High morale 1
Educ. -0.68	Years filing -0.11	Low comp. 0.02	# audits 0.37	Easy to file 0.32	Nothing back 0.58	Bribe % 0.14	Nilfilers 0.69	
Age 0.21	Previous -0.07	Increasing 0.13	Fine Y/N 0.43	Accountant 0.29	Satisf. -0.59	No trust 0.54	Evaders % -0.08	
Female 0.42	Trade 0.13	Turnover 0.51	# fines 0.40	Time on tax 0.32		No transp. 0.26	Neighbours 0.20	
	Dist. SRA 0.57	Bank 0.61	Interact Y/N 0.38	Books 0.54		Unfairness 0.51		
	Hhohho -0.63		# interactions 0.35	Email 0.35		Poor SRA 0.51		
	Lubombo 0.36		Own audit 0.13					
	Manzini 0.19		Others' audit 0.09					
	Shiselweni 0.27		# peers 0.23					
				Eigenvalues				
1.55	1.56	1.81	2.24	3.31	2.50	2.01	1.85	-
				Variance explained				
0.26	0.26	0.36	0.28	0.28	0.31	0.34	0.23	1

Table A11: Active vs Non-filers - Statistical Learning Results - Last Year's Behaviour

	(1) Original LPM	(2) LPM	(3) Stepwise	(4) CV Lasso	(5) AIC	(6) AICC	(7) BIC	(8) EBIC	(9) New LPM
				Deterrence					
Risk audit below median	-0.06**	-0.13***	-0.12***	-0.05	-0.09	-0.09	-0.07	-0.07	-0.06**
	(0.03)	(0.04)	(0.04)						(0.03)
				Compliance costs					
Difficult to file	-0.08***	-0.11**	-0.09**	-0.07	-0.06	-0.06	-0.04	-0.04	-0.09***
	(0.03)	(0.05)	(0.04)						(0.03)
Difficult to get in touch	-0.05*	-0.05		-0.01					
	(0.01)	(0.02)							
				Trust and political legitimacy					
Bribing above median	0.01	0.06							
	(0.05)	(0.07)							
Unfairness	-0.06**	-0.07*		-0.04	-0.03	-0.03			-0.07**
	(0.03)	(0.04)							(0.03)
No transparency	-0.01	0.02							
	(0.03)	(0.05)							
No trust above median	0.01	0.05							
	(0.03)	(0.04)							
				Fiscal exchange and reciprocity					
Services	-0.02	-0.03		-0.01	-0.01	-0.01			-0.02
	(0.01)	(0.02)							(0.01)
No fiscal exchange	-0.02	0.02							
	(0.03)	(0.04)							
Nothing in return	0.06*	-0.03		0.01					0.03
	(0.03)	(0.04)							(0.03)
				Social norms					
Evaders % above median	-0.01	0.01							
	(0.05)	(0.06)							
Peer pressure	0.10**	0.06		0.06					0.10**
	(0.04)	(0.06)							(0.04)
				Intrinsic motivation					
High tax morale	0.11***	0.10*	0.10**	0.07	0.06	0.06	0.03	0.03	0.11***
	(0.04)	(0.06)	(0.05)						(0.04)
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Business Char.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1009	505	505	1009	505	505	505	505	1009
RMSE out	-	0.44	0.43	0.41	0.43	0.43	0.43	0.43	-
RMSE in	-	0.40	0.41	0.42	0.42	0.43	0.43	0.43	-

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3, while each statistical learning method is presented in section 3.5.3.

Table A12: Determinants of Active Filing Behaviour - Probit Model

	(1) All	(2) Perpetuals	(3) All	(4) Perpetuals
		Deterrence		
Risk audit below median	-0.12***	-0.14***	-0.07*	-0.10*
	(0.03)	(0.04)	(0.04)	(0.06)
		Compliance costs		
Difficult to file	-0.11***	-0.11**	-0.10**	-0.14**
	(0.04)	(0.04)	(0.04)	(0.06)
Difficult to get in touch	-0.06**	-0.03	-0.06*	-0.02
	(0.03)	(0.04)	(0.03)	(0.03)
		Trust and political legitimacy		
Bribing above median	0.02	-0.03	0.02	-0.07
	(0.05)	(0.07)	(0.06)	(0.09)
Unfairness	-0.09**	-0.06	-0.09**	-0.06
	(0.04)	(0.04)	(0.04)	(0.06)
No transparency	0.02	0.04	-0.02	-0.00
	(0.04)	(0.05)	(0.05)	(0.06)
No trust above median	0.01	-0.00	0.01	-0.01
	(0.04)	(0.04)	(0.04)	(0.06)
		Fiscal exchange and reciprocity		
Services	-0.03	-0.02	-0.02	-0.03
	(0.02)	(0.02)	(0.02)	(0.03)
No fiscal exchange	0.00	-0.01	-0.03	-0.09
	(0.03)	(0.04)	(0.04)	(0.06)
Nothing in return	0.01	0.00	0.06	0.09
	(0.04)	(0.04)	(0.04)	(0.06)
		Social norms		
Evaders % above median	0.01	0.06	-0.02	0.10
	(0.05)	(0.06)	(0.06)	(0.08)
No peer pressure	0.11**	0.25***	0.13**	0.34***
	(0.05)	(0.07)	(0.06)	(0.08)
		Intrinsic motivation		
High tax morale	0.16***	0.13**	0.15***	0.15*
	(0.05)	(0.06)	(0.05)	(0.08)
Demographics	No	No	Yes	Yes
Business Char.	No	No	Yes	Yes
Mean of Y	0.513	0.644	0.513	0.644
Observations	1009	613	936	547

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are marginal effects from probit regressions evaluated at the mean. The operationalisaton of determinants is discussed in section 3.4.3.

Table A13: Active vs Non-filers - Tax Perceptions - Enumerators Fixed Effects

	(1) All	(2) Perpetuals	(3) All	(4) Perpetuals
		Deterrence		
Risk audit below median	-0.13***	-0.16***	-0.08***	-0.09***
	(0.03)	(0.04)	(0.03)	(0.03)
		Compliance costs		
Difficult to file	-0.08**	-0.08*	-0.08***	-0.08**
	(0.04)	(0.04)	(0.03)	(0.03)
Difficult to get in touch	-0.07*	-0.05	-0.05*	-0.02
	(0.04)	(0.04)	(0.03)	(0.03)
		Trust and political legitimacy		
Bribing above median	0.02	-0.02	0.00	-0.06
	(0.05)	(0.07)	(0.05)	(0.06)
Unfairness	-0.10***	-0.08*	-0.06**	-0.04
	(0.03)	(0.04)	(0.03)	(0.03)
No transparency	-0.00	0.01	0.00	0.04
	(0.04)	(0.05)	(0.04)	(0.04)
No trust above median	0.01	-0.01	0.00	-0.02
	(0.04)	(0.04)	(0.03)	(0.03)
		Fiscal exchange and reciprocity		
Services	-0.03	-0.03	-0.02	-0.02
	(0.02)	(0.02)	(0.01)	(0.02)
No fiscal exchange	-0.00	-0.02	-0.04	-0.05
	(0.04)	(0.04)	(0.03)	(0.03)
Nothing in return	0.00	-0.01	0.03	0.05
	(0.03)	(0.04)	(0.03)	(0.03)
		Social norms		
Evaders % above median	-0.01	0.05	-0.01	0.04
	(0.05)	(0.07)	(0.05)	(0.05)
Peer pressure	0.10*	0.23***	0.09**	0.19***
	(0.05)	(0.06)	(0.04)	(0.04)
		Intrinsic motivation		
High tax morale	0.18***	0.17***	0.13***	0.09*
	(0.05)	(0.06)	(0.04)	(0.05)
Enumerator FE	Yes	Yes	Yes	Yes
Demographics	No	No	Yes	Yes
Business Char.	No	No	Yes	Yes
Mean of Y	0.513	0.644	0.513	0.644
R-sq.	0.074	0.102	0.334	0.495
Observations	1009	613	1009	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

Table A14: Active vs Non-filers - Tax Perceptions - Survey Day Fixed Effects

	(1) All	(2) Perpetuals	(3) All	(4) Perpetuals
	Deterrence			
Risk audit below median	-0.12***	-0.11***	-0.07**	-0.07**
	(0.03)	(0.04)	(0.03)	(0.03)
	Compliance costs			
Difficult to file	-0.10***	-0.10**	-0.08**	-0.09***
	(0.03)	(0.04)	(0.03)	(0.03)
Difficult to get in touch	-0.06*	-0.05	-0.03*	-0.01
	(0.01)	(0.02)	(0.01)	(0.01)
	Trust and political legitimacy			
Bribing above median	0.02	-0.03	0.01	-0.05
	(0.05)	(0.07)	(0.05)	(0.05)
Unfairness	-0.08**	-0.06	-0.06**	-0.04
	(0.03)	(0.04)	(0.03)	(0.03)
No transparency	0.01	0.03	-0.01	0.02
	(0.04)	(0.05)	(0.03)	(0.04)
	Fiscal exchange and reciprocity			
Services	-0.03	-0.02	-0.02	-0.02
	(0.02)	(0.02)	(0.01)	(0.02)
No fiscal exchange	0.01	-0.02	-0.02	-0.05
	(0.03)	(0.04)	(0.03)	(0.03)
Nothing in return	0.01	0.00	0.04	0.06*
	(0.03)	(0.04)	(0.03)	(0.03)
	Social norms			
Evaders % above median	-0.01	0.05	-0.01	0.03
	(0.05)	(0.06)	(0.05)	(0.05)
Peer pressure	0.11**	0.23***	0.10**	0.20***
	(0.05)	(0.06))	(0.04)	(0.04)
	Intrinsic motivation			
High tax morale	0.15***	0.12**	0.11***	0.08*
	(0.04)	(0.06)	(0.04)	(0.05)
Day FE	Yes	Yes	Yes	Yes
Demographics	No	No	Yes	Yes
Business Char.	No	No	Yes	Yes
Mean of Y	0.513	0.644	0.513	0.644
R-sq.	0.065	0.086	0.330	0.489
Observations	1009	613	1009	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

A4 Mechanisms

Table A15: Active vs Non-filers - Interactions with Revenue Authority

	(1)	(2)	(3)	(4)
	All	Perpetuals	All	Perpetuals
# Years filing 2014-2018	0.03**	0.08***	-0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Far from SRA	0.02	-0.02	0.07**	0.05
	(0.03)	(0.04)	(0.03)	(0.04)
Ever audited	0.18**	0.10	0.10	0.00
	(0.09)	(0.08)	(0.08)	(0.07)
# audits	0.07	0.07	0.10*	0.05
	(0.07)	(0.05)	(0.06)	(0.06)
# audited peers	0.03	0.05	0.00	0.01
	(0.03)	(0.04)	(0.03)	(0.03)
Ever fined	0.03	0.07	0.01	0.06
	(0.09)	(0.09)	(0.08)	(0.07)
# fines	-0.03*	-0.02	-0.03	-0.02
	(0.02)	(0.02)	(0.02)	(0.01)
Interacted with SRA	0.03	0.01	-0.02	-0.02
	(0.08)	(0.09)	(0.07)	(0.07)
# interactions	0.01	0.01	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Info from SRA staff	-0.05	-0.07*	-0.03	-0.03
	(0.03)	(0.04)	(0.03)	(0.04)
Demographics	No	No	Yes	Yes
Business Char.	No	No	Yes	Yes
Mean of Y	0.513	0.644	0.513	0.644
R-sq.	0.102	0.157	0.331	0.467
Observations	1009	613	1009	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

Table A16: Active vs Non-filers - Risk Attitudes

	(1) All	(2) All	(3) Perpetuals	(4) Perpetuals	(5) All	(6) All	(7) Perpetuals	(8) Perpetuals
CRRA Lottery	-0.02		-0.04		0.01		0.01	
	(0.02)		(0.02)		(0.02)		(0.02)	
Risk reported > median		-0.03		-0.02		-0.03		-0.04
		(0.03)		(0.04)		(0.03)		(0.03)
Demographics	No	No	No	No	Yes	Yes	Yes	Yes
Business Char.	No	No	No	No	Yes	Yes	Yes	Yes
Mean of Y	0.513	0.513	0.644	0.644	0.513	0.513	0.644	0.644
R-sq.	0.002	0.001	0.006	0.000	0.300	0.301	0.447	0.448
Observations	1009	1009	613	613	1009	1009	613	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of risk variables is discussed in section A1.1.

Table A17: Active vs Non-filers - Single Tax Knowledge Questions

	(1) Q1 All	(2) Q1 Perp.	(3) Q2 All	(4) Q2 Perp.	(5) Q3 All	(6) Q3 Perp.	(7) Q4 All	(8) Q4 Perp.	(9) Q5 All	(10) Q5 Perp.
Panel A: without Controls										
Single question	0.22***	0.24***	0.07	0.02	0.16	0.14	0.20***	0.24***	0.16***	0.21***
	(0.04)	(0.05)	(0.10)	(0.12)	(0.12)	(0.14)	(0.04)	(0.04)	(0.03)	(0.04)
R-sq.	0.032	0.037	0.000	0.000	0.001	0.001	0.028	0.047	0.024	0.046
Panel B: with Controls										
Single question	0.12***	0.09**	0.01	-0.02	0.07	0.06	0.08**	0.06*	0.08**	0.09***
	(0.03)	(0.04)	(0.09)	(0.12)	(0.12)	(0.12)	(0.03)	(0.04)	(0.03)	(0.03)
Mean of Y	0.513	0.644	0.513	0.644	0.513	0.644	0.513	0.644	0.513	0.644
R-sq.	0.309	0.451	0.300	0.447	0.300	0.447	0.304	0.449	0.305	0.453
Observations	1009	613	1009	613	1009	613	1009	613	1009	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of the knowledge items is discussed in section 3.6.2.

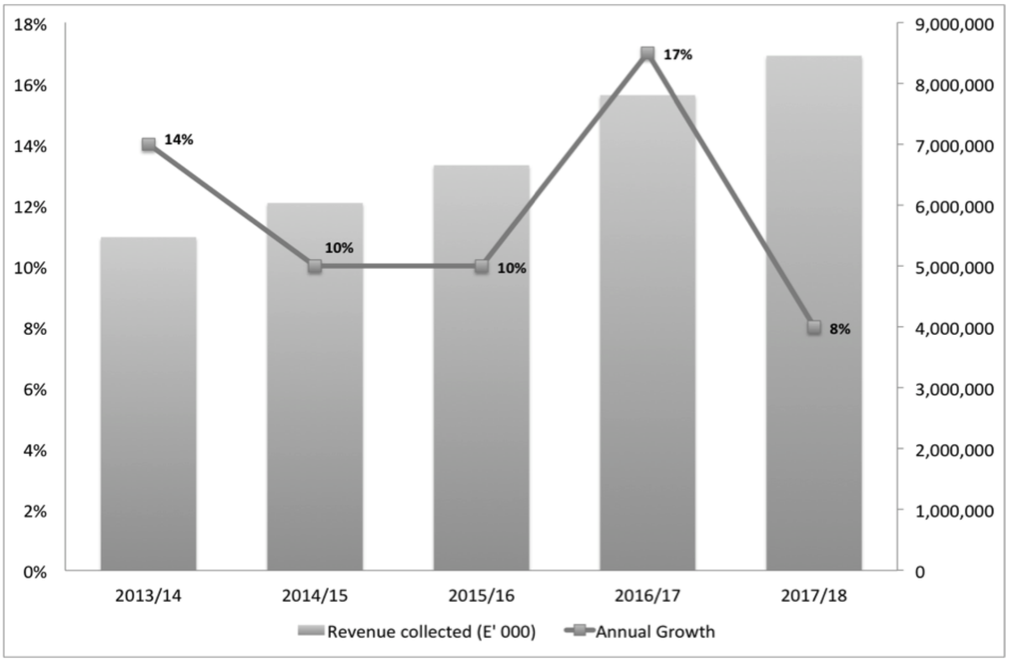
Table A18: Active vs Non-filers - Demographics and Business Characteristics

	(1) All	(2) Perpetuals	(3) All	(4) Perpetuals	(5) All	(6) Perpetuals
Female	-0.02	-0.03			0.01	-0.00
	(0.03)	(0.04)			(0.03)	(0.03)
Age group max=7	0.03**	0.04***			0.02	0.02
	(0.01)	(0.02)			(0.01)	(0.01)
Higher education	0.03	0.07*			0.01	0.03
	(0.03)	(0.04)			(0.03)	(0.03)
Swazi national	-0.11	-0.04			-0.09	-0.00
	(0.07)	(0.07)			(0.06)	(0.06)
Married	0.09**	0.07			0.07*	0.06
	(0.04)	(0.05)			(0.04)	(0.04)
Has tax accountant	0.19***	0.24***			0.16***	0.20***
	(0.03)	(0.04)			(0.03)	(0.04)
Time on tax > median	0.02	0.04			-0.01	0.00
	(0.04)	(0.04)			(0.03)	(0.04)
Bank account	0.16***	0.20***			0.05	0.07*
	(0.03)	(0.04)			(0.03)	(0.04)
Operative			0.38***	0.44***	0.31***	0.32***
			(0.03)	(0.04)	(0.03)	(0.05)
Years since registration			0.03***	0.04***	0.02***	0.04***
			(0.01)	(0.01)	(0.01)	(0.01)
Had a previous business			0.03	0.03	0.02	0.00
			(0.04)	(0.04)	(0.04)	(0.04)
Log(USD turnover)			0.04**	0.04**	0.01	0.01
			(0.02)	(0.02)	(0.02)	(0.02)
Hhohho			0.16***	0.24***	0.17***	0.23***
			(0.04)	(0.05)	(0.04)	(0.05)
Manzini			0.08**	0.17***	0.13***	0.20***
			(0.04)	(0.05)	(0.04)	(0.05)
Wholesale/retail trade			0.04	0.02	0.06**	0.05
			(0.03)	(0.03)	(0.03)	(0.03)
High Competition			0.05	0.06	0.05	0.05
			(0.03)	(0.04)	(0.03)	(0.04)
Less business			-0.03	-0.00	-0.03	-0.01
			(0.03)	(0.03)	(0.03)	(0.03)
Mean of Y	0.513	0.644	0.513	0.644	0.513	0.644
R-sq.	0.160	0.271	0.253	0.365	0.302	0.450
Observations	1009	613	1009	613	1009	613

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. *All* and *Perpetuals* refer to the total sample of taxpayers and the subsample of persistent taxpayers only, respectively. All coefficients are OLS estimates from a LPM. The operationalisaton of determinants is discussed in section 3.4.3.

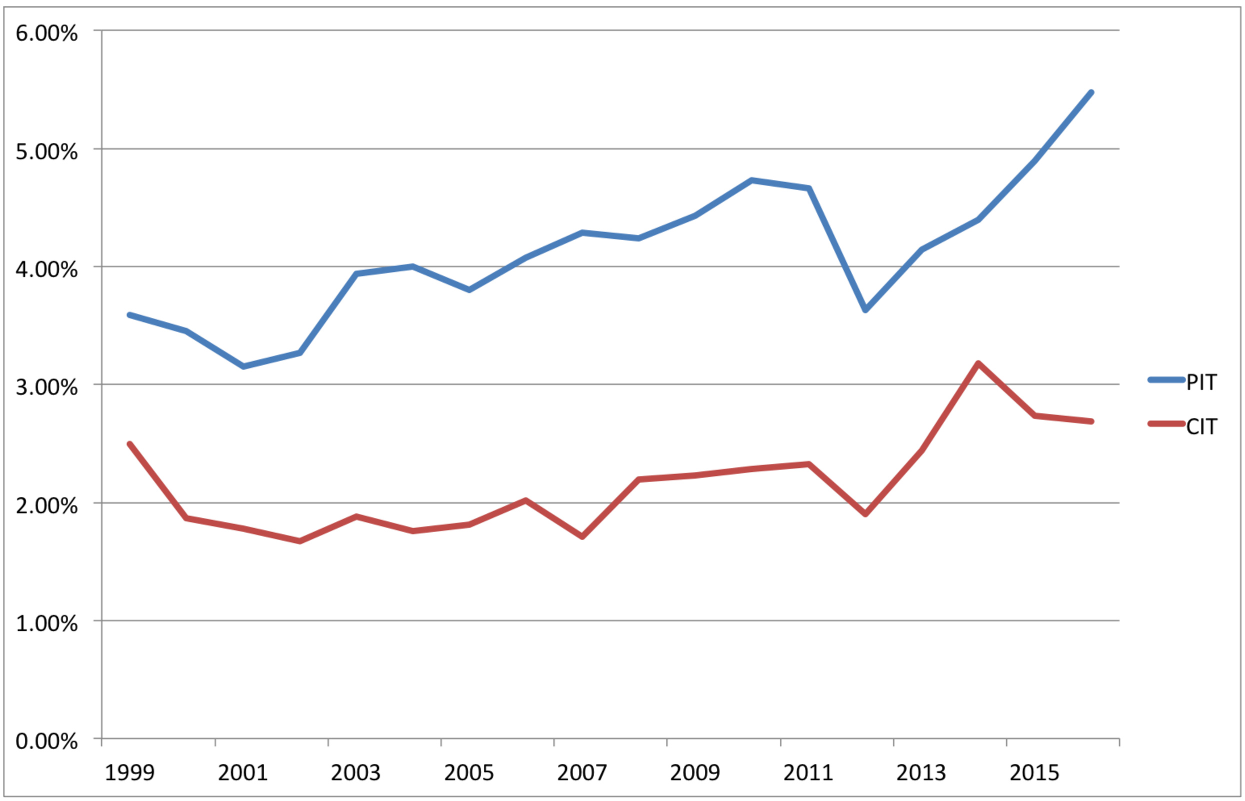
A5 Figures

Figure A1: Revenue collection in SZL '000 and revenue growth



Source: SRA (2018)

Figure A2: PIT vs CIT shares over GDP



Source: ICTD/UNU-WIDER (2020)

Figure A3: Survey sample location

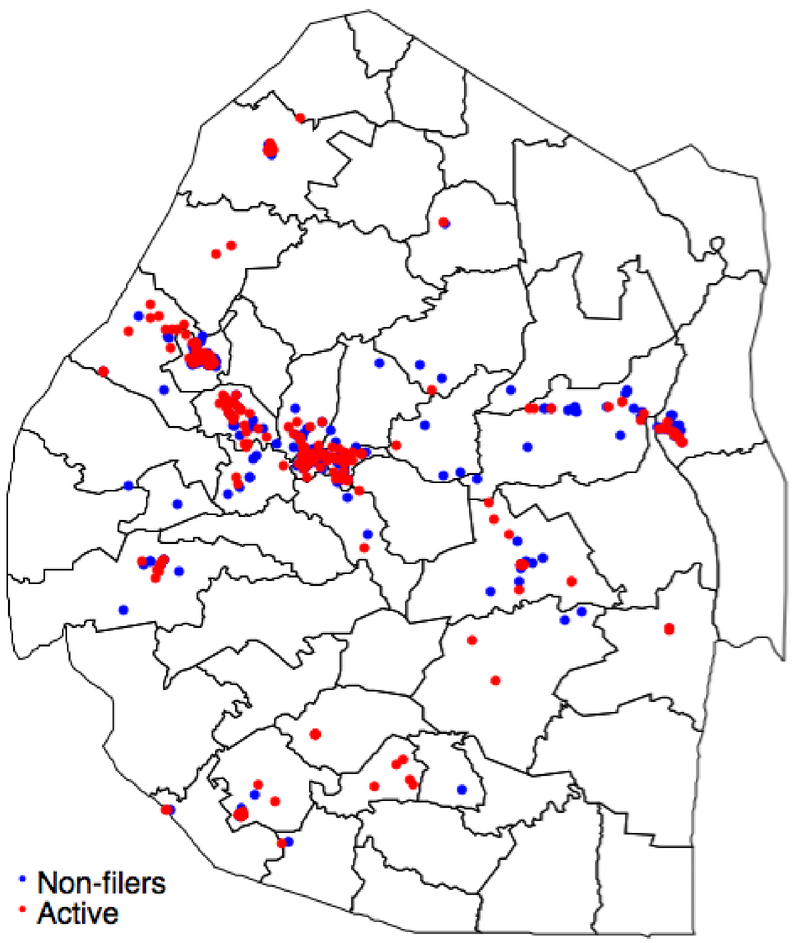


Figure A4: Self-reported monthly sales (USD) by study groups

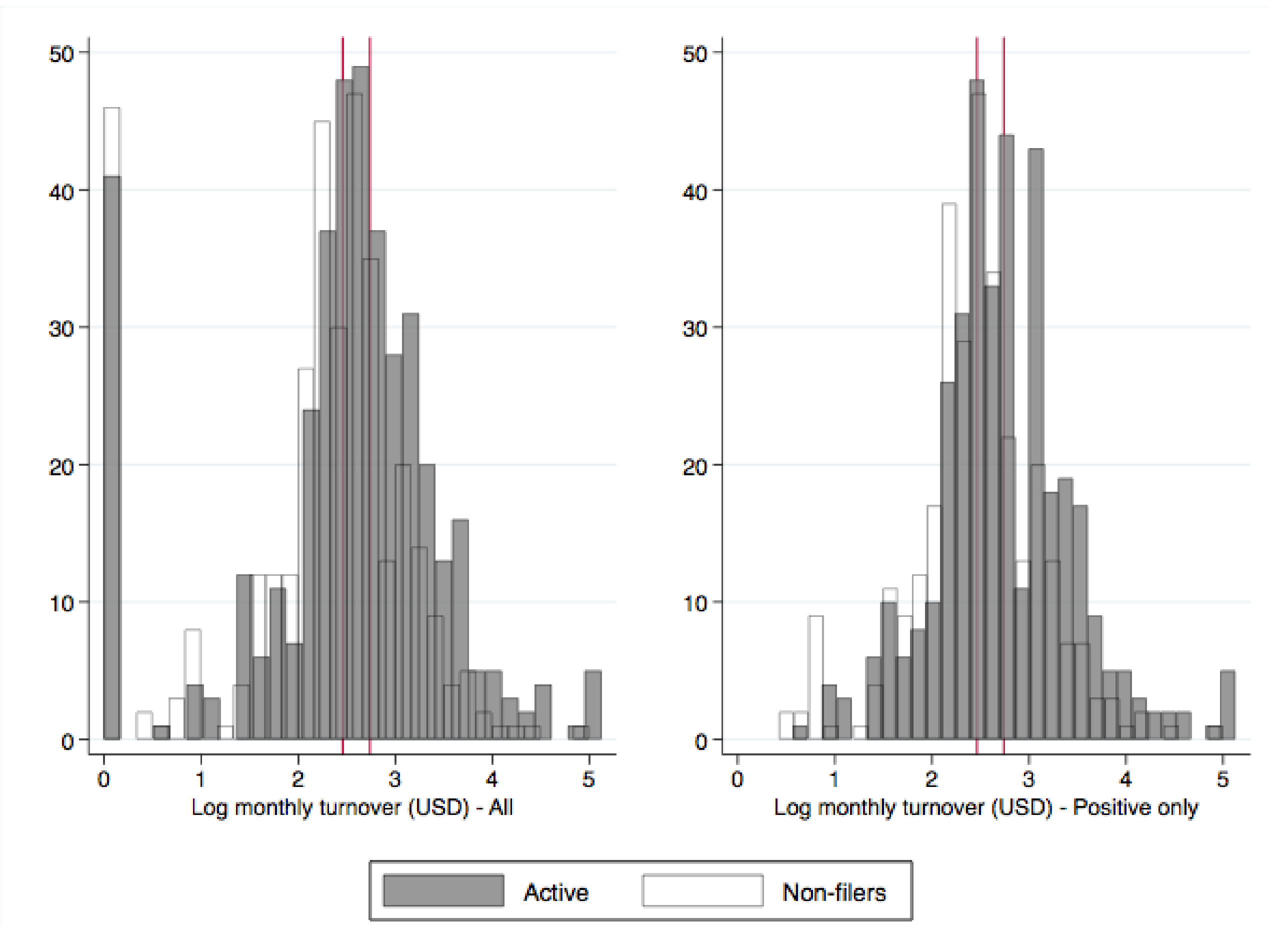


Figure A5: Reasons for registering with the authority

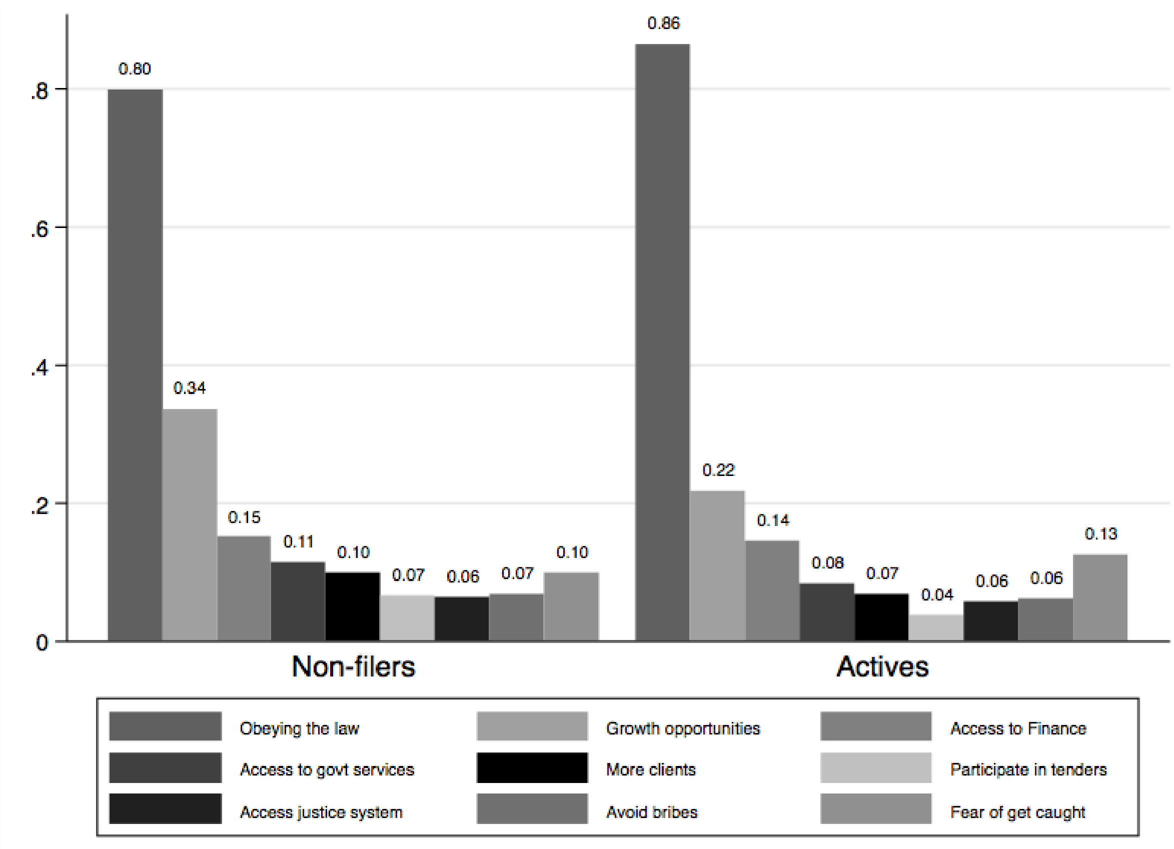


Figure A6: Perceived audit risk on respondent himself vs on other taxpayers

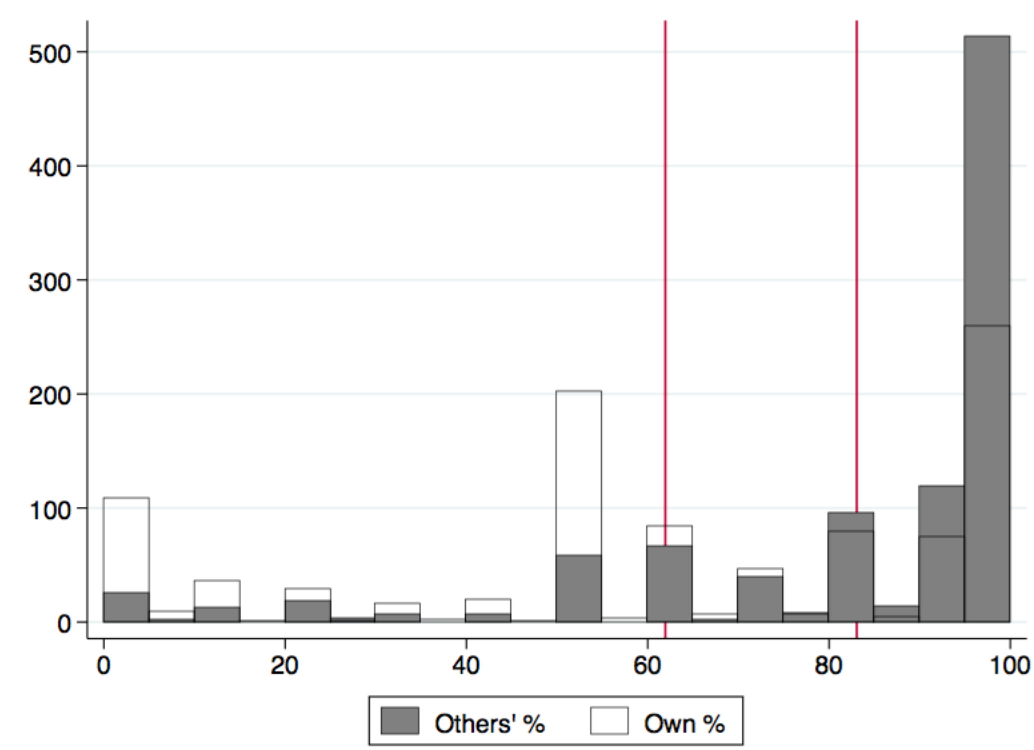


Figure A7: Determinants of Actively Filing a Return

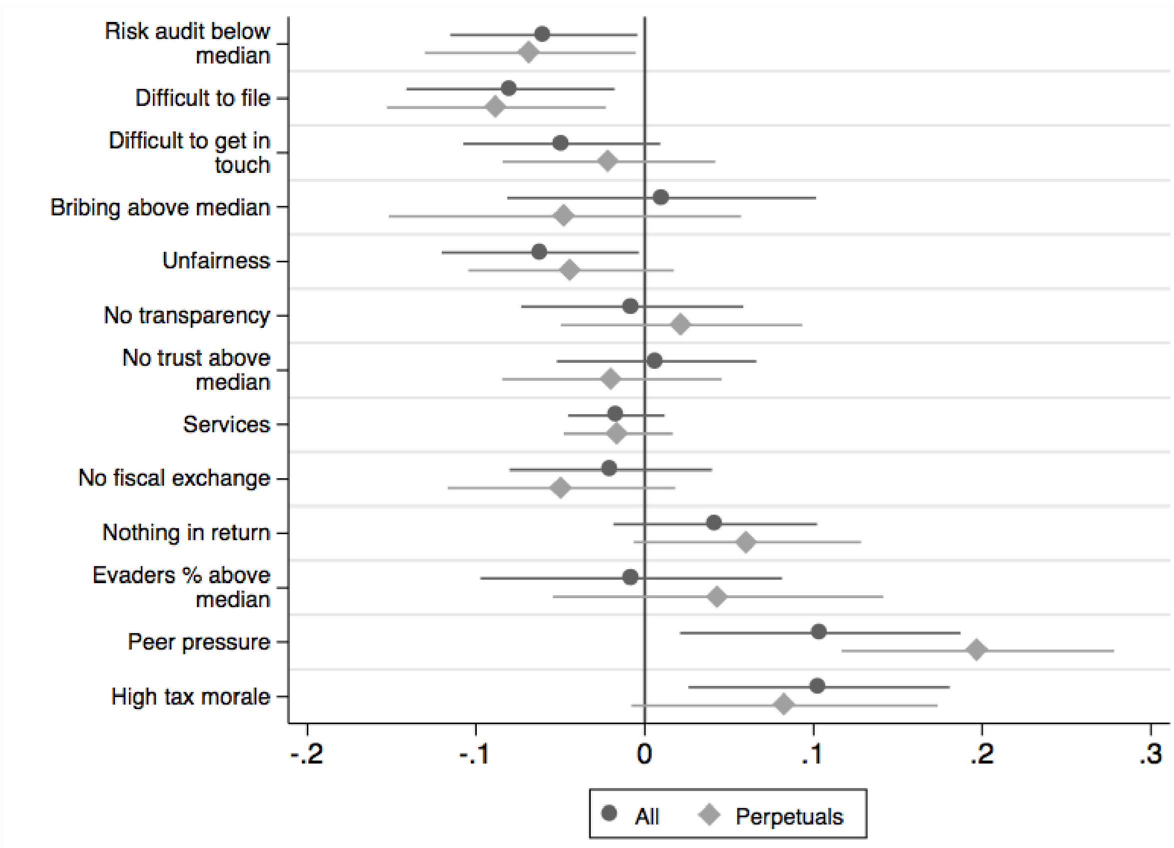


Figure A8: Determinants of Actual and Self-reported Compliance

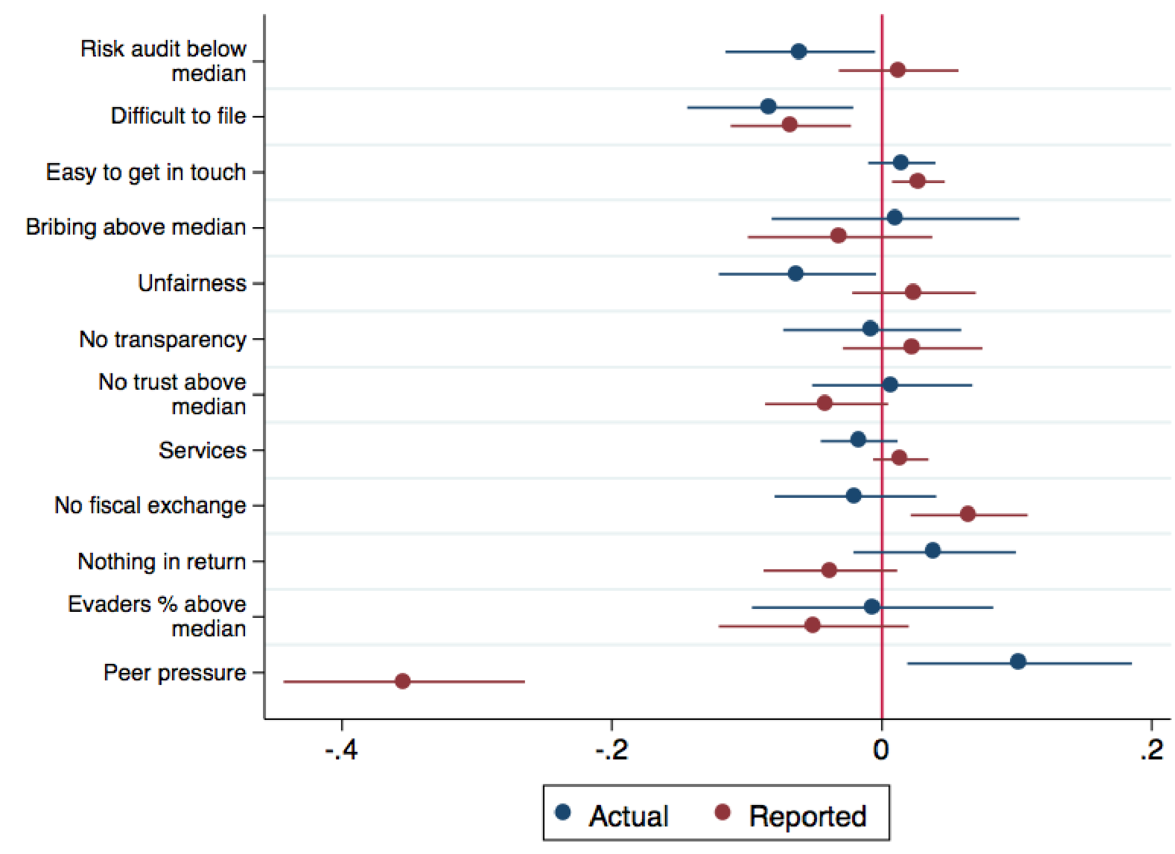
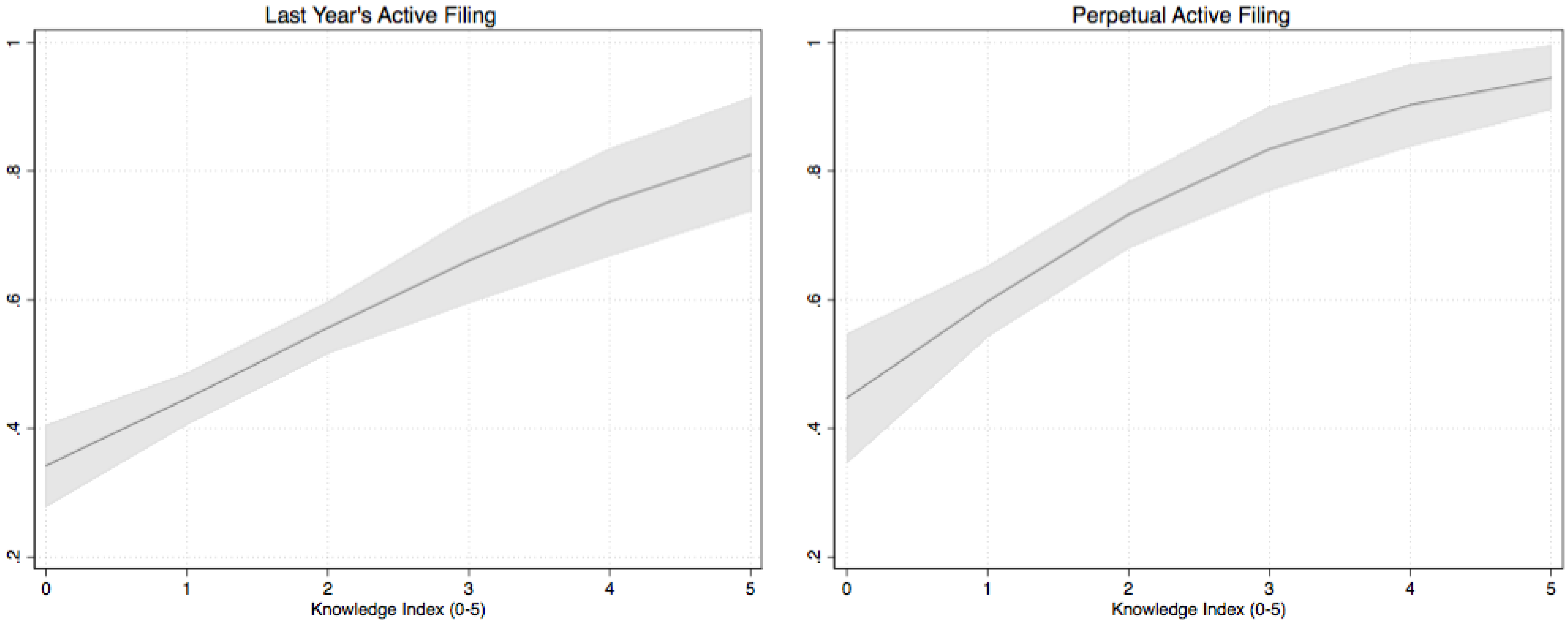


Figure A9: Actual tax knowledge and reported ease to file

Figure A10: Active Filing and Tax Knowledge, Predictive Margins



Chapter 4

How to Best Nudge Taxpayers? A Tailored Letter Experiment in Eswatini

Abstract

Very little is known about why taxpayers in sub-Saharan Africa (SSA) remit their taxes. In collaboration with the Eswatini Revenue Authority, this study implements a nation-wide randomised controlled trial nudging more than 20,000 income taxpayers with behaviourally-informed mailings. This study attempts to shed new light on which are the drivers of SSA taxpayers' compliance and how can these be leveraged by resource-constrained tax authorities. While the tax nudge literature has boomed in OECD countries and Latin America, only a handful of studies can be found in SSA; this paper contributes significantly to these. First, thanks to the wealth of administrative data available, this study is the first of its kind to target three different categories of taxpayers at the same time – non-filers, nil-filers and active – while most of the existing literature focusses on positive filers. Second, I tailor the content of letters to be specific to each taxpayer category. Third, I am able to target both companies and individuals and explore heterogeneity of results along a number of dimensions, including past filing behaviour. I find that non-filers significantly respond to the nudges, while nil and active filers do not. The best performing nudges build on the deterrence and taxpayer-assistance paradigms. Perverse responses are found from large companies. With the causal evidence produced, I am able to formulate policy recommendations on how to best target the complex ecosystem of income taxpayers.

4.1 Introduction

As one of the Sustainable Development Goals (SDGs), domestic revenue mobilisation is a fundamental objective of revenue authorities in low- and middle-income countries.[1] In sub-Saharan Africa (SSA), while revenue authorities are progressively increasing tax collections with respect to other low-income countries (Moore et al., 2018), the tax-to-GDP ratio has risen by only 2 to 3 percentage points of GDP in the past two decades (Akitoby et al., 2019).

As one of the key challenges to revenue collection, taxpayer compliance is far from being optimal. Some simple statistics clearly reflect the gravity of the problem. In Eswatini, the country under study, every year about half of income taxpayers fail to file (non-filers), while, among those that submit a return, about a quarter report zero income and zero tax (nil-filers). In addition, taxpayers are quite persistent over time in their filing behaviour. For example, 18 per cent and 10 per cent of the population of registered taxpayers in Eswatini either never filed or perpetually filed nil, respectively, since their registration with the authority.

These figures are economically relevant given that income taxes represented 35 pr cent of total tax revenue in 2018 (ICTD/UNU-WIDER, 2020). Similarly alarming figures are found in other SSA countries like Rwanda, Uganda, Malawi, Ethiopia and Nigeria (Chapter 1). In terms of their detrimental repercussions on domestic revenue mobilisation, non- and nil-filing (as long as the latter entails evasion) pose immediate challenges to already budget-constrained revenue administrations. At the same time, also active filers can be under-declaring their liabilities. The negative consequences of these filing decisions ultimately create economic inefficiencies and horizontal inequalities. More broadly, when tax evasion is involved, these decisions generate unfairness, lower the moral fibre of a society and eventually delegitimise the government.

In order to enhance revenue collection, revenue authorities usually implement traditional enforcement strategies, such as audits and penalties.[2] These strategies are however

[1] According to the International Monetary Fund, on average SSA will need additional resources amounting to 19% of GDP to finance the SDGs in education, health, roads, electricity, and water by 2030 (IMF, 2019).

[2] Additional, more sophisticated enforcement strategies, which are less likely to be implemented in revenue

costly, especially for revenue authorities in SSA who are historically constrained by limited budget and enforcement capacity (Besley and Persson, 2013; Pomeranz and Vila-Belda, 2019). In a context in which State legitimacy is already low as in SSA (Isbell, 2017), a deterrence-based compliance strategy may be detrimental – because it could reinforce distrust and taxpayer resistance (Fjeldstad and Semboja, 2001), and delegitimise the revenue administration even more. For these reasons, it has been argued that the optimal mix of tax instruments can diverge from what traditional public finance theory prescribes when tax capacity and State legitimacy are limited (Best et al., 2015).

More flexible and potentially highly cost-effective interventions, such as tax nudges, can represent a feasible alternative to foster voluntary compliance.[3] Tax nudges are behaviourally-informed tax compliance interventions that respect the taxpayer's freedom of choice and leave economic incentives intact while attempting to improve taxpayer's behaviour. Tax nudges, by building on the theoretical formulations of behavioural economics, also represent a theory-grounded tool to shed light on the pecuniary and non-pecuniary factors driving compliance and promptly suggest which of them governments should prime for revenue mobilisation. While tax nudges have been implemented in high-income countries for decades, limited evidence has been produced from low- and middle-income countries – and even less so in SSA (see section 4.2). It is likely that the drivers of voluntary compliance are likely to be different in SSA for a number of reasons, ranging from limited taxation capacity of revenue bodies (Mascagni, 2018; Moore et al., 2018) to low level and quality of public services (D'Arcy, 2011; Bodea and Lebas, 2016; Blimpo et al., 2018), from distrust into the State processes and tax officials (Bratton and Gyimah-Boadi, 2016; Isbell, 2017; Pirttilä, 2017), to poor knowledge of the tax system (Fjeldstad et al., 2012; Aiko and Logan, 2014; Isbell, 2017; Mascagni et al., 2019). However, very little is known on why African taxpayers comply with or evade their taxes.

Against this background, this study attempts to fill the gap in knowledge answering to the first-order question on which are the key drivers of compliance. As one main contribution, this study stems from a close research collaboration with the Eswatini Revenue

authorities with lower than optimal administrative capacity, refer to third-party information reporting, changes in the remittance regime, shaming through public disclosure, and take-up of benefits.

[3]Thaler and Sunstein (2008) define nudges as "choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives".

Authority (SRA), which provided access to a wealth of administrative data, assisted during the challenging implementation of the experiment and fully embraced the goals of the study. I am thus able to provide robust evidence from a lower-middle-income country in SSA, Eswatini, which has not been studied before (see section 4.3.1) and implement a field experiment in a region, Southern Africa, where there has not been a great progress in this sense (Ebrahim et al., 2019). I attempt to address this question by implementing a nation-wide Randomised Controlled Trial (RCT) that sent behaviourally-informed experiment letters to more than 20,000 income taxpayers. Tax compliance decisions of income taxpayers are particularly interesting from both a theoretical and practical point of view, since their tax evasion is particularly difficult to uncover – especially in contexts with lower tax capacity of the revenue authority – given that they self-report their income and have strong economic incentives to underreport it to reduce tax liabilities (Allingham and Sandmo, 1972; Slemrod and Yitzhaki, 2002; Sandmo, 2005). This challenge becomes even more delicate for revenue mobilisation purposes, since income tax (net of PAYE) usually represents the second largest contributor, after VAT, to tax revenue both in Eswatini and in SSA (ATAF, 2017).

The second contribution of this paper is that, departing from the existing evidence (see section 4.2), I tailor the content of letters to be specific to taxpayers' filing behaviour, thus considering three filing categories: non-filing – failure to file a return –, nil-filing – filing a return with zero turnover, taxable income and tax liability – and positive filing, plus the additional category of income filers reporting divergent information in their VAT returns.

The reasoning behind targeting is that, for instance, a nudge teaching how to file is more likely to benefit non-filers than the other two filing categories who, presumably, already know how to submit a return – so compliance costs nudges are only sent to non-filers. In the same fashion, the threat message displays only the relevant articles of the tax law that apply to the specific filing category, so to make the threat more salient and realistic. It is also interesting to understand whether the same type of soft nudge, such as the one on fiscal reciprocity, has a different impact on different taxpayers, such as non-filers and active, or, as with non- and nil-filers, whether a more service-oriented communication informing about the steps for deregistering from the system (an option which, in theory, is optimal for both categories) spurs similar or different reactions. To the best of my knowledge, this is the first time these three filing categories have been targeted in the same study. By targeting these categories in the same RCT, I test different theoretical motivations for each filing behaviour and present a broader, more complex picture of tax compliance than the existing literature, which mostly focusses on positive filers.[4] Given the negative repercussions that incomplete compliance has on revenue collection in SSA, both researchers and tax administrators from the African continent need to understand what drives these different behaviours.

A third contribution is given by the fact that, consistently, I consider filing outcomes that are specific to each category and study both the extensive (filing a return) and intensive (remitting a given tax amount) margin of compliance. Furthermore, I am able to target both companies and individuals and explore their different reactions to nudges, while most similar studies focus on either incorporated agents or individual traders (see section 4.2). Relatedly, thanks to the nationwide nature of the RCT, I am also capable of exploring the heterogeneity of results along different dimensions; some of these, such as previous filing behaviour, are poorly documented in the literature and highly policy-relevant.

The experiment letters represent a simplified, one-page version of standard SRA communication, and differ only in terms of the key message printed in a box at the centre of the page, and the envelope colour – this is different from ordinary SRA mailings to make the letter more salient (BIT, 2012). Following a stratified randomisation algorithm, non-filers are randomly assigned to five different treatment arms, that is deterrence, assistance on filing, assistance on deregistering, reciprocity and social norms. The smaller group of nil-filers is allocated to deterrence and assistance on deregistering arms. Due to the small sample size of nil-filers, I needed to make a choice on which treatment arms to test. Aiming at being consistent with the little experimental evidence existing on nil-filers (Mascagni et al., 2020), I opted for deterrence and de-registration as the main possible drivers of behavioural change for nil-filers. Active taxpayers are nudged with deterrence and reciprocity messages while taxpayers with VAT discrepancies are allocated to a deterrence letter highlighting the amount of the discrepancy.[5] For each category, a randomly created control group provided counterfactual outcomes.

[4] There are a few notable exceptions, summarised in section 4.2.

[5] The deterrent letters sent to different categories varied. Only the penalties and fines associated with the particular behaviour (such as failure to file or false statement) of the taxpayer in one category is responsible for are made salient. More details in section 4.4.

Using detailed administrative data on anonymised income tax filing for tax year 2019, as well on VAT, PAYE and previous years' income tax filing,[6] with the support of survey data from 1,000 sole traders collected in Chapter 3, I present four sets of results. First, the best performing treatments are those targeting non-filers. Compared to a very low filing rate in the control group (6.9%), which is suggestive of the persistent nature of non-filing over time, I find that filing rates increase the most, by 2.6 percentage points or a sizeable 38 per cent, when a deterrent tone is used. Similarly, more service-oriented nudges reducing the compliance costs of filing and providing information on deregistration, as well as moral appeals, boost the probability of filing by 1.3 percentage points or 20 per cent over the control group. Among the latter, the social norm nudge, despite referring to a rather low descriptive norm (66% filing rate from peers) significantly increases filing. These results are remarkably in line with what was found in a meta-analysis of tax compliance experiments run by Atinyan and Asatryan (2019), who find that deterrence nudges are more effective in improving filing rates that non-deterrence ones by an average of 1.5-2.5 percentage points.

In addition, thanks to the provision of letter delivery reports from the national post office, I am able to measure the actual treatment effect on the treated. When partial (45%) letter pick-up is taken into account, estimates of impact increase accordingly and keep their statistical significance, with (i) the deterrence letter doubling filing rates, (ii) the service-oriented arms increasing filing by 60 per cent, and (iii) the moral appeals improving filing by 156 per cent over the control group. These results may seem extremely large but, as explained above, are to be compared with a very low control group average, meaning that the majority of nudged non-filers still persist in the same behaviour. Importantly, non-filers are also significantly filing more past returns. In terms of immediate revenue considerations, nudged non-filers are not remitting more tax than controls, even if individuals are actually paying more taxes. In line with similar findings discussed below, the nudge backfires for non-filing companies – who remit less tax. Tax remitted was not the main outcome for non-filers in my pre-specified plan, as I gave priority to the extent to which they become visible to the authority. Non-tangible benefits from increased filing are achieved, which go beyond quantifiable extra revenue.

[6] To the best of my knowledge, this was the first time tax data was used for research in Eswatini, adding to the existing pool of tax studies in Southern Africa which mostly come from the analysis of administrative data from South Africa (Ebrahim et al., 2019).

Unfortunately the experiment letters do not encourage nil-filers to start filing positive tax, or, active taxpayers to increase their liabilities. If anything, the fiscal exchange nudge actually backfires for active companies, who reduce their tax due. This could be due to the fact that, according to survey data gathered in Chapter 3, active taxpayers have a lower average satisfaction with six different public services than non-filers. However, it is worth mentioning that letters actually increase the probability of active taxpayers filing positive tax, and much more significantly for companies – suggesting that they particularly shape the response of those taxpayers who are about to turn to nil-filing.[7] Once nudged, potential nil-filing companies eventually file positive, even if they do not remit more tax compared to the control group.

This last piece of evidence connects with a second set of results, referring to the significant differences in response between companies and individuals. For instance, non-filing companies are more likely to respond (41% increase) to the threat nudge than individuals (25%), who in turn react more when taught how to file. At the same time, individual non-filers are more prone to increase tax due, once filing, than companies – who actually decrease their liabilities. This evidence could suggest that companies are more likely to enjoy the service of tax accountants and therefore the lack of tax knowledge is not a relevant constraint for them. To support this hypothesis, I provide additional evidence on the probability of filing on time, according to which individual taxpayers, not companies, are more likely to file on time when reminded about the deadline through the compliance costs nudge. Another explanation for the negative response of companies could be companies' dissatisfaction with a taxpayer service deemed of inadequate quality – a one-page to-do list – which would have increased the annoyance of the recipients and hence their negative reaction.[8]

Third, and relatedly, allowing for heterogeneity provides a more complex picture of taxpayer responses, making the argument for a more tailored enforcement strategy. I explore different dimensions, thanks also to the wealth of administrative data available: (i) non-persistent taxpayers, i.e. those that are more unstable in their filing behaviour and

[7] As shown in Santoro and Mdluli (2019), nil-filing is more likely to take place among companies than individuals.

[8] An opposite explanation could be that non-filing companies in the control group mistakenly remit too much tax to the authority. However, it is not clear why this is not the case for control group individuals.

shift between the three filing categories over time, are more likely to file and even increase their tax due with respect to chronic/persistent taxpayers, (ii) newly registered taxpayers are less likely to respond than older ones, probably due to the fact that a one-page letter is not adequately incisive in shaping their filing decisions; (iii) rural taxpayers respond more to deterrent letters than to softer-toned ones, which in turn perform better in urban settings; and (iv) business size also matters – the largest active taxpayers are more likely to reduce tax amounts when nudged.

Fourth, it is worth stressing that nudges could lead to perverse responses. As a mechanism behind the back-firing effect of the fiscal exchange letter, active taxpayers significantly report more income and, at the same time, increase expenses and deductions. This compensating response implies that the final tax due does not change or even decreases. Similar responses have been observed in the literature (Ariel, 2012; Carrillo et al., 2017; Slemrod et al., 2017; Mascagni et al., 2017) and further attention should be devoted on how to avoid them.

Despite poor impact at the intensive margin, the overall cost-benefit ratio of the trial is of 1:11 and extra revenue associated with the experiment amounts to US$0.2 million. As explained in section 4.7, the extra revenue comes mostly from current and past returns of non-filers, as well as from the extra positive tax raised from those active payers who are prevented from nil-filing.

The results of this study have both practical and theoretical relevance. On a practical level, I attempt to inform the SRA's, and other similar revenue administrations', communication and enforcement strategies with robust causal evidence and policy recommendations grounded in the local reality (see section 4.7). Following the *economists as plumbers* framework of Duflo (2017), this study builds on collaboration with local policymakers to evaluate specific details of actual policies, and highlights margins for policy improvement that diverge from textbook models of tax compliance. In doing so, I broaden the current understanding of both researchers and tax administrators on two widely under-researched taxpayer profiles – non- and nil-filers. As summarised in section 4.2, these two profiles are widespread in high-income countries (Erard et al., 2018; Meiselman, 2018) and Latin America (Kettle et al., 2016; Brockmeyer et al., 2019), and have only recently been documented in SSA (Mascagni and Mengistu, 2016; Almunia et al., 2017; Ligomeka, 2019a; Mascagni et al., 2020). Further practical considerations arise from the very low letter take-up of individual non-filers who, when exposed to the treatment, are extremely responsive to the different nudges and therefore might need to be reached with alternative methods than mailings, such as SMS or phone-calls – but rigorous research on other delivery methods is needed to prove their efficiency.

On a theoretical level, I empirically test the validity of predictions on the drivers of compliance in a largely under-studied African country. The conceptual framework elaborated in Prichard et al. (2019) is directly tested in the field. According to the framework, the combination of three measures – enforcement, facilitation, and trust – is the way forward to encourage quasi-voluntary compliance in developing countries, in addition to generating political support for reform and building stronger fiscal contracts. The results presented in this paper, albeit focusing only on tax compliance, show a mixed picture of the framework's effectiveness. While the neoclassical approach of Allingham and Sandmo (1972), based on pecuniary motives and deterrence, seems to be working for non-filers and, partially, for active taxpayers (for the latter, on the probability of positive filing only), it is also ineffective in increasing the tax remitted by nil-filers and actives. At the same time, a service oriented approach, originally formulated in Alm et al. (2010) and empirically tested mostly with lab experiments,[9] is highly impactful, especially for individuals who arguably face larger costs than firms in dealing with the tax system. Lastly, recurring to a non-pecuniary, trust-based approach (Luttmer and Singhal, 2014) improves compliance of non-filers but backfires with active taxpayers. The latter effect is probably due to higher dissatisfaction of active taxpayers in how tax revenue is spent, as documented in Chapter 3, and questions the effectiveness of the fiscal exchange theory in contexts of misuse of public funds and inadequate quality of public services. Future avenues of research should continue testing the conceptual framework of compliance in similar developing settings.

The paper is organised as follows. Section 4.2 provides an overview of the existing evidence on tax nudges from both developed and developing countries, with a focus on SSA. Section 4.3 describes the institutional context, while the field experiment is addressed in detail in Section 4.4. Section 4.5 presents the results. Section 4.6 explores the mechanisms

[9]See Alm et al. (1992), Kosonen and Ropponen (2015), Vossler and McKee (2017), McKee et al. (2018). Field experiments are much less common. See Chetty and Saez (2013) and Mascagni et al. (2019).

underlying the main results and the final section concludes.

4.2 Tax nudges: what works and what does not?

4.2.1 Tax nudges in Western countries and Latin America

Natural field experiments on taxpayer communication first appeared in high-income countries. Apart from early applications (Schwartz and Orleans, 1967), modern tax experiments build on the seminal work of Blumenthal et al. (2001) and Slemrod et al. (2001) in Minnesota, USA. The authors nudge as many as 60,000 individual taxpayers and show both that the threat of audit increases compliance, even if for small taxpayers only, and that public services and descriptive norms letters have no effect. Soon after these initial trials, a large number of nudge experiments blossomed in Europe[10], Australia[11] and USA as well.[12] Nudge interventions have become mainstream in HICs thanks to the increased collaboration between researchers and revenue authorities, which granted access to their administrative data and invested a great amount of internal resources in the implementation of the nudges.[13] Overall, these studies show that enforcement is unambiguously effective in promoting compliance across different settings, with the exception of Ariel (2012), while the evidence from moral appeals and civic duty is rather inconclusive.[14] Also, reminders per se seem to be effective. For a more complete review of experiments in HICs, the interested reader could see Hallsworth (2014).

Outside of HICs, tax nudges have been increasingly implemented in the last decade, especially so in Latin America. These studies make an important contributions to the

[10] See Torgler (2004b), Hasseldine et al. (2007), Kleven et al. (2011), Fellner et al. (2013), Bott et al. (2014), Hernandez et al. (2017) and De Neve et al. (2019)

[11] See Wenzel and Taylor (2004) and Biddle et al. (2017).

[12] See Chirico et al. (2016), Perez-Truglia and Troiano (2016) and Meiselman (2018).

[13] In the most advanced settings, specific nudge units have been launched within institutions, such as the Behavioural Insights Team (BIT) in the UK and the White House's Social and Behavioral Science Team (SBST) in the USA. According to DellaVigna and Linos (2020), there are more than 200 such units globally.

[14] A noteworthy exception is the Hallsworth et al. (2017) study on norms nudges, in which sizeable impacts on tax payments can be find.

literature as they both manipulate message content in a novel way and also offer a more nuanced picture of nudges' effectiveness, when implemented in a less than optimal institutional context as the one in HICs. For instance, in Peru, disclosing information on the level of compliance of the subject reference group has a large positive impact on compliance with property tax (Del Carpio, 2014). In the same fashion, Kettle et al. (2016) run a large (N=43,387) nationwide RCT in Guatemala targeting non-filing income taxpayers with four different behaviourally informed tax letters. The authors show that the best performing nudges are a deterrent message framing non-filing as a deliberate choice and a social norms message. The latter surprisingly tripled tax receipts by referring to the (rather low share of) 64.5 percent of taxpayers that had already paid this tax and invited non-compliers to join the majority.[15] In other instances nudges do not boost compliance but incentivise negative, albeit fully rational, responses. In Ecuador, Carrillo et al. (2017) emailed about 10,000 corporate income tax (CIT) payers on the extent of discrepancies in their income tax return with third-party sources. While most firms failed to respond, those who increased revenue also reported higher costs on less verifiable items of the tax returns, ending up in minimally increasing tax liability.

A number of interesting lessons emerge from Latin America, which are also likely to be valid in SSA.[16] First, it results that deterrent nudges can be limited in low-enforcement environments, or, if anything, work only in specific contexts. In most lower- and middle-income countries, revenue authorities have limited budget resources to dedicate to audits with the risk of nudging taxpayers with threats that cannot be backed by credible enforcement (Carrillo et al., 2017). In others, the mere fact of being contacted by the tax authority by written communication, regardless of the content, has a positive impact on compliance (Ortega and Sanguineti, 2013). For this reason, in-person visits, putting tax collectors in direct contact with taxpayers, despite being costly, are also more effective (Ortega and

[15] The descriptive norm communicated in Kettle et al. (2016) is slightly lower than the one tested in this study (see section 4.4.2).

[16] Outside Latin America, Chetty et al. (2014) nudged more than 23,000 VAT firms in Bangladesh with 8 different version of letters, associated with different combinations of recognition cards, information disclosure to peer groups and descriptive information on compliance levels (the social norm nudge in this study). They show how impact differs by the ex-ante compliance levels, with an increase in tax payments of 17% in high-compliance clusters and zero effect in low-compliance clusters for those firms exposed to the possibility that information on their compliance could have been shared with their peers.

Scartascini, 2016a). Furthermore, explicitly stating the amount of fines according to the tax code reinforces impact (Castro and Scartascini, 2013).[17] Deterrent signals can also act as *scarecrows* and exploit inconsistencies in utility maximisation, especially in small taxpayers. In Uruguay, Bergolo et al. (2019) show that the threat of audit generates fear and induce probability neglect (Sunstein, 2003).[18] Second, it is also true that impacts can last over time, as shown by Kettle et al. (2016) in Guatemala and Brockmeyer et al. (2019) in Costa Rica, where enforcement emails increase tax payments two years after the trial.[19] Third, average effects may mask differences across individuals (Castro and Scartascini, 2013). Heterogenous impacts are highly common in these studies and dimensions such as past filing behaviour, income levels and peers' compliance rates tend to differently influence taxpayers' responses.

4.2.2 Tax nudges in sub-Saharan Africa

Despite having become the norm in Europe and North America and being increasingly tested in Latin America, evidence on tax nudges from SSA is almost non-existent. So far, only a handful of studies have rigorously tested tax nudges in the region and are therefore worth mentioning.[20]

Shimeles et al. (2017) study the impact on profit tax paid of letters delivered to 3,120 businesses in Addis Ababa, Ethiopia, testing both a threat and a more persuasive nudge. While the threat emphasises the risk of being audited and the penalty regime, the persuasion letter is patriotic in tone and lists flagship projects funded by taxes, in much of the same way as the fiscal exchange nudge in this study (see section 4.4.2). Increases in tax paid are remarkably large, 38 per cent for the threat and 32 per cent for the persuasion let-

[17] Both these aspects are tested in this trial as well, as discussed in section 4.4.2.

[18] Loewenstein et al. (2001) elaborate that individuals experiencing fear react quickly, intuitively and sub-optimally and thus neglect the underlying probabilities of a given event.

[19] It is not clear why impacts are sustained over the medium-term in some specific contexts, while, in the majority of cases, nudges are short-lived. One possible explanation can be due to the extent of credibility associated with the nudge, which is based on trust towards the revenue authority and is highly country-specific. Further research should be devoted to this aspect.

[20] It is also true that similar tax experiments are currently carried out in SSA, but to the best of my knowledge the results are not yet publicly available.

ter, and highly significant. A major explanation for these large impacts may be that letters are hand-delivered by tax officials, thus augmenting taxpayers' perception of being under the agency's radar and, consequently, their response. However, hand-delivering letters can be expensive and not feasible when the nudge strategy needs to be scaled up.

In another original field experiment run in Rwanda, Mascagni et al. (2017) test the impact on tax due of different message contents (deterrence, fiscal exchange or reminder) and delivery methods (letter, email or SMS), for a total of nine treatments. The authors target a total of 13,000 CIT and PIT payers located in the capital city Kigali. Results show that non-traditional channels of communication, such as SMSs and emails, are more effective than physical letters, while softer types of nudges, i.e. reminders and fiscal exchange, outperform deterrence. At the same time, smaller taxpayer are more responsive to a deterrent tone than larger ones.[21]

In a follow-up study, Mascagni et al. (2020) explicitly target about 7,000 income tax nil-filers by exploring two main reasons behind this important but understudied behaviour, much in line with what is studied in this trial: (i) tax evasion, which is tested with a threat SMS, and (ii) the need to deregistering from the system, which is tested with a more service-oriented SMS. An SMS reminding about the filing deadline is also added. The authors show that nudged nil-filers are more likely (2.3%) to switch to positive filing after receiving the deterrence message, even if significance dissipates after controls are added. Likewise, informing about the deregistration procedure has a significant effect on the probability of deregistering, even if small in magnitude (just under 1%). Importantly, the reminder SMS significantly reduces nil-filing and this could hint to the fact that reminders can be seen as a form of friendly deterrence.

Against the limited evidence on tax nudges in SSA, this paper contributes significantly in at least three ways. First, thanks also to the limited size of the country, all income taxpayers in Eswatini have been involved in the trial, or about 40,000 units, while studies in Ethiopia and Rwanda focus on capital cities only. Targeting the entire population of income taxpayers, this study allows for a more in-depth study of taxpayers' response in urban and rural areas, for which it is reasonable to expect different levels of compliance rates due to a number of factors, including higher tax knowledge in more sophisticated and

[21] This particular dimension of heterogeneity is explored in section 4.6.

urban settings as well as limited reach of the tax agency and lower audit probability in more remote areas. I explore variation across additional dimensions as well, such as income deciles, demographics and compliance history (see section 4.6).

Second, against much of the literature which focuses on a specific type of taxpayer (active/non-filer, individual/companies only) the reach of nudges is improved and three different filing categories are targeted in the same trial, following the rationale delineated in section 4.1. This approach goes in contrast with the studies discussed above, all of which primarily focussed on measuring the impact of different types of nudges on the same filing category. The policy relevance of this aspect of the study is important, since the SRA and other SSA tax authorities can better learn to direct their limited resources to those taxpayers who are more likely to respond. The nation-wide feature of this study discussed above enhances the generalisability of the results in the eyes of the tax agency, who will benefit from a comprehensive, albeit multifaceted, picture of taxpayers in Eswatini.

Thirdly, given its relevance in Eswatini (see section 4.3), the persistence of a filing behaviour over time is explicitly included in the analysis and the subgroup of *perpetuals* is identified, as opposed to non-perpetuals – taxpayers moving across the three filing categories over time. The study of perpetuals is important as past behaviour is likely to be a key determinant of actual compliance (Dunning et al., 2017; Tourek, 2020). Survey data from Eswatini also shows that perpetual non-filers are more rooted in their negative attitudes and perceptions than non-perpetual ones (Chapter 3). To the best of my knowledge, this dimension has not been explored in the literature thus far.

4.3 Anatomy of tax compliance in Eswatini

4.3.1 Institutional context

As already mentioned in section 3.3 of Chapter 3, the Kingdom of Eswatini is a lower-middle-income country in Southern Africa with an income per capita of $3,243 in PPP (2017).[22] Economic growth is estimated to have slightly risen to 2.4 per cent in 2018 from

[22]Source: World Bank World Development Indicators.

2 per cent in 2017, although growing fiscal challenges resulted in a projected growth rate of just 1.3 per cent for 2019 (World Bank, 2018).[23] The country is highly dependent on South Africa, which provides around 85 per cent of its imports and a market for about 60 per cent of exports. Its tax-to-GDP ratio (14.7%) is slightly below the average in SSA (15.0%), while being substantially lower than the 34.2 per cent average in OECD countries and about half of that of Southern Africa.[24]

The Eswatini Revenue Authority (SRA) is a semi-autonomous institution established by the Revenue Authority Act in 2008, officially taking over the function of revenue collection on 1 January 2011. For a summary of the main direct and indirect taxes collected in the country see section 3.3 of Chapter 3. This experiment targets income taxpayers registered for corporate income tax (CIT)[25] and personal income tax (PIT).[26] More specifically, PIT is remitted by three main groups: (i) sole traders (37.5% of all PIT registered payers), (ii) directors of companies (21.5%) and (iii) non-business employees taxed at source through PAYE (41%). In this study, I exclude PAYE payers, for the inherently different tax remitting mechanism, since employees do not self-report their earnings, and compliance factors shaping it. Removing PAYE, it results that income taxes represent a sizeable 35 per cent of total tax revenue in 2018 (ICTD/UNU-WIDER, 2020). As explained more in detail in section 4.4.1, as at July 2019 about 55,000 taxpayers were registered for income

[23]However, the country faces major development challenges and human development indicators may result weak compared to other middle-income countries. Based on the international poverty line of $1.90 a day, and the lower-middle income poverty line of $3.20 a day, it is estimated that 38 per cent of the Swazi population live in extreme poverty, and a total of 60.4 per cent are poor overall. This is accompanied by an unemployment rate of 23 per cent 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 per cent, with a total population of 1.2 million in 2018 (World Bank, 2018).

[24]Table 1.1 of Chapter 1 reports key fiscal and governance indicators for Eswatini and Southern Africa.

[25]CIT is levied at a standard rate of 27.5 per cent, and imposed on taxable income from corporate business activities. Taxable entities include companies, whether incorporated or not, as well as foreign-incorporated entities of a similar nature, whether resident or non-resident; permanent establishments of non-residents; trusts and partnerships. Some entities are exempted from CIT; exemptions must be authorised by SRA's Commissioner General.

[26]PIT has a progressive structure – a maximum marginal rate of 33 per cent and exemptions for income below SZL41,000 ($2,848).

tax.

In terms of filing obligations and deadlines – already mentioned in section 3.3 of Chapter 3 but worth to be discussed again for their practical relevance in this experiment –, income tax returns must be submitted according to a staggered timeline. Non-VAT-registered small and medium enterprises are expected to furnish their returns by 31 October each year, individuals have to file by 30 November, and large companies and VAT registered entities must submit their returns by 31 December. The tax year ends on 30 June. Filing can take place either in person or through the online e-tax system.[27]

Importantly for this study, the law mandates that every registered taxpayer is required to file their return regardless of whether they are operative during the year. Strict sanctions are imposed by law for non-filing and for false assessment. Anyone who fails to furnish a return within the stipulated period may be liable on conviction to a fine of SZL10,000 ($719) and/or imprisonment for a period of up to one year. Those making false assessments with an intention to evade are liable to a fine of SZL50,000 ($3,591) or imprisonment for up to five years.[28] These amounts are discouraging, representing about a quarter of the total annual income of PIT payers and 9 per cent of the total turnover of companies. Overall, also thanks to a sparser population, the enforcement capacity of the tax administration seems to be higher than the SSA average: according to ATAF (2017), Eswatini has a ratio of labour force to tax administration staff of less than 500:1, while most countries in Africa have a ratio of about 3,600:1. However, this positive indicator is somehow muted by the fact that, as mentioned in Chapter 3, in 2017 auditors accounted for 6.5% of total tax administration staff, well below the SSA average of 12% and the 30% international benchmark (Gallagher, 2004).

[27] The revenue authority is encouraging taxpayers to register for the online system (registration is not compulsory and there are no specific thresholds to be eligible for it) which would reduce compliance costs and facilitate the sharing of information with the authority. However, as of July 2019, when the randomisation took place, only 2,700 taxpayers were on the e-tax system (two thirds of which are companies in urban areas, which suggests that registration can be associated with lower compliance costs and higher tax knowledge), out of the 55,000 in total (see section 4.4.1). One of the treatment arms for non-filers also aims to push taxpayers to register for the online system, in the hope of reducing their compliance costs.

[28] However, more drastic measures such as imprisonment are rarely, if ever, implemented.

4.3.2 Patterns of compliance

Appendix Figure A1 shows the trend of PIT and CIT collection over time, according to which CIT collection reported a 14 per cent below-target gap in 2017/2018, while individual income tax performed fairly well, being 13 per cent above target.[29] Apart from broader indicators of revenue collection, a closer look at tax returns data can help better identify the patterns of compliance in the country. Filing categories are directly observable from administrative data in a quite straightforward fashion, even if the algorithm used in this paper is rarely embedded in the monitoring and data mining processes of SSA tax authorities. Administrative data provides a picture of what taxpayers decide to disclose to the authority – unreported income is not observed – and it does not include the informal sector, which by its very nature is invisible to the authority.[30] The following patterns of compliance can be derived in the period 2013-2018 as at March 2019 (about 3 to 5 months after the most recent filing deadline):

- *Active taxpayers*: conditional on filing, the six-year average of active (non-nil) returns is 70.5 per cent for CIT and 74 per cent for PIT payers. Figure 4.1 displays the trend over time as a share of all filing taxpayers (in turn a subset of all taxpayers required to file, see *non-filers* below), which is quite stable for CIT and more oscillatory for PIT payers. Perpetually active taxpayers amount to 52 per cent and 61 per cent of the CIT and PIT *filing* population (5,214 and 14,637 taxpayers, respectively). Out of the total 68,000 income taxpayers registered with the authority, less than a third persistently file non-zero returns.

- *Nil-filers*: the share of nil-returns is derived from the share of active described above. Every year in 2013-2018, about 29.5 and 26 per cent of CIT and PIT returns are nil, thus remitting zero tax. This amounts to 4,707 (45%) and 9,327 (32%) CIT and PIT payers filing nil at least once over the period, respectively. When considering

[29] However, this performance was underpinned by higher PAYE collections mainly due to an increase in employee numbers in the public administration and manufacturing sectors (SRA, 2018).

[30] Although the most recent estimate of informality in Eswatini is five years old, and it may have reduced, the informal sector made up roughly 40% of national income on average for years 2005-2015 according to estimates of Schneider and Medina (2018), a much higher share than the overall regional average of 32% (Table 1.1).

the persistence of this behaviour over time, 2,895 and 4,185 CIT and PIT payers file nil every year in which they file a return. The share of perpetual nil-filers is sizeable: they represent the 29 per cent and 17 per cent of CIT and PIT filing populations and, overall, more than 10 percent of all registered taxpayers. For a more detailed discussion on CIT nil-filers in Eswatini, see Santoro and Mdluli (2019), who show that nil-filing is more prominent in some sectors (construction, ICT and services) than others (public administration) and much more frequent in younger firms (46%) – firms in their first year after registration – than less younger ones (26%). Also, survey data collected in Chapter 3 shows that the main reason behind nil-filing seems to be that the firm is not operative yet.[31] Eswatini is not an exception in SSA and further evidence on nil-filing has been produced in other countries as well.[32]

- *Non-filers*: failure to file – despite being compulsory by law and punished with fines (subsection 4.3.1) – seems to be a much more widespread phenonemon in Eswatini, if not the rule. For CIT and PIT respectively, the six-year average of missing returns is 43 per cent and 57 per cent, as the share of all taxpayers eligible to file an income tax return (see Figure 4.1, right). This implies that more than a third of CIT payers (5,334) and more than half of PIT payers (24,386) supposed to file a return in a given year fail to do so. Considering the persistence of this behaviour over time, as many as 2,339 companies and 10,035 individuals are persistent non-filers, meaning that they have not filed a return since registration. In sum, about 18 per cent of all registered taxpayers have never lodged a tax return while about half of all taxpayers failed to file at least once. Again, this behaviour is not peculiar to Eswatini only. Non-filing has

[31] 88% of the sample agree with this hypothesis, while 63% think it hides evasion.

[32] In Rwanda, 53 per cent and 19 per cent of CIT and PIT returns are nil in 2013-18 (Mascagni et al., 2020), while Mascagni et al. (2019) show that 35 per cent of VAT returns from July 2016 to June 2017 have both zero VAT on sales and zero VAT on purchases. In Ethiopia, about 23 per cent of CIT returns filed in 2006/2013 are from nil-filers (Mascagni and Mengistu, 2016). Likewise, in Uganda 27 per cent of PIT returns are nil over the period 2013-2018[33] and, according to Almunia et al. (2017), 15 per cent of VAT returns in 2012-2015 are nil.

been documented in high-income countries[34] and Latin America.[35] Most relevantly, non-filing is common practice in SSA as well.[36]

In conclusion, compliance with income tax in Eswatini is far from optimal. Only about a third of the total taxpayer population regularly files positive returns and remits non-zero taxes. The remaining two-thirds are characterised by high instability in filing patterns. The majority of taxpayers either intermittently file, and then only in some cases file positive taxes, or mostly just fail to submit a return.

A first consideration is that deregistering from the tax system does not seem to be the preferred option, with only a few hundreds taxpayers exiting the system every year (see Appendix Figure A2). Either due to lack of knowledge, which is addressed experimentally through a specifically tailored letter, or an over-optimistic hope to grow in the future (Santoro and Mdluli, 2019), most of nil- and non-filers keep remaining formally registered despite not contributing any tax. Another anecdotal explanation for the negligible extent of deregistrations is that the taxpayer has to clear all pending tax obligations when deregistering, thus involving remittance of past tax due or the extra compliance cost of filing all missed returns (Santoro and Mdluli, 2019).

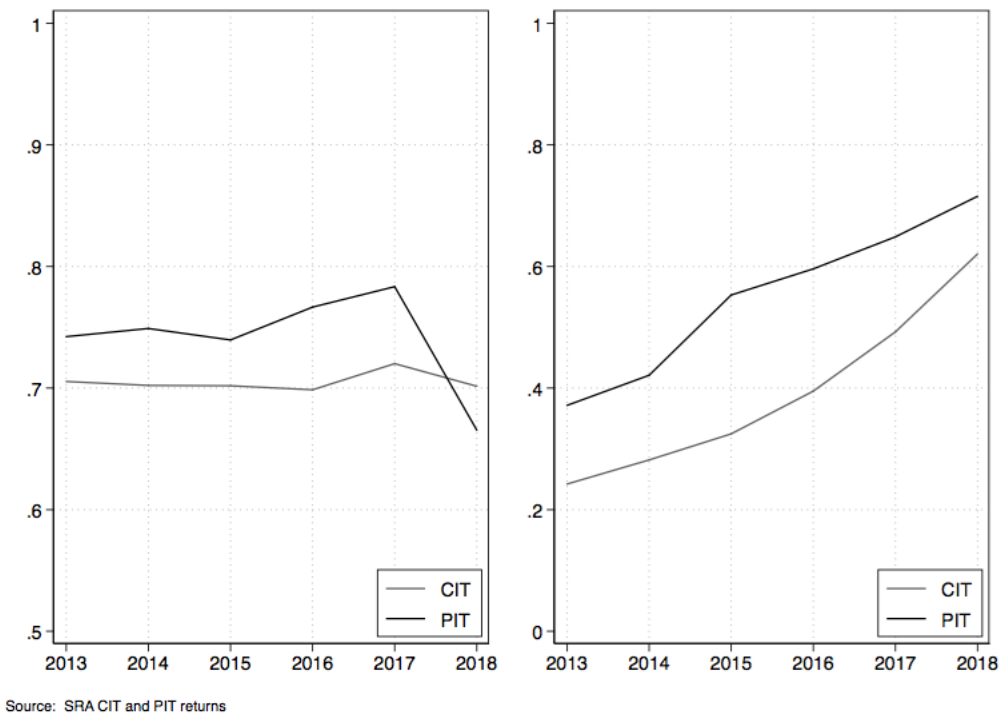
Second, nil-filing seems to be more common for CIT payers while non-filing is the

[34]Meiselman (2018) shows that non-filing nears 50 per cent for local taxes in Detroit, USA, while the share of non-filers for the US federal individual income tax is about 7 per cent for the period 2000-2012 (Erard et al., 2018).

[35]In Guatemala, the share of non-filers of income tax in 2013 is 39 per cent (Kettle et al., 2016); in Costa Rica, 50 per cent of registered firms failed to file in the period 2006-2014 (Brockmeyer et al., 2019); in Venezuela, the non-filing rate for the municipal income tax is 20 per cent (Ortega and Scartascini, 2016a).

[36]In Rwanda, over three-quarters of individuals supposed to file for the fiscal year 2018 failed to do so, while about half of companies did the same. Figures from Uganda are even higher, with the average rate of PIT non-filing being 86% over the period 2014-2018 (figures from Rwanda and Uganda have been computed by the author in parallel studies on tax compliance, drawing on detailed tax returns data). In Malawi, almost 50 per cent of income taxpayers have filed no tax return and/or made no tax payment over the period 2014-2016 (Ligomeka, 2019a). Additional descriptive evidence from Kenya shows that out of the over nine million registered taxpayers, only 3.5 million filed their 2018 returns (see `https://www.businessdailyafrica.com/lifestyle/profiles/what-to-expect-file-nil-return/4258438-5232858-oiqmom/index.html`, accessed on June 10, 2020). Lastly, Moore (2020) notes, non-filing rates in Nigeria are exceptionally high: 98% for PIT; 94% for CIT; and 95% for VAT.

Figure 4.1: Compliance trends 2013-2018: Active (left) and Non-filers (right)



Note: active taxpayers (left) are measured as a share of all filing taxpayers. Non-filers (right) are measured as a share of all taxpayers required to file for income tax.

preferred choice of PIT payers – where non-filers outnumber active taxpayers every single tax year (Chapter 3). At the same time, an additional layer of complexity is due to the fact that nil-filing seems to be correlated with non-filing. As discussed in Santoro and Mdluli (2019), non-filers who eventually end up filing are more likely to declare nil, as they are suggested by tax officials to do so to put their records in order.

4.4 Research Design

4.4.1 Data sources

Different data sources have been used in this study. On one side, detailed administrative data represent the most important source in order to run the impact analysis. On the other, two additional sources, the post office delivery reports and the survey data collected in Chapter 3, proved to be extremely useful for estimating the impacts on the actually-treated subgroup and to provide more information on the mechanisms at play, respectively.

I have direct access to administrative data, granted by a confidentiality agreement signed with the revenue authority. I merge together different types of datasets. First, I refer to the taxpayer registry, which contains background information on the universe of registered taxpayers, such as size, location, sector of activity. The registry is also used to gather contact information of the taxpayers to be nudged. By the end of July 2019, when the sample randomisation took place, the registry contained a total of 55,462 taxpayers, 30 per cent of which are companies and the remaining individuals. Second, I use income tax returns of corporations (CIT) and self-employed (PIT) during the period 2013-2019, in which: (i) tax year 2018 serves as a baseline year to identify the filing categories, (ii) year 2019 is used to observe outcomes after the intervention, and (iii) years 2013-2017 are considered to measure pre-trends in behaviour. Returns data amounts to about 6,000 companies and 12,000 self-employed filing per year on average. This data includes all line items from the tax return form with particular detail, including the main financial variables, such as turnover, gross profits, and tax liability. Third, VAT returns for the same period are observed as filed by the subpopulation of about 4,000 VAT registered entities in the country. Fourth, information on deregistrations from the tax system are available. A unique taxpayer identification number (TIN) is assigned to each taxpayer, which is consistent across all datasets and used as a key identifier in merging them.

Administrative data serve two main purposes. First, it is needed to identify, locate and randomly assign the taxpayers to the treatment arms. Second, the data allows to measure pre-treatment compliance and identify the three categories under study. The filing behaviour is classified by looking at the most recent tax year, 2018. This means that a taxpayer labelled as *active* positively filed a 2018 tax return, while a nil-filer reported zero and a non-filer failed to file in that year. Furthermore, I am able to observe the filing behaviour in the previous 5-year period and create the *perpetual* sub-category, i.e. taxpayers who keep filing in the same way every year: for each filing category – non-filers, nil-filers, active filers – there will be a relatively large subgroup of perpetual non-/nil-/active filers who have been consistently filing in the same way since registration. To the best of my knowledge, this is the first time that (any type of) tax data in Eswatini has been used for research purposes.

Additional relevant data come from the delivery reports of the post office. Uncollected letters (more on the delivery process below) within a period of 2 weeks since delivery were returned back to the post office who unambiguously linked each letter to a recipient's TIN. With this data, I am able to know whether letters have been effectively collected and identify the taxpayers actually treated by the nudge. This piece of information is essential to estimate the local average treatment effect (LATE), or *treatment on the treated*, as explained more in depth in section 4.4.3. Unfortunately, the post office has still not released the data on the actual dates of pick-up but only information on whether the letter has been returned uncollected or not. It is also not clear whether a letter is returned to the post office because the postman failed to deliver it to the designated box or because the recipient did not collect from their mailbox.[37]

As a final source of data, I recur to the taxpayer perception survey fielded in November 2019 on 1,009 PIT payers, evenly split by active and non-filers. The survey was meant to be representative of the PIT payers population. Furthermore, it is directly linked with tax returns data through taxpayer identification numbers. The data gathered in this effort are used to study the determinants of compliance in Chapter 3 parallel study and represent a source of information used to explore the mechanisms of impact more in depth (section 4.6).

[37]For more information on the process of delivery, see the paragraph on logistics below.

4.4.2 Study Design: Letter Message Experiment

Experiment letters and testable hypotheses Using the terminology of Harrison and List (2004), my RCT can be labelled as a natural field experiment, since the taxpayers under study do not know that an experiment is occurring. The letter experiment consists of a one-page letter mailed to the taxpayer. I experimentally manipulate the content of the letter by highlighting alternative key messages in a box at the centre of the page, so to make it more salient to the recipient (BIT, 2012). For the same purpose, a clear subject line is added to the letter, such as *Comply to avoid penalties and fines for failure to file.* Everything apart from the box and the subject line, such as the headings and footers, the introductory 2-line paragraph and the final text with information on how to communicate to SRA (phone number, fax number and email address) remain constant across all treatment arms.[38] Also, the format of the letter mirrors that of SRA standard mailings used in their regular correspondence with taxpayers. To enhance credibility, experiment letters have been signed by the SRA Director of Compliance. Finally, in line with practitioners' recommendations (BIT, 2012) the letter is kept simple and easy to understand. Convoluted and wordy sentences are removed. Letters are personalised in the sense that they mention the name of the taxpayer at the top of the page, and use active language, addressing the reader as *you*.

An example of the experiment letter is provided in Appendix Figure A3. While the text may seem long, it is also true that I needed to strike a balance between the need to be concise/effective and the need to convey complete, exhaustive information. This is not different from what usually happens in the literature. The length of this experimental letter is shorter than in Carrillo et al. (2017), Ortega and Scartascini (2016a), Bergolo et al. (2019), and more in line with Del Carpio (2014), Pomeranz (2015) and Mascagni et al. (2017).

The SRA standard letter was extended along five dimensions: threat/deterrence (T1), taxpayer assistance on how to comply (T2) and how to deregister from the system (T3), a reciprocity appeal stressing the fiscal exchange aspect of taxation (T4) and social information (T5). The control group is assigned to an untreated mailing condition (T0) and

[38]In the introductory 2-line paragraph an element of tailoring is introduced, by which the SRA communicates to a given taxpayer that it is aware of the taxpayer's filing category.

does not receive any letter. Mostly due to power reasons, I was not able to test the impact of a placebo letter - the impact of receiving correspondence from the authority - as often done in similar studies (Del Carpio, 2014; Ortega and Scartascini, 2016a). Also, it would have been complicated to isolate the effect of a placebo letter in a context in which the tax agency does not usually send mailings to taxpayers (more on this below). Furthermore, what is usually adopted as placebo letter is a neutral message reminding about the filing deadline, so presumably acting as a facilitation nudge in contexts where tax knowledge is very low (Chapter 2). Despite the lack of a placebo group, I attempt to measure the impact of receiving any letter in section 4.6.1.

The deterrence nudge T1 emphasises the size of the penalties associated with a given wrongdoing and increases the pecuniary costs of evading. It could also be the case that taxpayers are overestimating such costs and then the intervention would imply a downward revision of this component, thus making evasion even more likely. However, survey evidence from Chapter 3 shows that about 98 per cent of the sample does not know how much these penalties amount to. An element of targeting is introduced as the content of T1 changes by filing category. For non-filers, the penalties on failing to declare are highlighted, while, for nil- and active filers, penalties on false declarations are included. The nudge message for non-filers consists in adding the following paragraph:

- *SECTION 66 of the INCOME TAX ORDER (1975): a taxpayer who fails to submit a return within the stipulated period commits an offence and may be liable on conviction to a fine of* ***E10,000****, or imprisonment for a period of up to* ***one year****, or both.*
- *SECTION 40 of the INCOME TAX ORDER (1975): a taxpayer who defaults in submitting a return for any year of assessment is liable to pay additional tax of an amount equal to* ***twice the tax chargeable*** *in respect to his taxable income for such year of assessment.*

Please, comply with your tax obligations and to ensure your declarations are correct to avoid fines and penalties.

Nil- and active filers received the same message with the only difference of the fines imposed on false statements (SZL50,000 or imprisonment up to five years) and a final sentence saying that *if all information you reported is complete and correct, you do not need to make any changes.* Taxpayers with VAT discrepancies are showed the turnover reported in both the income tax and VAT returns, as well as the amount of the discrepancy arising between the two. Consistently, the following hypothesis on T1 is derived:

> **H1**: *receiving a deterrence letter T1 increases both the probability to file for non-filers and the tax declared for nil-filers and active. The impact should extend to previous years' returns and spill over other taxes, such as VAT. It should also be larger for smaller taxpayers, given that the cost of evading is now made more salient and is proportionally higher for them.*

The compliance costs nudge T2, delivered to non-filers only, communicates all the steps necessary to file a return, including a reminder of the upcoming filing deadline and indications on how to register for the e-tax system (section 4.3). The overall goal is to reduce the compliance costs that make it extremely difficult for taxpayers to comply with tax rules (Slemrod 2007). Compliance costs can be seen as the burden taxpayers have to bear to be compliant, which includes the time and cost associated with preparing tax returns, filing, paying, acquiring the relevant tax knowledge and interacting with tax authorities. These costs are likely to be large in Eswatini. Survey evidence in Chapter 3 shows that 83.5 per cent of the non-filers in the sample do not know when the next deadline is, compared to 31 per cent of surveyed active taxpayers. More broadly, active filers perform 35 per cent better than non-filers in a six-item tax quiz embedded in the survey.[39] This evidence demonstrates how compliance costs could be more relevant for non-filers, hence the targeting on this category. The tone is friendly and instructive, with no reference to penalties for non-compliance. A blank tax form, as well as a blank form for e-tax registration are enclosed with the letter. Both for its content and attachments, T2 can be seen as a more complete educational nudge than the reminder-only type of messages usually tested in the literature (see section 4.2). The following text is introduced:

> *You just have to follow these simple instructions:*
>
> 1. *Obtain the CIT/PIT Income Tax Return form from a SRA Service Center countrywide or download it from the SRA website. For your convenience, also find the tax return form attached to this letter.*
> 2. *Fill in the attached form correctly so that your taxes can be calculated*

[39]The difference in knowledge is statistically significant at the 1% level.

3. *Follow instructions on the tax rate applicable to get the correct tax payable*
4. *Sign and submit the Tax Return to any SRA Service Center before the deadline XXX*[40]

Alternatively, you can also save time and travel costs of coming to SRA to physically submit your Income Tax Return by filing online. All you need to do is to register for e-tax with SRA. Fill in the e-tax registration form attached and submit tax return online before the deadline. Please, note that e-tax registration takes 5 days and only after receiving registration confirmation from SRA you will be able to file online. Please, follow the steps above and file your past and future returns. ***Learn how to file in order to be a compliant taxpayer.***

The following hypothesis is formulated:

> **H2**: *receiving a compliance costs letter T2 increases the probability to file for non-filers. The impact should extend to previous years' returns as well. It should also be larger for smaller taxpayers, given that compliance costs are usually regressive in nature.*

The deregistration nudge T3, targeted at non- and nil- filers, similarly shows the service-oriented side of the tax authority. In T3, I explain which are the steps to follow to exit the tax system in a clear and direct way, as the current deregistration process may be considered confusing to go through without assistance. Similarly to T2, a blank deregistered form is included in the mailings. The goal is to reduce taxpayers' ignorance and perceived complexity associated with deregistering:

If your business is not operating, you can easily deregister from SRA. You just have to follow these simple instructions:

1. *Obtain a de-registration certificate for the business from the Registrar of Companies office at the Ministry of Commerce.*
2. *Visit any on the SRA service centers with the following documents:*
    - *De-registration certificate*
    - *A signed Taxpayer Declaration Form which is obtainable from SRA.*
    - *If the business is also registered for VAT, it is a mandatory requirement for you to bring the original copy of your VAT registration certificate.*

[40]The deadline varies according to the taxpayer's type (see section 3.3).

3. Once your documents have been accepted, you will receive confirmation on your deregistration status from an SRA official within 5 working days. As you deregister you are also reminded to comply with all your past tax obligations (tax returns and any payments due).

In case you are no longer trading and don't plan to be operative in the future, deregistering with SRA is advisable as you will no longer have periodic obligations to file returns or make payments for the deregistered TIN. If your business is dormant but you still intend to keep your business registered with SRA, you can disregard this message and continue to file your returns as required.

The following hypothesis stems:

> **H3**: *receiving a deregistration letter T3 increases the probability to exit from the system. The impact should be larger for perpetual non- and nil-filers, since it is more likely that they ceased operations, thus it is optimal for them to just deregister.*

The fiscal exchange nudge T4 is targeted at PIT non-filers and active payers[41] and appeals to the taxpayers' morality by stressing the importance of tax revenue in order to finance national development. First, the recipient is told that their contribution directly affects all Emaswati's lives. Second, the text gives examples of flagship development projects. These projects have been directly financed by the taxpayers' money and help to make the reciprocity link between taxes and public services more concrete. Lastly, the message ends with an open question rhetorically addressing the taxpayer:

Your tax payment contributes to the funding of publicly financed services that make the lives of Emaswati better.
For example, public infrastructures are directly funded by the taxes you pay: last year, the Lower Usutu Smallholder Irrigation Project in Lubombo Region has been successfully constructed and is now fully operational. In addition, rural electrification, rural water and many other development projects, which the Government has embarked on, are financed by the taxes you pay. Therefore, tax compliance by all stakeholders has positive effects on the lives of Emaswati. Please, declare your taxes correctly.
Are you going to support the building of a better Eswatini for all?

[41]Originally intended to nil-filers as well, it was not implementable because of the too small final size of the nil-filing subgroup. The same is true for CIT non-filers.

In relation to T4, the following hypothesis is testable:

> **H4**: *receiving a fiscal exchange letter T4 may increase or decrease compliance, depending on the level of satisfaction with government spending.*

Lastly, the social norm nudge T5, mailed to PIT non-filers only for reasons of sample size, gives information about the extent of filing in the taxpayers' region in the period 2013-2018. Following Cialdini and Goldstein (1991), the descriptive norm, i.e. what others do, rather than the injunctive norm, i.e. what others believe or approve, is communicated. The main idea is to to encourage non-filers to join the compliant majority by increasing the moral costs of non-compliance (Myles and Naylor, 1996; Traxler, 2010; Frey and Torgler, 2007) mostly through feelings of shame and guilt for being part of a small minority of evaders (Elster, 1989; Wenzel and Taylor, 2004). T5 is framed with respect to the region level, rather than the national level, to make the norm more specific to the individual's context or group to which she belongs (Hallsworth et al., 2017). A reference level of compliance of 66 per cent is presented. This compares to survey evidence from Chapter 3 showing that the perceived extent of compliance is about 53 per cent. In this sense, T5 may update upwards the taxpayers' perception and ideally push them to join the majority. A the same time, T5 risks to back-fire for some taxpayers if the compliance rate revealed in the message is lower than what previously thought by them. T5 reads:

> *Do you actually know that the majority of your peers in Eswatini regularly submit their declaration for Personal Income Tax (PIT)?*
> *According to SRA's administrative records, in the period 2013-2018* ***two thirds (66%)*** *of PIT payers living in your Region declared their income tax. You are currently part of the minority of PIT taxpayers in your Region who are yet to declare for this tax for the year 2018.*
> ***Please, declare all your taxes and be part of the majority.***

Relatedly, hypothesis H5 is formulated:

> **H5**: *receiving a social norm letter T5 may increase or decrease compliance, depending on the perceived descriptive norm on compliance held ex-ante.*

The final hypothesis tested considers that the mere fact of receiving a letter from the authority may increase perceived enforcement and thus compliance:

> **H6**: *receiving* any *letter from the authority signals that the taxpayer is under the authority's radar and therefore increases both the probability to file for non-filers and the tax declared for nil-filers and active. Spillover and heterogenous effects should work in the same direction as in* ***H1****.*

Sample The sample is extracted from the registry data described in subsection 4.4.1. Inclusion criteria are applied, such as: (i) the taxpayer has registered by December 2017, so to be liable to file an income tax return for the year 2018 and therefore allow for categorisation in a filing category;[42] (ii) the registration status is labelled as active, meaning that the taxpayer did not deregister and exit the tax system by December 2017 (therefore she is still liable to file), as well as the taxpayer is not exempted from income tax;[43] (iii) she has been uniquely identified and duplicates are removed;[44] (iv) she has valid address information and is located in Eswatini. A clean population of about 40,000 income taxpayers is derived from the original taxpayer registry.

The population of taxpayers is then merged with CIT/PIT returns for 2018 and categorised according to their filing behaviour. Crucially, non-filers are those taxpayers that appear as registered with the authority but for which a CIT/PIT return is not found in 2018. Non-filers are, in principle, potential filers who have not submitted their returns yet. The share of non-filing is therefore a moving target and depends on the specific date at which the data is analysed. In this study, sample creation and randomisation took place at the end of July 2019, hence non-filers had not file yet by that time, about nine months after deadline. At the same time, active and nil-filers are observed directly from the CIT/PIT tax returns data. In line with the compliance patterns in the country (see section 4.3), non-filers and PIT payers represents the majority of taxpayers. Given the limited size of the target population, no further sampling strategy is pursued and the randomisation involves all 40,000 income taxpayers. As explained above, the latter represent the *cleaned* total of actively registered agents who are all liable to file a return, since it does not contain dupli-

[42] Nine percent of the observations are removed, as registered in 2018 or 2019. Also, an additional 20% of taxpayers are discarded since they are not registered for income tax.

[43] About 2,000 taxpayers are dropped as officially suspended, plus only 32 entities are officially exempted.

[44] Duplicates can occur when the taxpayer is registered for multiple taxes or due to technical mistakes in the registration process. About 12,000 duplicate observations are dropped.

cates, deregistered or exempted taxpayers. In this sense, the experiment can be considered already *at scale* and representative of the entire taxpayer population in the country.

The summary statistics of key characteristics of the sample, as derived from the tax authority records, are reported in Table 4.1 below. Some relevant differences emerge across groups. First, non-filers are less likely to be incorporated than nil- and active taxpayers, in line with what described in section 4.3.1. Second, groups are roughly comparable in terms of business age, with nil-filers being slightly younger. Third, non-filers are less likely to be in the trading sector and, perhaps relatedly, less likely to be registered for VAT and e-tax. More specifically, while about a fifth of active taxpayers are registered for VAT and 7 per cent overall in the sample, very few (2%) taxpayers in the sample are registered for e-tax. Fourth, as many as 60 per cent of taxpayers in the sample keep filing in the same way since registration. While it is positive to see that 76 per cent of active are persistent in their behaviour, it is also concerning that more than half of non- and nil-filers have been non- and nil-filing since they entered the tax system. Lastly, categories are not too unevenly distributed across the four districts.

Randomisation into treatment is performed through a replicable algorithm in Stata and takes place within each filing category, since the nudges are tailored. The choice of the treatment arms to assign to filing categories depended on power considerations. While I was powered enough to capture small effects for non-filers and actives, the smaller size of the nil-filing group implied that maximum two treatments could be tested on them.[45] Non-filers are allocated to treatments T1 to T5, while nil-filers receive T1 and T3 and active taxpayers are assigned to T1 and T5 only (see Table 4.2 below). Stratified randomisation is implemented in order to achieve better balance, increase statistical power and allow heterogeneity analysis (Glennerster and Takavarasha, 2013). The variables chosen as strata are those expected to influence the outcome. For nil- and non-filers, strata includes the district in which the business is located and whether the taxpayer is at his first year since registration. Size is not used in this case since taxpayers in these categories are never

[45]Pre-specified power calculations (power = 80% and α = 5%) indicated that a minimum detectable effect of 5% and 4% for non-filers (reduction in non-filing) and active (increase in tax due) could be found. A MDE of 20% for nil-filers was implied as well, which is not extremely high but still more difficult to reach when compared to the other two categories.

Table 4.1: Experiment Sample - Summary Statistics

	Non-filers		Nil-filers		Active		Total	
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE
CIT	25,525	0.18	3,575	0.40	11,405	0.33	40,505	0.24
		(0.00)		(0.01)		(0.00)		(0.00)
Reg. year	25,525	2012.63	3,186	2012.86	9,225	2011.77	37,936	2012.44
		(0.02)		(0.06)		(0.04)		(0.01)
# returns	25,525	4.34	3,575	3.97	11,405	4.34	40,505	4.31
		(0.01)		(0.03)		(0.01)		(0.01)
Hhohho	25,525	0.38	3,575	0.41	11,405	0.42	40,505	0.39
		(0.00)		(0.01)		(0.00)		(0.00)
Lubombo	25,525	0.15	3,575	0.12	11,405	0.11	40,505	0.14
		(0.00)		(0.01)		(0.00)		(0.00)
Manzini	25,525	0.39	3,575	0.42	11,405	0.39	40,505	0.40
		(0.00)		(0.01)		(0.00)		(0.00)
Shiselweni	25,525	0.08	3,575	0.06	11,405	0.07	40,505	0.07
		(0.00)		(0.00)		(0.00)		(0.00)
Trading	25,525	0.13	3,575	0.27	11,405	0.27	40,505	0.18
		(0.00)		(0.01)		(0.00)		(0.00)
VAT reg.	25,525	0.02	3,575	0.08	11,405	0.20	40,505	0.07
		(0.00)		(0.00)		(0.00)		(0.00)
Etax reg.	25,525	0.02	3,575	0.05	11,405	0.18	40,505	0.07
		(0.00)		(0.00)		(0.00)		(0.00)
Perpetual	25,525	0.53	3,575	0.59	11,405	0.76	40,505	0.60
		(0.00)		(0.01)		(0.00)		(0.00)

large.[46] For PIT, a dummy for whether the individual is a sole trader (see section 4.3) or not is used as well. For active taxpayers, the strata above are used and an additional dummy for whether the taxpayer is large is added as well. For the subgroup of active taxpayers with VAT discrepancy, district, first year and compliance type (active, nil- and non-filer) are used as strata.[47] The final random allocation is produced as shown in Table 4.2.

[46]Unsurprisingly, large taxpayers only rarely fail to file or file zero, given that this behaviour would almost automatically trigger an audit from the authority.

[47]Business size is not included as highly correlated with being active - see previous footnote.

Table 4.2: Treatment groups

Category	*T0 Control*	*T1 Deterrence*	*T2 Costs*	*T3 De-reg.*	*T4 Exchange*	*T5 Norms*	*Total*
Non-filers	15,266	2,431	2,429	2,424	1,373	1,366	25,289
Nil-fllers	1,182	1,162		1,164			3,508
Active	3,607	3,574			3,578		10,759
VAT disc.	477	472					949

Sample balance on observables As expected, the randomisation is successful in creating balanced groups. Appendix Section A1 provides a set of tables that compare the pre-treatment balance of characteristics between the different nudge types, and also shows p-values from F-test of the joint significance of characteristics in discriminating between groups. Balance is achieved in statistical terms: of the 145 tests presented in Section A1 in only seven cases the null hypothesis of equality of means is rejected at the 10 per cent level, well below what is expected from pure chance.

Logistics and timeline The field experiment was organised in close collaboration with the revenue authority, which provided invaluable support throughout implementation. Since the nudge represented an official communication from the authority, the content was ratified by the Legal Office as well.[48] Before the experiment started, the author briefed all tax officials working in the processing of the mail, as well as those involved in the call centres. The aspect of confidentiality of the research has been stressed so that taxpayers were not aware that they were part of a study. All taxpayers' queries have been directed to the call centre who followed a pre-specified protocol in addressing the calls and clarified that the letter was to be considered as an important piece of communication from the authority, even if it did not necessarily translate into an actual audit.[49] Also, the taxpayer TIN was used as a tracking identifier for the corresponding letter. A similar nation-wide

[48] The trial also obtained ethic clearance from the University of Sussex (ER/FS294/1).

[49] Anecdotal evidence also shows that some nudged taxpayers tried to reach SRA staff informally, and, in few cases, approached the Office of the Commissioner General directly. These cases were not the norm and, despite introducing an element of uncertainty and bias in the identification strategy, can be considered as negligible.

tax communication effort has never been carried out in the country before.

Before this experiment, the tax authority did not send standard mail to broad categories of taxpayers, but rather ad-hoc letters to specific recipients and for specific reasons, such as informing about taxes due, the outcome of a tax assessment, prepayment of taxes and, more relevant to this study, to nudge VAT non-filers to submit their returns.[50] Appendix Table A7 reports all the types of notices sent by the authority. Nudging usually takes place through newspapers and billboards, while most physical letters inform about payments.

The letters have been sent using registered mail, so that the delivery could be tracked for each taxpayer. Each letter was assigned an identification number that was uniquely linked to a taxpayer's TIN. For this reason, it is almost impossible that a letter has been delivered to a taxpayer in the control group, since control TINs have not been shared with the post office or linked to any letter. The mailing process in Eswatini consists in the post office leaving the letter into the recipient's box, which is located in the nearest post office. The letter remains in the box for 14 days and is returned to the central post office if uncollected in that time window. It is the responsibility of the taxpayer to routinely check their postbox to see if any mail has been delivered.[51] The letters, one page long, have been folded in SRA-labeled envelopes by a dedicated team of SRA staff and interns. A different colour of the envelope has been chosen so to make the letter more salient to the taxpayer and differentiate it from SRA standard mail. SRA staff, in collaboration with the national post office, processed the posting in the field.

In terms of timing of the experiment, the letters have been mailed in autumn 2019, after the end of tax year (June 2019), so to study changes in reporting behavior, while not affecting production decisions. The letters were posted in three waves. Given the staggered filing deadline for income taxpayers (see section 4.3), letters were grouped in batches and sent 40 days before the recipient's deadline. That means that for small CIT payers, whose deadline falls at the end of October, letters were sent by mid-September (round 1). Likewise, PIT non-VAT registered payers and large CIT/PIT entities received the nudge by mid-October (round 2) and mid-November (round 3), respectively. The 40-

[50]However, due to resource constraints not all VAT non-filers are systematically nudged in each period.

[51]Despite the peculiarity in the mailing delivery process in Eswatini, which entail extra costs in reaching the nearest post office to collect the letter, the SRA and post office were quite confident that taxpayers, especially business people, check their mailbox on a regular basis.

day timeframe has been chosen to reach a balance between the need to make the nudge punctual and the need to nudge taxpayers before they actually file a return. The closer to the deadline the recipient receives the letter, the more likely his response will be shaped by the nudge, while, the further from the deadline, the more likely it is that taxpayers have not declared yet and can be influenced by the nudge. The tax returns data for the period 2013-2018 shows that while 37 per cent and 25 per cent of CIT and PIT payers file after the deadline on average every year, an additional 50 per cent and 60 per cent file in the last 40 days. That translates in about 85 per cent of total taxpayers who have not filed 40 days before deadline, as summarised in Appendix Figure A4.

4.4.3 Identification strategy

Main specification To comply with international research standards, I pre-registered my trial with the AEA RCT Registry (ID number AEARCTR-0004753). The identification strategy is quite straightforward as it relies on the fact that the study is a randomised controlled trial. I regress my compliance outcomes on treatment dummies and taxpayer controls. The estimating equation is run for each filing category and, for non- and nil-filers, writes:

$$Y_i = \alpha + \sum_{j=1}^{n} \beta_j Treat_i + X_i \Gamma + \epsilon_i \tag{4.1}$$

While, for active taxpayers, the equation takes the following form:

$$Y_i = \alpha + \sum_{j=1}^{n} \beta_j Treat_i + \theta_i Y_{i0} + X_i \Gamma + \epsilon_i \tag{4.2}$$

Where the outcome Y is the ex-post compliance behaviour of taxpayer i, as measured at three months after the experimental interventions, unless stated otherwise. The variable $Treat_i$ indicates which treatment nudge j taxpayer i has been assigned to. Therefore, β_j stands for the intention-to-treat (ITT) estimate of impact of nudge j. The set of control variables X_i includes the strata used in the randomisation (see section 4.4.2) to assure valid inference (McKenzie, 2012) and additional controls to increase power, such as the sector of activity and the frequency of filing in the previous five years. The error term

ϵ_i is clusterised at the taxpayer level and robust to heteroskedasticity. When it comes to active taxpayers, for which non-zero filing behaviour is observed before the experiment, I also include the baseline outcome variable Y_{i0} (lagged tax) as a pre-treatment control, in an ANCOVA estimation which helps reduce the variance of the error term and thus results in gains in statistical power (McKenzie, 2012).

An additional specification calculates the impact of letters on recipients who actually collected them. While the main specification above provides ITT estimates, LATE estimates are derived by using an instrumental variable model where where the actual collection of the treatment letter is instrumented by the random assignment to it. LATE refers to the impact on compliers, which would be larger than the ITT estimates, given the partial letter take-up (Angrist and Pischke, 2009). Lastly, in section 4.5.7 I examine the robustness of my core findings to: (i) clusterising the error terms at the district level, (ii) using different operationalisation of tax amounts (for actives only), and (iii) allowing for delays in the timing of the mail delivery.

Outcomes The outcomes of interest differ by taxpayer category. For non-filers, the main outcomes are the probability to file a tax return and the probability to deregister from the system. For nil-filers, I consider the probability to switch to non-zero filing as well as the probability to deregister. Lastly, for active taxpayers and those in the VAT discrepancy group, I focus on the tax declared. The latter is framed both as the probability to declare more than previous year, thus to show an increase in reported liability, as well as in the level of tax declared, adequately transformed in logs and in hyperbolic sines and converted in US dollars.

Additionally, I consider secondary outcomes as: (i) the probability to file or amend previous returns, (ii) spillover effects on VAT/PAYE declaration for the subset which is VAT/PAYE registered (mostly active), (iii) for active only, both turnover and costs declared, as well as the probability to positive filing, given the high instability in filing patterns (section 4.3) (iv) for non-filers only, the probability to file on time as well as the probability to register for the e-tax system.

Choice of model The choice of the model depends on the outcome of interest. For dichotomous outcomes, such as the probability to file a return, I implement a linear probability model which provides easier interpretations for the marginal effects than logistic regression models.[52] As discussed in the robustness section, results do not change if I use a logit model. On the other hand, if the outcome is tax liability, I adopt a tobit model which best addresses censored distributions, as widely used in the tax literature Alm et al. (2010); Slemrod and Weber (2012); Alm and McClellan (2012); Mascagni et al. (2018).

Heterogeneity Despite being already quite targeted to each filing category, the nudges under study can nevertheless show heterogenous effects across subgroups. For this reason, I perform the analysis of taxpayers' responses across four sub-groups for which I am sufficiently powered: (i) perpetual taxpayers, (ii) top income decile at the baseline (available for active taxpayers only – 1,100 of which are in the top decile), (iii) being a new registered taxpayer (i.e. being in the first year after registration) and (iv) being located in urban areas.[53] Variables (ii) and (iii) are used as strata and are balanced by design. However, also (i) and (iv) results to be balanced, as shown in Appendix Section A1.

4.5 Results

4.5.1 Do non-filers switch to filing their returns?

In this subsection I measure the response of non-filers. As explained in section 4.3, these taxpayers failed to file a return for tax year 2018, in contrast with what the tax code prescribes. Figure 4.2 shows the rate of income tax filing and payment over time by treatment status, for mailing round 1 (SMEs) and 2 (individuals) taxpayers.[54] The start of the experiment and the corresponding deadlines are indicated by vertical lines. While pre-intervention trends in the treatment and control groups are almost identical for both groups, a small positive treatment effect on filing and payment start to emerge after the

[52]While the assumption of homoskedascity does not hold in a LPM, calculating *robust* standard errors controls for that (Angrist and Pischke, 2009). Moreover, LPM does not restrict predicted values within the 0-1 interval, but the share of such values is not high, ranging from a minimum of 0% to a maximum of 5% of the sample.

[53]For individual taxpayers only, I also look at age and whether the taxpayer is married.

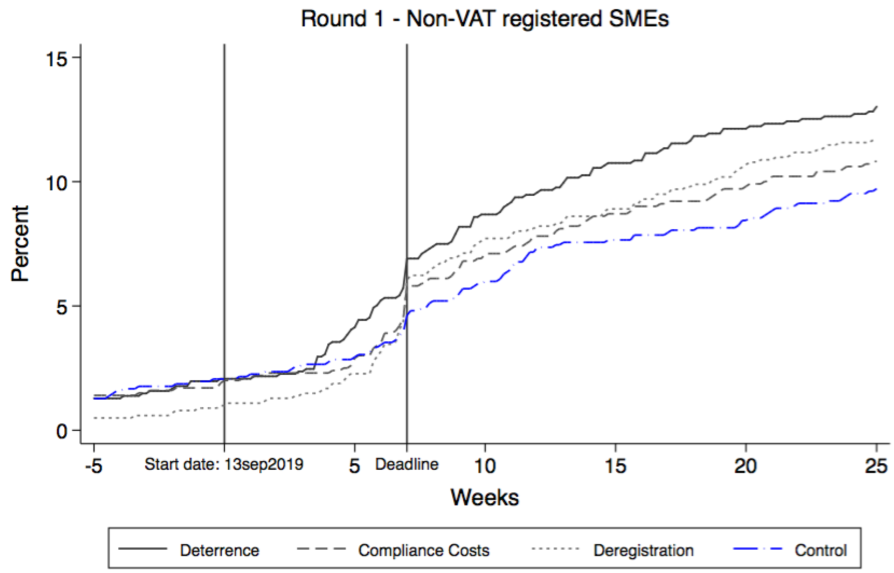
[54]No data comes from mailing round 3 (large taxpayers) as there are no non-filers due to file in that round.

official start of the experiment. The rising in filing is not immediate. In some cases, like with the deterrence nudge in round 1, nudged taxpayers start filing much before the deadline and continue doing so after it, probably due to the threat of a penalty. A similar pattern emerges from individuals (round 2) nudged with the compliance costs letter, even if the jump is much closer to the deadline and supportive of the reminder mechanism explored in section 4.6. In all other cases, the increase in filing is sizeable but appears much after the start of the experiment, mostly even after deadline. This could hint to the fact that significant delays in delivery took place in the field, for which I run a robustness check in section 4.5.7. In any case, the increase in filing takes a stable trend around the deadline, about seven weeks after the experiment start date, and remains approximately constant during the next 15-20 weeks.

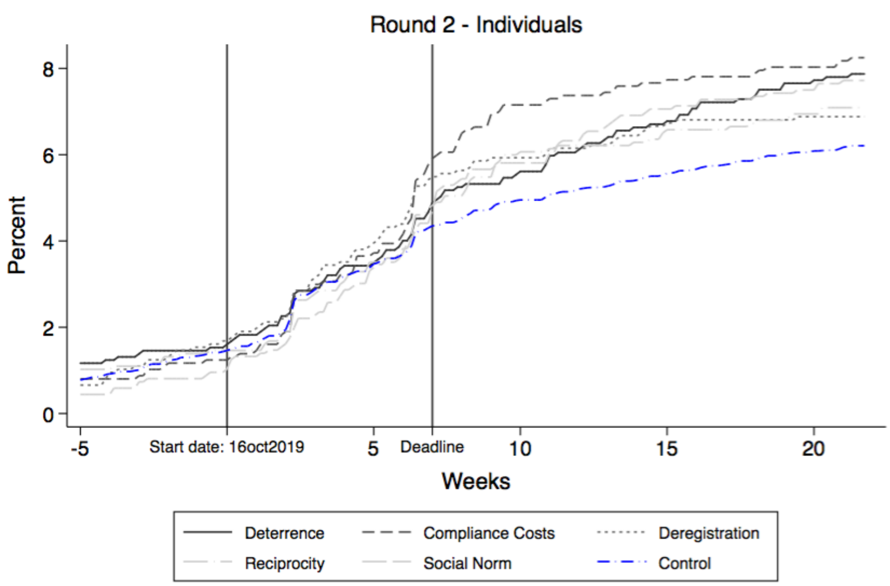
Turning to the OLS estimates, Table 4.3 reports the coefficients of impacts of the treatment letters on the probability to file a return. Standard errors are presented in parentheses. Column 1 and 2 pool CIT and PIT together, without and with controls, respectively. T4 and T5 arms, which are tested on PIT only, are dropped from the pooled sample. Columns 3 and 4 consider CIT payers, while columns 5 and 6 look at PIT ones considering only treatments T1-T3 which are common to both categories, so to enhance comparability across the two. Lastly, columns 7 and 8 report all treatments for PIT payers. Coefficients from regressions with controls are also displayed in Figure 4.3. All coefficients must be interpreted as incremental changes with respect to the control group. The control group mean, i.e. the filing average at the endline for not nudged non-filers, is reported at the bottom. By construction, the filing average of the control group at baseline is 0 per cent and therefore not reported in the table. As expected, very few (7%) taxpayers who failed to file in 2018 filed for 2019, thus suggesting again how non-filing can be a well-rooted behaviour persisting over time.

Results from the table show that all letters increase the probability to file in the pooled sample (col. 1), and significantly so. When controls are added (col. 2), deterrence increases filing probability by 2.6 percentage points (p.p.) or a 38 per cent increase over the control group. The compliance cost nudge as well positively pushes non-filers to submit a return, even if by a smaller increment (1.3 p.p. or 19%). Interestingly, the deregistration nudge, who was aimed at increasing the *exit* from the tax system, has the unintended effect of

Figure 4.2: Declaration Rates over Time by Treatment Group: round 1 (a) and round 2 (right)



(a)



(b)

Table 4.3: Non-filers - Impact on Filing Probability

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	CIT	CIT	PIT	PIT	PIT	PIT
Deterrence	0.036^{***}	0.026^{***}	0.038^{***}	0.039^{***}	0.017^{**}	0.017^{**}	0.017^{**}	0.017^{**}
	(0.007)	(0.007)	(0.014)	(0.014)	(0.008)	(0.008)	(0.008)	(0.008)
Compliance costs	0.024^{***}	0.013^{**}	0.007	0.006	0.018^{**}	0.018^{**}	0.018^{**}	0.018^{**}
	(0.006)	(0.006)	(0.013)	(0.013)	(0.008)	(0.008)	(0.008)	(0.008)
De-registration	0.024^{***}	0.014^{**}	0.021	0.021	0.009	0.009	0.009	0.009
	(0.006)	(0.006)	(0.013)	(0.014)	(0.007)	(0.007)	(0.007)	(0.007)
Fiscal Exchange							0.012	0.013^{*}
							(0.008)	(0.008)
Social Norm							0.013^{*}	0.014^{*}
							(0.008)	(0.008)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.069	0.069	0.095	0.095	0.067	0.067	0.067	0.067
R-sq.	0.003	0.016	0.002	0.009	0.001	0.015	0.001	0.015
F Joint Test	0.000	0.000	0.036	0.027	0.015	0.015	0.017	0.014
Observations	22550	22425	4231	4106	18319	18319	21058	21058

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed an income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

pushing non-filers to become visible and still be part of that system. The size of the impact of the deregistration nudge is similar to that of the compliance costs one. So far, the trial's results highlight that both the stick of the threat and the carrot of a more service-oriented approach improve compliance, even if the former performs twice as better.

Considering columns 4 and 6/8, a number of observations can be made. First, companies seem to be affected by the deterrence letter only, with the softer nudges losing significance. The impact of T1 is sizeable (col. 4): 3.9 p.p or 41 per cent of the control group mean. The deterrence nudge significantly improves individuals' compliance as well, even if by a smaller increase (1.7 p.p. or 25%).[55]

[55]The evidence on PIT payers is linked with the survey data collected in Chapter 3 and it is likely to

Second, compliance costs seem to matter more for individuals than companies. PIT payers receiving T2 experience an increase of 1.8 p.p (25%), while companies are not affected. This could be due to the fact that companies might be more likely to have tax accountants and therefore lack of tax knowledge is not a relevant constraint for them. This first evidence links back to the survey data on PIT payers in Chapter 3. From it, it results that for 89 per cent of surveyed non-filers lack of knowledge is an important obstacle to filing. The same figure for active payers is 72 per cent. Likewise, tax knowledge as measured in a tax quiz embedded in the survey results to be strongly significantly correlated with filing. While CIT payers are not covered, the survey evidence is consistent with the large impacts of the compliance costs nudge found in this trial.

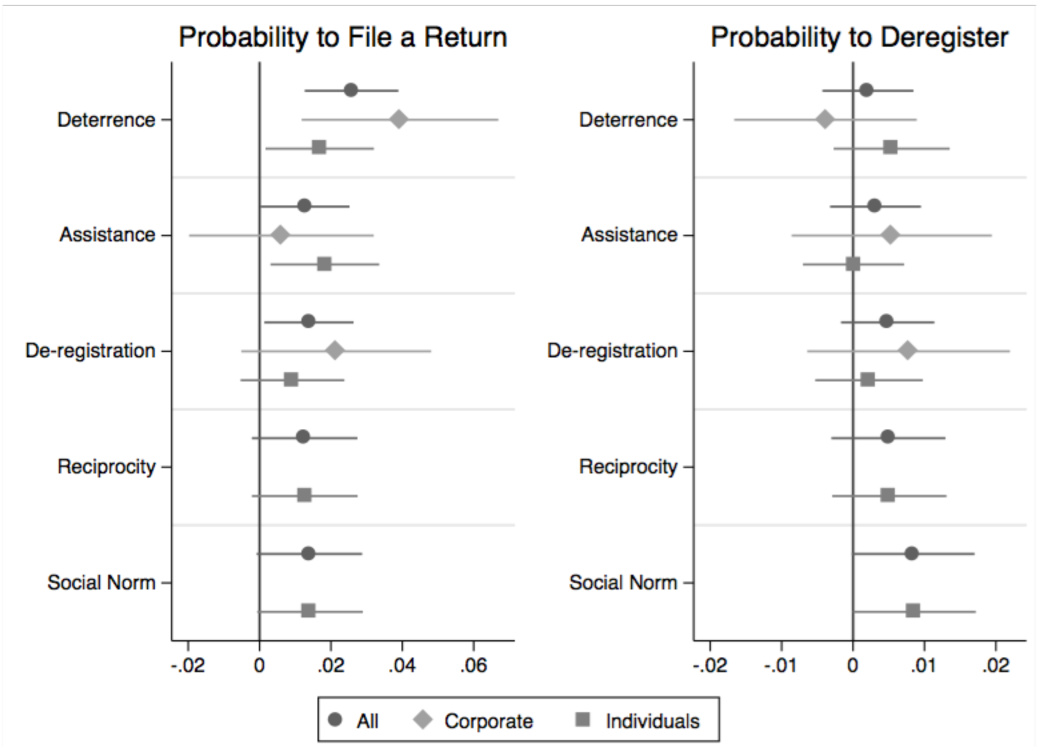
Third, the deregistration nudge remains significant for companies only, probably due to the fact that the exit for the system is more costly for them than for individuals. Fourth, the additional tax morale nudges T4 and T5 both have a weak positive effect on individuals of about the same size (20%), as reported in column 8. This means that while deterrence matters, also moral appeals as well as facilitation are effective, in line with the recently formulated framework of tax compliance presented in Prichard et al. (2019). These three components under study work in a complementary fashion and none of them back-fires.

A second main outcome of the non-filer trial is the probability to deregister from the tax system. As explained in section 4.4.2, a specific nudge, T3, teaches how to navigate the deregistration process. Nudge impacts on this outcome are presented in Appendix Table A8. Overall, less than 2 per cent of control group deregistered, in line with previous years. On top of that, the deregistration nudge increases the exit from the system by 0.7 p.p. (col. 1) or about a 30 per cent increase. However, this estimate turns insignificant when controls are added (col. 2). Surprisingly enough, the social norm message for PIT payers increases exits by 0.9 p.p. or more than 50% of the control group mean. Figure 4.3 also plots impact estimates on deregistration rates. As explained in section 4.3, while lack of knowledge of the process is addressed with the corresponding nudge, the little impact of it may be explained either by the over-optimistic belief to still be operative in the future or, more likely, by the extra monetary and time costs of clearing all past filing and payment

be explained by an increase in the audit probability, rather than improved knowledge of the penalty structure. The survey data shows that audit likelihood is a key determinant of the decision to file, while awareness of the size of penalties is not. More on this in section 4.6.1.

obligations. The fact that the social norm is pushing some individuals to deregister may hint to the fact that, for a certain group of taxpayers, the descriptive norm of non-filing is quite high and suggestive of the idea to leave the system. This goes in contrast with the positive impact of the social norm letter on filing, meaning that for another group of taxpayers that norm is actually low and they are induced to join the filing majority.

Figure 4.3: Treatment effects - Nonfilers



Given the large magnitude of these estimates, one may wonder to what extent they reflect real additional tax revenue. To address this question, it is sensible to enquire into how much tax non-filers are induced to remit when nudged. Appendix Table A9 reports the impact estimates on the log tax declared. The dependent variable is set to zero for those taxpayers who do not file at all. While the pooled results show no significant impacts (col. 1-2), allowing for heterogeneity between corporate and individual taxpayers shows how companies are remitting less taxes than the control group, and sizeably and significantly so when allocated to the compliance costs letter, while individuals react positively.

Unincorporated taxpayers remit 0.06 more log taxes when taught how to file a return, an estimate which is significant at the 10 per cent level. This amounts to about 50 per cent more of what control group individuals remit. In contrast, non-filing companies are reducing taxes in much the same vein as active companies, as it will be described below. This negative reaction could be due to either lack of clarity in the nudge, which in turn irritated recipient taxpayers who respond negatively out of spitefulness, or, less plausibly, to the fact that control group companies are actually remitting more tax to the authority out of confusion and hassle costs (Benzarti, 2015). However, it is not clear a priori why I do not see the same pattern in the control group individuals.

4.5.2 Do nil-filers start to positively file?

When it comes to nil-filers – taxpayers who filed nil returns at baseline – the evidence on the nudges is inconclusive. Table 4.4 reports the impact of the deterrence and deregistration nudges on nil-filers' probability to switch to positive filing. First of all, it can be noticed that 17 per cent of the control group positively filed in 2019, with individuals being three times more likely to switch to positive filing than companies. Relatedly, only 56 per cent (or 1,969) of the nil-filers subsample actually filed a return. The remaining 44 per cent are included anyway in the model with the dependent variable taking a value of 0. These two pieces of evidence underline once again the high instability in filing behaviour over time, which is exacerbated for nil-filers in particular. The fact that a sizeable portion falls back to non-filing while an additional 17 per cent positively files makes it difficult to find precisely estimated impact coefficients, both due to a smaller sample size available (taxpayers actually filing) and a relatively high control group mean.

In terms of impact, the threat letter seems to back-fire, thus producing more nil-filing, even if the coefficient is not precisely estimated. Third, the deregistration nudge has a positive impact (in line with the evidence from Mascagni et al. (2020) in Rwanda), but again not significant. In sum, nil-filing seems to remain a puzzle. As explained in Santoro and Mdluli (2019) nil-filing is largely explained by a context in which firms are not operative yet – and the lack of response to threat may indicate that businesses are actually not operating – and by a sub-optimal equilibrium in which, on one side, the authority encourages filing (even if reporting zero) in order to avoid stiff penalties for non-filing and, on the other side, taxpayers choose to be in a safe position through nil-filing.

Table 4.4: Nil-filers - Impact on Positive Filing Probability

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT
Deterrence	-0.002	-0.002	-0.010	-0.012	0.003	-0.002
	(0.015)	(0.015)	(0.017)	(0.017)	(0.022)	(0.022)
De-registration	0.004	0.003	-0.006	-0.005	0.012	0.007
	(0.016)	(0.015)	(0.017)	(0.017)	(0.022)	(0.022)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.168	0.168	0.075	0.075	0.225	0.225
R-sq.	0.000	0.011	0.000	0.030	0.000	0.017
F Joint Test	0.912	0.958	0.830	0.773	0.856	0.924
Observations	3508	3508	1352	1352	2156	2156

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed a non-zero income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

If anything, a weak impact can be found when considering exits from the system. The deregistration nudge significantly affects PIT payers only, more than doubling exit rates. However, this estimate turns insignificant when controls are added. Results are reported in Appendix Table A10.

4.5.3 Do active taxpayers increase tax liability?

Table 4.5 considers the impact of the deterrence and the fiscal exchange nudge on active taxpayers' probability to increase their tax liabilities, conditional on filing.[56] These taxpayers reported a positive taxable income and tax liability at baseline. Overall, 30 per cent of the control group show an increase in reporting one year later. However, the two nudges are not impactful in increasing taxes (col. 2). The deterrence nudge is not effective

[56] A sizeable 20% of active taxpayers failed to file, equally balanced across CIT-PIT categories. This finding shows once more how the three categories are exchanging taxpayers with each others over time.

in raising tax declared probably due to the fact that recipients do not consider the threat as credible. While perceived enforcement may be effective in increasing tax compliance at the extensive margin (more filing) as shown in Chapter 3, it seems not to be enough to improve the intensive margin of compliance, i.e. tax declared. As a second possible mechanism, additional survey evidence from Chapter 3 also hints to the fact that knowledge of the penalty structure is not a key driver of filing. This is confirmed by the failure of the deterrent letter, which aimed at increasing the salience of the penalties, to have an impact.

The reciprocity letter is also ineffective. CIT payers actually decrease their tax declared once nudged with this letter, while PIT payers are unaffected. This trial then shows how nudges can backfire, at least for certain categories of taxpayers. The fact that the same effect does not hold for individuals may suggest that companies are less satisfied with how tax revenue is used to fund public services, probably because they are more likely to bear the costs of poor infrastructure and other public services in the country. Consistently, survey data gathered in Chapter 3 show that active taxpayers have a lower average satisfaction with six different public services than non-filers. While survey data are available for PIT payers only, the backfire effect for CIT hints to the fact that dissatisfaction is common to companies as well. Likely because non-filers are free-riding on tax-funded public services, they are more satisfied with the their provision. However, the reciprocity nudge generates a reaction from them (Table 4.3), probably due to a sense of guilt (Andreoni et al., 1998). At the same time, active filers who have contributed to the public purse may feel dissatisfied with how taxes are spent and believe that the right thing to do is to retaliate against the tax collector. In section 4.6, I will try to disentangle the null effects to explore any underlying mechanisms at play.

Similarly, the subgroup of taxpayers with VAT discrepancies do not react to the deterrence nudge, with results reported in Appendix Table A11. The null findings of this trial may suggest that the discrepancies found in the cross-checking exercise are legitimate or that the nudged taxpayers do not believe that the authority will credibly enforce the threat communicated in the letter.

Table 4.5: Active - Impact on the Probability to Increase Tax

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT
Deterrence	0.008	0.008	0.015	0.015	0.005	0.006
	(0.012)	(0.012)	(0.023)	(0.023)	(0.014)	(0.014)
Fiscal Exchange	-0.011	-0.012	-0.056**	-0.055**	0.006	0.007
	(0.012)	(0.012)	(0.023)	(0.023)	(0.014)	(0.014)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.300	0.300	0.348	0.348	0.280	0.280
R-sq.	0.000	0.014	0.004	0.013	0.000	0.017
F Joint Test	0.000	0.000	0.000	0.000	0.000	0.000
Observations	8678	8678	2497	2497	6181	6181

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

4.5.4 Treatment on the treated

In this subsection, I exploit the availability of letter delivery reports to calculate LATE estimates (see section 4.4.3). While ITTs are more relevant from a policy perspective since they are informative on the effect size of implementing an intervention that cannot be mandated (Bloom, 2008), LATEs are telling of the actual impact of being exposed to the nudge. As explained in section 4.4.2, delivery reports indicate whether a given letter was returned uncollected or not, while it is not clear whether this is due to failure to delivery or failure of collection from the recipient conditional on delivery. Not surprisingly and consistent with similar studies, letter take-up was far from optimal, 45 per cent. However, quite unexpectedly delivery rates widely varied for companies and individuals. Treatment exposure was almost total for companies, about 93 per cent, and comparable across categories: 93.1 per cent of non-filing, 94.8 per cent of nil-filing and 93.2 per cent of active companies actually collected the letter. In contrast, take-up rates for individuals were stunningly low, 21 per cent: letters were collected by only 12.1 per cent of non-filing, 33.4 per cent of nil-filing and 30.1 per cent of active individuals.

One possible explanation for the low take-up from individuals may consist in the fact that a much larger bulk of letters, about 13,000, was processed for individuals in round 2 (see section 4.4.2). However, nil-filers and actives are equally targeted in round 2 and they show much higher pick-up rates than non-filers. A more reasonable explanation may be due to the fact that individual non-filers are a quite peculiar category of taxpayers in the tax system, quite distinct even from non-filing companies. Contact information for this group may be incomplete or outdated. Likewise, they may feel distant from or not interested into what the authority has to communicate to them. Despite the incomplete implementation of the experiment, delivery rates were not unbalanced across treatment groups, so they are not inducing any bias in the estimation strategy.[57]

Table 4.6 below reports the LATE coefficients for non-filers, while Appendix section A3 displays the results for nil- and positive filers. Unsurprisingly, the already insignificant impact on nil- and positive filers remain so in Tables A12 and A13, with the exception of the backfiring effect of positive filers when nudged with the fiscal exchange nudge. The LATE estimate for T2 does not change much from the ITT one in Table 4.5 due to the almost complete take-up of companies.

It is worth noticing that LATE estimates for non-filers in Table 4.6 remain highly significant and larger in magnitude, especially so for individuals, commensurate with the lower relative probability of treatment. The pooled LATE estimates in col. 2 imply that the deterrence nudge more than double declaration rates compared to the control group, while the two service-oriented nudges provoke a 60 per cent increase. LATE impacts for CIT slightly increase with respect to Table 4.3. In contrast, the estimates for the individual subgroup, for which exposure to the treatment was particularly low, are about 8 times larger than the ITT ones and still significant. For example, both the deterrence and the compliance costs letter increase filing rates by 15-16 percentage points which translate into three times larger filing rates than the control group. Moral appeals also increase filings by about 150 per cent.

[57] Non-filers CIT: T1 92.7%, T2 93.4%, T3 93.3% (p-value 0.807). Non-filers PIT: T1 11%, T2 11.2%, T3 12.1%, T4 12.2% and T5 13.4% (p-value 0.496). Nil-filers CIT: T1 93.3%, T2 96.2% (p-value 0.058). Nil-filers PIT: T1 34.6%, T2 32.1% (p-value 0.321). Active CIT: T1 92.6%, T2 93.8% (p-value 0.291). Active PIT: T1 30.1%, T2 30.2% (p-value 0.949).

Table 4.6: Non-filers - Impact on Filing Probability - LATE

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT	(7) PIT	(8) PIT
Deterrence	0.084^{***}	0.071^{***}	0.042^{***}	0.044^{***}	0.148^{**}	0.146^{**}	0.148^{**}	0.146^{**}
	(0.012)	(0.014)	(0.015)	(0.015)	(0.062)	(0.062)	(0.062)	(0.062)
Compliance costs	0.056^{***}	0.041^{***}	0.008	0.007	0.162^{**}	0.162^{**}	0.162^{**}	0.162^{**}
	(0.012)	(0.014)	(0.015)	(0.015)	(0.064)	(0.064)	(0.065)	(0.064)
De-registration	0.057^{***}	0.043^{***}	0.022	0.023	0.074	0.076	0.074	0.076
	(0.012)	(0.014)	(0.015)	(0.015)	(0.060)	(0.060)	(0.060)	(0.060)
Fiscal Exchange							0.102^{*}	0.105^{*}
							(0.060)	(0.060)
Social Norm							0.101^{*}	0.105^{*}
							(0.054)	(0.054)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.069	0.069	0.095	0.095	0.067	0.067	0.067	0.067
R-sq.	0.005	0.017	0.002	0.009	0.001	0.012	0.003	0.014
F-stat	5333.19	3979.39	4791.95	4607.51	621.37	663.46	386.31	423.43
Observations	22550	22425	4231	4106	18319	18319	21058	21058

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed an income tax return. All coefficients are LATE estimates where actual exposure to the treatment is instrumented by the random assignment to it. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

These results are promising in the sense that they indicate that individual non-filers are particularly responsive to nudges from the authority.[58] If the collection of the letters had been optimal, the intervention would have induced at least three times higher filing rates with respect to the status quo. The revenue authority must take this aspect into account when designing future communication strategies for individual taxpayers. Individual non-filers are responsive, but need to be reached more effectively.

[58]LATE estimates for those additional non-filer outcomes whose ITT estimates are discussed in the next sections are omitted for brevity. The pattern is consistent with what shown in this section, i.e. that coefficients of impact for individuals are 7-8 times larger when accounting for low take-up.

4.5.5 Fiscal externalities

While this trial was specifically targeted at influencing filing behaviour with CIT and PIT in 2019, nudges could also have impacts that spill over other types of taxes or previous declaration periods. Taxpayers might perceive the nudge as a broader increase in enforcement, and thus comply more with other tax obligations as well. On the other hand, they might perceive the intervention as pertaining to income tax only, and increase compliance with the this type of tax only while reducing compliance with other tax to compensate for the income loss. In considering possible fiscal externalities, I recur to a rich set of administrative data on (i) filing for other types of taxes, such as VAT and PAYE, as well as (ii) amendments of previous tax returns.

Results on other types of taxes are not supportive of any spillover effect. If I consider VAT and the active category, Appendix Table A14 shows that both hard- and soft-toned messages increase the probability to file and the amount of VAT due, but not significantly so. This is mostly due to the reduction in power I get when restricting the analysis on VAT-registered active taxpayers, who are only 15 per cent of the total. For the same reason, this exercise is not feasible for non- and nil-filers.[59] When it comes to PAYE, 27 per cent of active taxpayers are also remitting it. At the same time, no significant spillover effects are found in this case as well.

Fiscal externalities can also refer to tax returns corresponding to previous years. Given the data available (see section 4.4.1), I am able to see whether nudged taxpayers amend previous returns in the period 2013-2018. Interestingly, I see a significant impact for non-filers. This category, when nudged, is significantly more likely to file for previous years' returns as well, in addition to the direct effect on 2019 filing (Table 4.3). Table 4.7 below shows the nudge effects on the probability to file at least one previous returns in the period 2013-2018. The deterrence mailing presents the largest results (1.6 pp), amounting to more than half of the control group filing probability of 3 per cent (col. 2). Also, the threat letter seems to work for companies only (col. 4). Likewise, informing about deregistration options has same positive effects (1.1 pp) on filing rates noticed in Table 4.3, again driven by companies. Individuals seem more likely to respond to moral appeals (fiscal exchange and social norms). Notably, the social norm letter doubles the filing rates with respect to

[59]In the study sample, just 6% and 1% of nil and non-filers are registered for VAT, respectively.

non-treated individuals and strongly significantly so (col. 6). A similar impact is found when considering the number of past returns filed after the nudge as outcome, reported in Appendix Table A15.

The fact that the deregistration message is consistently improving past compliance may be due to the fact that taxpayers need to clear their tax duties before deregistering from the authorities. However, this does not seem to be the mechanism in place here. If anything, taxpayers that eventually deregistered after the trial are *less* likely to file previous returns (1.4% vs 3.6% of non deregistered ones) and file on average a smaller number of past returns (0.2 vs 0.8). It is more plausible that the deregistration nudged acted more as reminder to file for past years, as it contained the explicit instruction to do so in order to exit the system (section 4.4.2). The other two taxpayer types do not amend their previous returns in any meaningful fashion.[60]

More in general, non-filers seem to be more malleable to the experiment. It can be the case that, once a non-filer files for the current tax year, her compliance costs for filing for past years are negligible or, as stated before, that this category increases enforcement perceptions once nudged, to the point of complying with past filing obligations as well.

4.5.6 Unexpected impact on secondary outcomes

On top of the main outcomes and the fiscal spillovers described above, nudges seem to have induced an unexpected response in a number of additional, secondary outcomes. First, non-filers increase registrations to the online e-tax filing system. In principle, I would have expected an impact from the compliance costs nudge only, since it also included a reference to the online system. However, as shown in Table A16, T2 is significant without controls in col. 1 but turns insignificant in col. 2 when controls are added. Other nudges such as deregistration and fiscal exchange ones are more likely to push non-filers to register. Most importantly, the threat nudge boosts registration by 1 p.p. over a control mean of 1.4 per cent, and strongly significantly so (col. 2).

Second, active taxpayers are more likely to positively (non-zero) file once nudged. This finding is similar to what Mascagni et al. (2017) get when nudging active taxpayers in

[60]Only 1% and 0.6% of nil and active filers amended previous returns, a very negligible response.

Table 4.7: Non-filers - Impact on Past Filing Probability

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT	(7) PIT	(8) PIT
Deterrence	0.024***	0.016***	0.024**	0.024**	0.007	0.006	0.007	0.006
	(0.005)	(0.005)	(0.011)	(0.011)	(0.005)	(0.005)	(0.005)	(0.005)
Compliance costs	0.013***	0.005	0.001	0.000	0.005	0.005	0.005	0.005
	(0.004)	(0.004)	(0.010)	(0.010)	(0.005)	(0.005)	(0.005)	(0.005)
De-registration	0.019***	0.011**	0.020*	0.021*	0.000	0.000	0.000	0.000
	(0.005)	(0.005)	(0.011)	(0.011)	(0.005)	(0.005)	(0.005)	(0.005)
Fiscal Exchange							0.014**	0.014**
							(0.006)	(0.005)
Social Norm							0.026***	0.026***
							(0.006)	(0.006)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.029	0.029	0.054	0.054	0.027	0.027	0.027	0.027
R-sq.	0.002	0.010	0.002	0.003	0.000	0.005	0.002	0.008
F Joint Test	0.000	0.001	0.042	0.038	0.436	0.449	0.000	0.000
Observations	22550	22425	4231	4106	18319	18319	21058	21058

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed a past (2013-2018) income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Rwanda, even if stronger in significance. Results are reported in Table A17. Non-zero filing was not a first order outcome for active taxpayers (sec. 4.4.3) and it is true that 94 per cent of control group positively file. However, both hard and soft toned letters increase positive filing significantly and much more so for companies, who are also more prone than individuals to fall into zero filing (Santoro and Mdluli, 2019). It results that while nudges do not increase tax liabilities, at least they partially prevent active taxpayers from falling back to zero filing and reporting no tax.

4.5.7 Robustness checks

My results are robust to a number of checks. First, in order to partially control for spillover effects that have shown to matter in similar studies (Drago et al., 2015; Carrillo et al., 2017), I clusterise the error terms at the town level, and allow for error correlation within town. Cameron and Miller (2015) recommend clustering at the highest subnational level, which in the case of Eswatini is the district. However, there are only four districts in the country and the number is too small to have precise estimates. For this reason, I refer to towns, the second highest administrative area and larger in number than the typical cutoff of 50 clusters (Imbens and Rubin, 2015).[61] While this is only an imperfect solution to the bias that information spillovers can produce in a small country like Eswatini, it is also true that in the survey evidence from Chapter 3 shows that only 4% of the sample get information on tax matters from other taxpayers.[62] Also for this reason, I do not clusterise the error terms in my preferred specification discussed in section 4.4.3. To partly confirm that spillovers are not a main threat to the analysis, the main results discussed in the previous section do not change when I clusterise the error terms. Appendix Tables A19, A20 and A21 report the coefficients from the clusterised regressions and compare them with the main ones from my preferred specification. No significant changes are noticed.

Second, for active taxpayers only, I operationalise the main outcome, i.e. tax liability, in alternative fashions: (i) amounts in USD, winsorised at the 99^{th} percentile, (ii) the log of the tax liability with the addition of one unit so to include also cases of zero tax declared,[63] and (iii) the inverted hyperbolic sine transformation (IHS).[64] Using all these alternative tax outcomes the sign of the coefficients in Table 4.5 remain the same. However, none of the estimates is significant, both with more (winsorised amounts) and less (log and IHS)

[61] Non-filers are spread over 63 towns, nil-filers are located in 54 different towns while active in 57.

[62] The majority of the sample, 58%, gets information from newspaper and radio. This is reassuring since the study was not made public. Also, 29% of the sample gets information directly from SRA, which was prepared to handle taxpayers' queries according to a common protocol (see section 4.4.2).

[63] However, by construction, log transformation cannot be performed on negative values of tax declared, which happens for a sixth of total declarations

[64] IHS is defined as $\log(tax + (tax^2+1)^{1/2})$. Except for very small values of y, the inverse sine is approximately equal to log(2y) or log(2)+log(y), and so it can be interpreted in exactly the same way as a standard logarithmic dependent variable. It also accounts for negative values of tax.

skewed distributions. Appendix Table A18 reports the nudge impacts when using the IHS transformation, while other tables are omitted for brevity.

Third, as a last check, I consider only declarations taking place from four weeks before the deadline onwards. As explained in section 4.4.2, letters were mailed about six weeks before deadline. However, there is reason to believe that the experiment has been implemented with delays in the field. This is also shown in the response trend over time in Figure 4.2. While more detailed delivery reports providing information on the exact date of delivery are not available yet, the fact that most responses happens quite after the start of the experiment may hint to the existence of significant delays. By dropping the first two weeks since the start of the experiment, a period in which it is unsure whether letters actually reached taxpayer mailboxes, I can be more confident that filing responses are due to the treatment. Results remain consistent with those from my preferred specification, as shown in Tables A19, A20 and A21.

As a final consideration, it can be noticed that I have one main outcome for each filing category, therefore correcting for multiple hypothesis testing may be not necessary. A second typical correction consists in controlling for the randomisation process, through the randomisation inference approach. However, in this case the samples are fairly large, with the smallest group (nil-filers) consisting of 3,500 units, so that the process of randomisation should not have influenced the estimates I get. On top of that, there are currently no available programmes able to perform randomisation inference with multiple treatments. Therefore, I do not implement such check in this paper.

4.6 Further Analysis and Mechanisms

4.6.1 Being contacted by the authority with *any* type of message

The results presented in section 4.5 are disaggregated by the type of nudge. However, the mere experience of receiving any communication from the revenue authority can plausibly influence compliance decisions as well, since taxpayers may perceive themselves to be more under the agency's radar. In order to test for this hypothesis (also formulated in section

4.4.2), I rerun my main specification by pooling all treatment arms together. Results are consistent with those in section 4.5. Namely, non-filers are mostly affected by receiving any type of message. As shown in Table 4.8, receiving any nudge increases the likelihood to file by 1.6 p.p. (col. 2), and more for companies (2.2 p.p.) than individuals (1.4 p.p.). On average, that implies a 20-24 per cent increase over the control group. Again, nil-filers (Table A22 Panel A) and active (Table A22 Panel B) taxpayers are unaffected.

Table 4.8: Non-filers - Impact on Filing Probability - Any Treatment

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT
Any Treatment	0.023^{***}	0.016^{***}	0.022^{**}	0.022^{**}	0.014^{***}	0.014^{***}
	(0.004)	(0.004)	(0.011)	(0.011)	(0.004)	(0.004)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.069	0.069	0.095	0.095	0.067	0.067
R-sq.	0.002	0.016	0.001	0.008	0.001	0.015
Observations	25289	25164	4231	4106	21058	21058

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed an income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

From Table 4.8 it emerges that it may be beneficial for SRA to just communicate with non-filers in order to improve their filing rates. This is linked to the survey evidence described in Chapter 3, which shows that non-filers are often neglected by the tax authority. PIT non-filers are much less likely than active taxpayers to have had at least one interaction with SRA in a whole year, 29 per cent versus 51 per cent. They are also half as likely to be audited (9% vs 20%) or fined (17% vs 33%). If I interact the experience of an audit with the perceived audit probability, survey data supports the idea that taxpayers significantly increase their perception of enforcement when contacted by the authority.[65] In turn, the same survey data shows that the perceived audit probability is one of the key factors correlated with active filing (Chapter 3). More in general, there is wide evidence on taxpayers having considerable uncertainty over audit probabilities, and almost system-

[65]The perceived audit likelihood is 74% for audited taxpayers versus 54% of non-audited ones.

atically overestimating actual audit probabilities (Andreoni et al., 1998). In conclusion, when contacted by the authority non-filers may update their priors on the risk of audit, even if the actual audit probability does not change for them, and start complying with their filing obligations.

4.6.2 Compliance costs and reminder effect

Considering more in depth the most significant results of this paper, the ones on non-filers, additional evidence on the underlying mechanisms at work can be derived by focusing on timely declarations. I construct a new outcome variable that takes value one for on-time declarations and zero otherwise.[66] Coefficients of impact on on-time declarations are displayed in Appendix Table A23. Overall, about 4 per cent of the control group declare on time. The incremental impact of the nudges differ widely by companies and individuals and by nudge time. First of all, the compliance costs nudge, which is the only nudge clearly stating the declaration deadline, is statistically significant only for individuals. This nudge implies a sizeable 28 per cent increase in on-time filings with respect to the control group. Secondly, the compliance costs nudge is not affecting companies which are instead responding to the deterrence nudge with a 62 per cent increment over the control group mean. Thirdly and unsurprisingly, the moral appeals do not influence on-time filings, despite impacting the probability to declare (Table 4.3).

This results shed light on how the compliance costs nudge improves compliance. The reminder explanation seems to be a reasonable candidate for individuals, which are less likely to benefit from the services of tax accountants than companies. As shown in Chapter 3, knowledge of the deadline is much worse for individual non-filers than active filers.[67] Companies, on the other hand, might be well aware of the filing deadline but strongly respond when the penalty associated with the late filing is made salient. The deterrent letter may have worked as a reminder as well, given that reference was made to failure to file by the deadline. However, only companies seem to link the threat of a fine with the timely filing, while individuals seem to interpret the content differently.

[66] Taxpayers who did not file yet are considered as *late*, since they are potential filers.

[67] Only 16% of non-filing individuals in the survey sample are aware of the fling deadline, compared to 30% of active filers.

4.6.3 Do active filers report more deductions to offset tax due?

The immediate answer seems to be yes. Focusing on active taxpayers only, Table A24 reports how log incomes increases for both treatment arms, and significantly so for T1 deterrence. Therefore it seems that deterrence is disclosing more income, which would otherwise be underreported. This is at odds with previous findings on null (or even negative) impacts on tax due (Table 4.5). It could well be that active taxpayers are reporting more income and offsetting this increases with more costs and deductions. Results from Appendix Table A25 confirm this hypothesis and show that reported log costs increase as well.[68] When it comes to deductions, coefficients are positive as well but not statistically significant. This finding is in line with similar evidence on negative responses from previous studies (Ariel, 2012; Carrillo et al., 2017; Slemrod et al., 2017; Mascagni et al., 2017) and needs further scrutiny. Both researchers and tax administrators could devote more attention to understanding how tax compliance strategies unfold.

4.6.4 Are active filers targeting previous tax declarations?

As another attempt to understand the reaction of active taxpayers, I focus now on the decision to *target* past tax amounts declared. As many as 44% of the experimental group of active taxpayers filing in 2019 report the exact same amount of tax liability as in 2018. This share goes above the expected rate of consistency in tax due given the growth of the Eswatini economy (see section 4.3.1) and is indicative of a strategic behaviour. This phenomenon is labelled as targeting and suggests a situation in which taxpayers may lack accurate record-keeping, thus being unable to correctly measure the tax liability and relying on targeting as a compliance heuristic, or may be strongly affected by cash-flows considerations, according to which they prefer to remit a predictable amount of taxes each year.

This phenomenon has been studied in Rwanda by Tourek (2020), which remains the only quantitative study exploring this behaviour. The evidence from Eswatini confirms the pattern in Rwanda, where the author finds that in about half of all administrative income tax filings over a ten-year period (2008-2017) the amount paid by a taxpayer is identical

[68]Due to limitations in the administrative data, expenses are available for CIT payers only.

to the amount paid by the same taxpayer the year before (Tourek, 2020).

In order to account for this behaviour, I rerun my main specification for active taxpayers removing that 44% targeting past liabilities. Results are reported in Table A26 in col. 3-4 and compared to those from all actives in col. 1-2 (the same reported in Table 4.5). Surprisingly enough, it can be noticed that, once targeters are removed, the deterrence nudge shows a significant positive impact of 3.5 p.p., as expected. The impact corresponds to an increase of about 7% in the probability to remit more taxes over the control group. The fiscal exchange nudge also produces a positive impact, but insignificant.

In conclusion, the enforcement paradigm seems to be impactful for active taxpayers as well – as it was with non-filers –, once I take into account the existence of targeters. This evidence adds a further layer of complexity to the current understanding of tax compliance in low-income countries. Much more evidence is needed in this direction.

4.6.5 Heterogenous treatment effects

As a final exercise, I rerun the main specification in section 4.4.3 by splitting the sample according to different dimensions. As described below, overall average effects hide a number of underlying impact within subgroups. Based on this evidence, tax administrators may want to tailor nudging by those specific subgroups who are more likely to respond.

Filing history A first important dimension is a taxpayer's filing history. It can be argued that the way in which taxpayers have filed in the past is likely to affect present compliance. For this reason, I split the sample in two subgroups according to whether taxpayers are persistent in their behaviour, i.e. they have kept filing in the same way in the 6 years before the experiment (2013-2018) or whether they kept changing their behaviour from one option to another, such as from positive to nil-filing or from non- to positive filing.

For what concerns non-filers, 54 per cent of which are perpetual (sec. 4.3), Table 4.9 shows how perpetuals are less responsive than non-perpetuals when nudged with deterrence, compliance costs, fiscal exchange. Impacts on non-perpetuals are about twice larger in magnitude, suggesting that they are more malleable or easily intimidated by the tax authority, probably because they exaggerate audit probabilities. This evidence can be interpreted as suggesting that perpetual non-filers have much more deeply rooted motivations for their behaviour which are more difficult to be addressed with a nudge.[69] These motivations can include, for example, a Bayesian updating of beliefs about tax debt collection enforcement. This translates into a perception of non-credibility of the nudging strategy, given that previous experience with the tax authority presumably convinced perpetuals that the SRA may not act on its sanction threats. Interestingly, the nudges for which perpetuals react more are the deregistration and social norms ones. The findings on the social norm nudge are supported by the survey data in Chapter 3, in which adherence to a social norm is much more relevant for perpetuals than non-perpetual taxpayers.

At the same time, non-perpetual actives (19% of the total) are much more likely to react than perpetual ones. Appendix Table A27 Panel A reports that the deterrence nudge implies a higher probability to increase taxes, which rises by about 25 per cent the control group mean. Fiscal exchange has positive impact as well, but insignificant. From this evidence two considerations can be formulated. First, non-persistent active taxpayers are probably underreporting their taxes and worth to be targeted. Second, the negative/null results found in Table 4.5 are mostly driven by perpetual actives. This suggests that persistent active taxpayers are irritated by the nudges probably due to the fact that they have always reported positive taxes and consider both the deterrent and the fiscal exchange message as inappropriate. Lastly, nil-filers show no differential responses across the dimension of filing history, as reported in Appendix Table A27 Panel B.

From this exercise, it can be learnt that filing history matters in explaining nudge responses. Taxpayers with a more irregular filing history are apparently more likely to be shaped by communication from the authority, which in turn could exploit this finding by giving more weight to previous filing behaviour when targeting taxpayers.

New registered taxpayers It is fair to believe that newly registered taxpayers (or *young*) can be substantially different than those who have navigated the tax system for a longer period of time and interacted more with the tax authority. Considering the impact of nudges across this dimension further confirms that new taxpayers are a particularly important category to monitor (Mascagni et al., 2019), especially given that the filing

[69]One motivation could be that the firm is closed, for which deregistering might be the optimal choice. However, nudging impacts on de-registrations are not different for perpetuals, and results are omitted for brevity.

Table 4.9: Non-filers - Impact on Filing Probability

	(1) All	(2) All	(3) Perpetual	(4) Perpetual	(5) Non-perp.	(6) Non-perp.
Deterrence	0.036***	0.024***	0.029***	0.018**	0.043***	0.031***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.011)	(0.011)
Compliance costs	0.024***	0.011*	0.017**	0.005	0.033***	0.023**
	(0.006)	(0.007)	(0.007)	(0.007)	(0.011)	(0.011)
De-registration	0.024***	0.012*	0.024***	0.014*	0.023**	0.012
	(0.006)	(0.006)	(0.007)	(0.008)	(0.010)	(0.011)
Fiscal Exchange	0.010	0.013*	-0.001	0.002	0.023*	0.026**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.013)	(0.013)
Social Norm	0.012	0.014*	0.017*	0.021**	0.004	0.007
	(0.008)	(0.008)	(0.009)	(0.009)	(0.012)	(0.012)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.069	0.069	0.044	0.044	0.098	0.098
R-sq.	0.002	0.016	0.002	0.034	0.002	0.024
Observations	25289	25164	13557	13465	11732	11699

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed an income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of perpetual and non-perpetual taxpayers, *Perpetual* indicates non-filers who have been failing to file since registration with the authority, while *Non-perp.* stands for taxpayers who have alternated non-filing with other filing behaviour (nil-filing, positive filing).

behaviour in the first year is likely to influence future behaviour (Dunning et al., 2017; Mascagni et al., 2019).

As reported in Table A28 Panel A, results for non-filers (5% are new) are mostly driven by old taxpayers, which are strongly affected by treatments T1 to T4. Given the importance of tax education from similar settings (Chapter 2), the fact that new taxpayers do not benefit from the compliance costs treatment is concerning. It could well be that the education nudge was not enough to address the knowledge barriers young taxpayers may face. Surprisingly, fiscal exchange backfires for young non-filers (-6.4 p.p.). This can also explain the rather weak effects found for this type of nudge in Table 4.3. Conversely, new taxpayers are mostly influenced by the social norm message (11.5 p.p.), which almost doubles filing rates. Social norms do not work for older taxpayers. Interestingly, the only significant result in the whole trial on nil-filers comes from new taxpayers (see Appendix Table A28 Panel B): the deterrence nudge actually backfires with them (-0.09 p.p.) and significantly so, while is positive but insignificant for older nil-filers, thus summing up to the insignificant overall impacts (Table 4.4). New active taxpayers do not show any differential impact with respect to older ones (Table A28 Panel C).

In sum, it is important for the tax authority to carefully choose the correct communication strategy with newly registered taxpayers. For example, tax knowledge could be increased for this subcategory by recurring to more intensive intervention than a letter, such as in-person trainings (Chapter 2). Deterrence is unlikely to be effective and could even crowd-out any intrinsic willingness to comply. At the same time, more research is needed in trying to understand why different types of soft-toned behavioural theories, such as fiscal exchange and social norms, work differently for this category.

Urban and rural taxpayers Another dimension of heterogeneity is taxpayers' location. Disentangling the main impacts by an urban/rural indicator, it emerges that, while for nil-filers and active taxpayers there are no significantly different patterns (Appendix Table A29), nudges have significantly different impact for non-filers, as reported in Table 4.10 below.

First, deterrence is more impactful in rural (3.2 p.p.) rather than urban (2 p.p.) areas, even if highly statistically significant in both categories. This may suggest that taxpayers in remote areas are more likely to see the threat as credible perhaps due to fewer interactions with the revenue authority. Survey data confirms this hypothesis: urban taxpayers are 43 per cent likely to have any type of interactions with the SRA while rural taxpayers are 35 per cent likely.[70] It is also true that revenue authorities channel limited auditing resources to those areas who are considered to be more profitable, such as urban centres. However, this study shows how a cheap intervention can easily reach more neglected taxpayers and still nudge them to comply.

[70]For example, taxpayers in Hhohho district, the district of the capital Mbabane, are 29 per cent likely to receive an audit, while taxpayers in more remote districts such as Lubombo and Shiselweni are about 5% likely only.

Second, urban taxpayers are more induced to file when receiving the compliance costs (1.7 vs 0.3 p.p.), deregistration (2 vs 0.2 p.p.) and social norm (2.3 vs 0.1 p.p.) nudges. Rural taxpayers are totally unaffected by messages other than deterrence. Importantly, the social norm nudge, that was not strongly significant in the Table 4.3, now turns significant at 5% level and larger in magnitude for urban taxpayers, inducing a 30 per cent increase in the filing rate with respect to the control group.

Table 4.10: Non-filers - Impact on Filing Probability

	(1) All	(2) All	(3) Urban	(4) Urban	(5) Rural	(6) Rural
Deterrence	0.036^{***}	0.024^{***}	0.030^{***}	0.020^{**}	0.047^{***}	0.032^{***}
	(0.007)	(0.007)	(0.008)	(0.008)	(0.011)	(0.011)
Compliance costs	0.024^{***}	0.011^{*}	0.028^{***}	0.017^{**}	0.018^{*}	0.003
	(0.006)	(0.007)	(0.008)	(0.008)	(0.010)	(0.010)
De-registration	0.024^{***}	0.012^{*}	0.030^{***}	0.019^{**}	0.016	0.002
	(0.006)	(0.006)	(0.008)	(0.008)	(0.010)	(0.011)
Fiscal Exchange	0.010	0.013^{*}	0.014	0.016	0.006	0.008
	(0.008)	(0.008)	(0.010)	(0.010)	(0.012)	(0.012)
Social Norm	0.012	0.014^{*}	0.019^{*}	0.023^{**}	0.000	0.001
	(0.008)	(0.008)	(0.010)	(0.010)	(0.011)	(0.011)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.069	0.069	0.065	0.065	0.075	0.075
R-sq.	0.002	0.016	0.003	0.015	0.003	0.020
Observations	25289	25164	15168	15082	10121	10082

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed an income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of urban and rural taxpayers, *Urban* indicates non-filers located in urban areas, while *Rural* stands for taxpayers are located in rural areas.

This exercise is informative in the way it suggests that urban and rural taxpayers are different in the nature of the motivational inducements that are more likely to have an impact. The fact that rural taxpayers are not influenced by non-deterrent messages is a finding that could possibly be explored in future research.

Business size For what concerns active taxpayers, it is possible to divide them by business size, proxied by declared income at baseline.[71] Taxpayers in the top decile of the income distribution are categorised as large and compared to taxpayers in the remaining deciles. These taxpayers are highly relevant in terms of revenue generation. At baseline, the top income decile remitted a twelve times larger income tax than the remaining nine deciles, or US$4,367 vs US$361. In other terms, as much as 60% of total income tax at baseline was raised by the top-income decile. For this reason, it is important to realise that large taxpayers are reducing their tax due when nudged with any treatment, and significantly so when receiving the fiscal exchange one (see Appendix Table A30). The backfiring effect of the fiscal exchange treatment noticed in Table 4.5 is totally driven by top-income taxpayers. Small and medium taxpayers either slightly increase reported liabilities (with the deterrence nudge) or do not respond at all (with the fiscal exchange one).

This result can be interpreted as larger taxpayers being more politically relevant and therefore being in a better position to reduce tax liability if frustrated by a wrong message. Large taxpayers enjoy more bargaining power in negotiating with the government (Giertz and Mortenson, 2014), and are more likely to engage in aggressive tax planning schemes and recruit the services of financial experts to reduce their tax liability (Tanzi, 2012). Also, this evidence is in line with the main results in Table 4.5. Companies are more likely to be large and get irritated by the fiscal exchange nudge if they are not satisfied with the provision of public services (such as infrastructures) on which their productivity depends on. Lastly, large taxpayers are usually more targeted by audits and when receiving a non-deterrent message that appeals to national development they might infer that, as Bardach (1989) puts it, "the enforcement system cannot cope and must resort to rhetoric instead".

Individual taxpayers' demographics As a final exercise, I consider individual (PIT) taxpayers for whom demographic characteristics are available from administrative data, such as age and marital status.[72] While active taxpayers show very little heterogeneity, with some only descriptive evidence that nudges to older taxpayers are more likely to

[71]This exercise is not possible for non and nil-filers, since their baseline income is either missing or zero at baseline.

[72]Demographics are not available for CIT payers. Also, non all PIT payers have valid information. Age is available for 86% of the taxpayer population, while civil status is known for 53% only.

backfire (Table A31 Panel C),[73] and nil-filers do not react differently by age (Table A31 Panel B), more informative trends come up when looking at non-filers (Table A31 Panel A). With respect to the main results in Table 4.3, older taxpayers present larger effects of the fiscal exchange letter (2 vs 0.2 p.p. of younger).[74] Conversely, the significance of the social norm nudge is mostly due to younger individuals. There is also an only descriptive evidence that older taxpayers react more when receiving the compliance costs nudge, probably due to the larger costs they face ex-ante.[75] Being married correlates with age and therefore the results for married taxpayers resemble those of older ones: married non-filers respond more to reciprocity appeals and assistance with filing a return, as reported in Appendix Table A32 Panel A. However, these figures should be considered with more caution given that marital status is available for 53% of the PIT sample only and cannot be used for the nil-filer subsample, given its small size to begin with. Results for active taxpayers are not significantly different across subgroups and reported in Appendix Table A32 Panel B.

4.7 Conclusions and policy recommendations

This paper contributes to the large literature on tax compliance, and specifically to the burgeoning literature on behavioural responses to tax nudges. Despite the presence of several recent studies on tax nudges, the focus has been mostly cast upon developed countries, while little knowledge has been produced from sub-Saharan Africa. In collaboration with the Eswatini Revenue Authority, this study implements a nation-wide nudging experiment by targeting different categories out of the universe of 40,000 corporate and personal income taxpayers, labelled as non-filers, nil-filers and active taxpayers, and causally estimating the impact of nudges on a range of compliance outcomes. The content of nudges is built on the main theoretical formulations on behavioural tax compliance, some of whom are confirmed

[73] This is consistent with the fact that larger and incorporated taxpayers are likely to backfire, since the owner of these businesses are likely to be older.

[74] Older taxpayers are defined as those having an age larger than the median in the population (about 50).

[75] Survey data in Chapter 3 shows that the average index score for the young group (below 50 years of age) is of 1.63 out of 5. The corresponding figure for older taxpayers is 1.47, or 10% less. The difference in the distribution of knowledge is statistically significant at the 5% level.

in the field. By doing so, this study adds to the limited existing evidence on the drivers of compliance in SSA, by exploring a country (Eswatini) that has not been studied before. Furthermore, it is the first study of its type targeting three different filing categories at the same time, on top of including both companies and individuals (as already done in Mascagni et al. (2017)). Much of the relevant literature, with few exceptions, has focused on positive filers only.

The results of the RCT present a nuanced and multi-faceted picture of tax compliance in Eswatini. While non-filers substantially increase their filing once nudged, submitting returns for both the current and previous tax years, nil-filers are not responsive – still remaining a puzzle as discussed in Santoro and Mdluli (2019) – and active taxpayers show perverse (but fully rational) reactions, mostly driven by top-income entities.

Despite the poor impact at the intensive margin of compliance, the extra revenue generated by the experiment well compensate for the costs of implementation of about US$18,000.[76] In order to calculate the overall revenue gain associated with the experiment, I consider the statistically significant ITT results (section 4.5). First, non-filers in the treatment groups are more likely to file than those in the control group (Table 4.8). About 235 extra non-filers actually filed a return thanks to the experiment. These non-filers remitted a total of US$109,303. Second, the same non-filers, once nudged, started filing for previous years as well (Table 4.7). The extra revenue generated by these extra taxpayers (187 in number) amounts to US$25,505. Lastly, the nudges significantly pushed a small group of active taxpayers to continue to file positive tax, while they would have remitted zero tax in the absence of the intervention (Table A17). These 86 extra active taxpayers declared a total of US$69,269. Therefore, the extra revenue gains causally associated with the experiment sums up to US$0.2 million, thus making the intervention highly cost-effective, with an overall cost-benefit ratio of about 1:11.

The aggregate revenue gains are small when compared to similar studies in SSA (Mascagni et al., 2017). However, these studies focused on active taxpayers only and thus were more likely to generate larger revenue. The cost-benefit ratio becomes more consistent when compared to nudge studies targeting non-filers only (Kettle et al., 2016; Brockmeyer et al.,

[76]Total costs include postage fees, envelopes, printing and additional fees paid to SRA staff for their hours of work spent on the implementation.

2019). At the same time, the cost-benefit ratio above is likely to represent a lower bound of the actual one as it is difficult to quantify the benefits associated with improved compliance at the extensive margin. Beyond immediate revenue considerations, the trial tackled important detrimental fiscal consequences of non-filing such as horizontal inequality, perceptions of unfairness and production inefficiencies. On top of that, non-filers start to be visible to the authority and share valuable information with the SRA. All these impacts, despite being crucial in building a culture of compliance, are almost impossible to quantify.

This study has clear policy implications for the Eswatini Revenue Authority and for other tax agencies in low- and middle-income countries. First, non-filers represent an easy target for the authority. By simply contacting them with any type of nudge, the authority could achieve a 20-24 per cent increase in filing (Table 4.8). This impact could rise to about 80 per cent if all recipients are actually exposed to the treatment. This evidence links back to the survey data in Chapter 3 according to which non-filers are less likely to be reached by the authority. This result gains economic significance given that non-filers represent the majority of registered taxpayers in Eswatini, as well as in other SSA countries. In particular, the authority could implement a three-tiered approach in targeting non-filers, as formulated in the conceptual framework of voluntary compliance in Prichard et al. (2019) as well as in Alm et al. (2010). In this setting, three complementary paradigms, i.e. deterrence, taxpayer assistance and moral appeals are equally important for a tax administration that wishes to encourage voluntary compliance. The parallel study of Chapter 3 consistently documents that these three strategic pillars are strongly correlated with the probability of filing in Eswatini. Relatedly, the deterrent pillar of the framework seems to be more effective in boosting compliance from companies, which seem to pay more attention to the pecuniary motives, while taxpayer assistance seems a viable solution to increase compliance of individuals (Table 4.3). The latter strongly increase the probability to file on-time, on top of the likelihood of lodging a return at all, when nudged with the compliance costs nudge (Table A23). In this sense, short nudges through SMSs, automatically sent a little before the deadline, would act as an impactful reminder for individuals.

Second, alternative approaches should be pursued in increasing tax liabilities for nil- and active filers. For the former, this study contributes to the almost non-existent evidence on how to address nil-filing. The external validity of this trial is reinforced by the results of a similar RCT from Rwanda (Mascagni et al., 2020), where both the vast majority of nil-filers continued to do so after being nudged (Table 4.4) and neither deterrence or deregistration messages produced any impact. The trial on Eswatini nil-filers is ineffective as well: deterrence significantly backfires for newly registered taxpayers (Table A28) and that may suggest that other non-conventional messages are more likely to work with this category. The argument can be made that these taxpayers are actually not operating yet and legitimately filing nil (Santoro and Mdluli, 2019), but further research is needed to shed light on this category.

For active taxpayers, this trial shows how difficult it could be to build on non-deterrent motivations, such as reciprocity (Table 4.5), in a country where satisfaction with public services is low (Chapter 3). Alternative messages could be based on comparisons to peers, which proved to play a role in the parallel survey study of Chapter 3. At the same time, the authority could also explore alternative enforcement channels. While Chapter 3 documented that standard audit perceptions are linked with the extensive margin of compliance, i.e. with filing a return, this trial suggests that they are unlikely to be effective at the intensive margin as well, i.e. increasing tax due. If anything, active taxpayers seem both to update their perceived enforcement likelihood and strengthen their willingness to contribute to national development to the point of being more likely to positively file (avoiding nil returns, Table A17), but not as much as to increase their tax due. More credible and salient enforcement messages than the one-page letter of this study could be tested on this category, in much the same vein as already explored in in Latin America (Pomeranz, 2015; Bergolo et al., 2019). At the same time, the presence of targeters largely dilutes the impact of the nudges. As a matter of fact, deterrence is pushing active filers to remit more taxes once targeters are removed – as expected. It could be derived that the SRA should think about a strategy to address tax targeting - for example, by just informing these taxpayers that it is aware and suspicious of their behaviour.

Third, the inherent heterogeneity in this large nationwide experiment should be taken into account by the SRA when designing its compliance strategy. The authority could think of more tailored communication strategies building on the results presented in this paper so to focus on those subgroups that are more likely to respond, without wasting (the already limited) resources on those taxpayers whose filing behaviour is more resistant to change. Immediate recommendations may consist in, for example: (i) implementing more incisive interventions for both perpetual and newly registered taxpayers, such as direct contact from the authority for the former and workshops and trainings for the latter, which also proved to be effective in similar contexts (Chapter 2);[77] (ii) exploiting more standard deterrent messages in rural areas while appealing to moral factors in more urbanised settings, (iii) targeting top-income taxpayers with more sophisticated deterrent measures, such as deploying the most qualified auditors, and (iv) acknowledging that, more in general, companies and individuals are very likely to be driven by different sets of motivations, values, perceptions – as discussed further in the concluding chapter of this thesis – and thus implementing different compliance strategies for different taxpayer types.

Lastly, the nudging method could be improved as well and the authority could find better ways to reach certain filing categories. While letter delivery was successful for companies, it was dismally low for individuals, especially non-filers. Much larger compliance gains could have been achieved if individual taxpayers were reached adequately. One way forward could be to update the contact information (especially location and mail box) of individual taxpayers in the registry. A second option would explore alternative tools such as emails or SMSs.[78] More specifically, emails have the advantage of being cheaper and more easily to track but also inevitably rely on the quality of taxpayer contact information. Before implementing this trial I ran a quality check on the taxpayer registry, realising that email addresses were available in 25 per cent of the cases and poorly representative of the entire population.[79] The authority could devote more effort in cleaning its registry which, in turn, could serve as a valid basis for reaching individual taxpayers in future nudging exercises. Future experiments could test alternative methods characterised by higher salience

[77] Relatedly, the fact that non-perpetual active and non-filers are more responsive may suggest that the authority could benefit from an automated system tracking filing behaviour and flagging cases in which a previously active taxpayer stop filing or, conversely, a previous non-filer suddenly starts filing. In the latter case, as this trial shows, the the tax authority could investigate whether previous returns also should have been filed.

[78] In principle, SMSs represent a viable option as well, but they are limited in size and do not allow for sending more official documentation. However, they could potentially be used to send short deadline reminders, a mechanism that seems to be at play with non-filers.

[79] Active, large and urban taxpayers were more likely to have a valid email address.

of the delivery device and add to the corresponding limited literature (Doerrenberg and Schmitz, 2015; Ortega and Scartascini, 2016a; Mascagni et al., 2017).

I conclude by highlighting three directions for future research. First, it will be important to track nudged taxpayers over time and observe their filing behaviour in the following years, so to gauge whether the gains are long-lived. Tax returns data for future years will be analysed so to add to the growing evidence on mid-to-long term effects of letter interventions (Allcott and Rogers, 2014; Cronqvist et al., 2018; Brockmeyer et al., 2019).

Second, this study produced both ITT and LATE estimates, with the second being measurable thanks to the delivery reports shared by the national post office. However, future data on the actual date of delivery to the recipient's mailbox will hopefully be provided by the post office and used to more carefully calculate treatment effects.

As a final point, future research could dig deeper in disentangling different components at work within a given treatment letter. For example, the deterrence arm could work either through increases in the salience of the penalty structure or in the perceived probability of audits (Bergolo et al., 2019). At the same time, the combination of different treatments, such as those reducing compliance costs and those appealing to national development, could be more effective in increasing compliance. For reasons of limited sample size, this study was unable to test additional, more specific or combined, treatments. There is much exciting future research to be carried on in this direction.

Appendices

A1 Randomization balance

Table A1: Balance Table - CIT Non-filers

	(1)		(2)		(3)		(4)			T-test	
	Control		T1 Deterrence		T2 Costs		T3 Dereg.			P-value	
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(1)-(4)
Hhohho	1065	0.45	1055	0.45	1056	0.45	1055	0.45	0.95	0.96	0.95
		(0.02)		(0.02)		(0.02)		(0.02)			
Lubombo	1065	0.10	1055	0.10	1056	0.10	1055	0.10	1.00	0.95	1.00
		(0.01)		(0.01)		(0.01)		(0.01)			
Manzini	1065	0.42	1055	0.42	1056	0.42	1055	0.42	0.92	0.94	0.92
		(0.02)		(0.02)		(0.02)		(0.02)			
Shiselweni	1065	0.04	1055	0.03	1056	0.03	1055	0.03	0.66	0.66	0.66
		(0.01)		(0.01)		(0.01)		(0.01)			
First year	1065	0.07	1055	0.07	1056	0.07	1055	0.07	0.72	0.78	0.72
		(0.01)		(0.01)		(0.01)		(0.01)			
# returns	1065	4.23	1055	4.21	1056	4.14	1055	4.25	0.73	0.22	0.83
		(0.05)		(0.05)		(0.05)		(0.05)			
Perpetual	1065	0.49	1055	0.49	1056	0.54	1055	0.50	1.00	0.03**	0.76
		(0.02)		(0.02)		(0.02)		(0.02)			
Trading	1065	0.35	1055	0.33	1056	0.36	1055	0.33	0.42	0.49	0.39
		(0.01)		(0.01)		(0.01)		(0.01)			
F-test of joint significance (p-value)									0.98	0.41	0.99
F-test, number of observations									2120	2121	2120

Notes: The value displayed for t-tests and F-tests are p-values. ***, **, and * - 1, 5, and 10 percent level.

Table A2: Balance Table - PIT Non-filers

	(1)		(2)		(3)		(4)		(5)		(6)				T-test		
	Control		T1 Deterrence		T2 Costs		T3 Dereg.		T4 FE		T5 Norms				P-value		
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(1)-(4)	(1)-(5)	(1)-(6)
Sole trader	14201	0.17	1376	0.18	1373	0.16	1369	0.17	1373	0.17	1366	0.19	0.67	0.33	0.87	0.90	0.19
		(0.00)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)					
Hhohho	14201	0.36	1376	0.37	1373	0.36	1369	0.36	1373	0.36	1366	0.35	0.63	0.96	0.93	1.00	0.25
		(0.00)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)					
Lubombo	14201	0.16	1376	0.16	1373	0.16	1369	0.15	1373	0.17	1366	0.17	0.84	0.97	0.41	0.70	0.69
		(0.00)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)					
Manzini	14201	0.39	1376	0.39	1373	0.39	1369	0.40	1373	0.37	1366	0.40	0.63	0.74	0.40	0.41	0.26
		(0.00)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)					
Shiselweni	14201	0.09	1376	0.08	1373	0.08	1369	0.09	1373	0.10	1366	0.09	0.17	0.53	0.82	0.36	0.63
		(0.00)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)					
First year	14201	0.06	1376	0.05	1373	0.05	1369	0.06	1373	0.06	1366	0.05	0.48	0.57	0.83	0.98	0.33
		(0.00)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)					
# returns	14201	4.36	1376	4.39	1373	4.35	1369	4.37	1373	4.37	1366	4.41	0.39	0.76	0.74	0.78	0.10*
		(0.01)		(0.03)		(0.03)		(0.03)		(0.03)		(0.03)					
Perpetual	14201	0.54	1376	0.54	1373	0.55	1369	0.55	1373	0.54	1366	0.53	0.99	0.49	0.81	0.58	0.28
		(0.00)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)					
Trading	14201	0.08	1376	0.08	1373	0.07	1369	0.08	1373	0.08	1366	0.09	0.47	0.18	0.50	0.97	0.19
		(0.00)		(0.01)		(0.01)		(0.01)		(0.01)		(0.01)					
F-test of joint significance (p-value)													0.81	0.84	0.99	0.99	0.48
F-test, number of observations													15577	15574	15570	15574	15567

Notes: The value displayed for t-tests are p-values. The value displayed for F-tests are p-values. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table A3: Balance Table - CIT Nil-filers

		(1) Control		(2) T1 Deterrence		(3) T2 Dereg.	T-test P-value	
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)
Hhohho	453	0.47	448	0.47	451	0.47	0.97	0.89
		(0.02)		(0.02)		(0.02)		
Lubombo	453	0.08	448	0.08	451	0.08	0.94	0.98
		(0.01)		(0.01)		(0.01)		
Manzini	453	0.41	448	0.42	451	0.42	0.84	0.85
		(0.02)		(0.02)		(0.02)		
Shiselweni	453	0.04	448	0.03	451	0.04	0.74	0.87
		(0.01)		(0.01)		(0.01)		
First year	453	0.14	448	0.14	451	0.14	0.98	0.90
		(0.02)		(0.02)		(0.02)		
# returns	453	3.97	448	4.00	451	3.99	0.80	0.90
		(0.09)		(0.08)		(0.09)		
Perpetual	453	0.79	448	0.76	451	0.78	0.30	0.78
		(0.02)		(0.02)		(0.02)		
Trading	453	0.32	448	0.31	451	0.36	0.75	0.27
		(0.02)		(0.02)		(0.02)		
F-test of joint significance (p-value)							0.99	0.98
F-test, number of observations							901	904

Notes: ***, **, and * - 1, 5, and 10 percent level.

Table A4: Balance Table - PIT Nilfilers

	(1) Control		(2) T1 Deterrence		(3) T2 Dereg.		T-test P-value	
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)
Sole trader	729	0.41	714	0.41	713	0.41	0.99	0.94
		(0.02)		(0.02)		(0.02)		
Hhohho	729	0.36	714	0.36	713	0.37	0.98	0.96
		(0.02)		(0.02)		(0.02)		
Lubombo	729	0.14	714	0.14	713	0.14	0.95	0.98
		(0.01)		(0.01)		(0.01)		
Manzini	729	0.42	714	0.43	713	0.42	0.94	0.97
		(0.02)		(0.02)		(0.02)		
Shiselweni	729	0.07	714	0.07	713	0.07	1.00	0.84
		(0.01)		(0.01)		(0.01)		
First year	729	0.05	714	0.04	713	0.04	0.20	0.20
		(0.01)		(0.01)		(0.01)		
# returns	729	3.84	714	3.92	713	3.97	0.33	0.09*
		(0.06)		(0.05)		(0.05)		
Perpetual	729	0.52	714	0.47	713	0.49	0.07*	0.27
		(0.02)		(0.02)		(0.02)		
Trading	729	0.22	714	0.23	713	0.22	0.69	0.82
		(0.02)		(0.02)		(0.02)		
F-test of joint significance (p-value)							0.82	0.88
F-test, number of observations							1443	1442

Notes: The value displayed for t-tests and F-tests are p-values. ***, **, and * - 1, 5, and 10 percent level.

Table A5: Balance Table - CIT Active

	(1) Control		(2) T1 Deterrence		(3) T2 FE		T-test P-value	
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)
Turnover 2018	1041	12.44	1026	12.39	1027	12.54	0.79	0.54
		(0.13)		(0.13)		(0.12)		
Tax 2018	1041	1.83	1026	1.36	1027	1.66	0.14	0.59
		(0.22)		(0.23)		(0.22)		
Hhohho	1041	0.44	1026	0.44	1027	0.44	0.91	0.93
		(0.02)		(0.02)		(0.02)		
Lubombo	1041	0.09	1026	0.09	1027	0.09	0.96	0.95
		(0.01)		(0.01)		(0.01)		
Manzini	1041	0.42	1026	0.43	1027	0.43	0.85	0.83
		(0.02)		(0.02)		(0.02)		
Shiselweni	1041	0.05	1026	0.05	1027	0.05	0.52	0.52
		(0.01)		(0.01)		(0.01)		
First year	1041	0.09	1026	0.08	1027	0.08	0.83	0.89
		(0.01)		(0.01)		(0.01)		
# returns	1041	4.61	1026	4.69	1027	4.68	0.31	0.42
		(0.05)		(0.05)		(0.05)		
Perpetual	1041	0.80	1026	0.79	1027	0.81	0.35	0.53
		(0.01)		(0.01)		(0.01)		
Trading	1041	0.38	1026	0.37	1027	0.38	0.74	0.77
		(0.02)		(0.02)		(0.02)		
F-test of joint significance (p-value)							0.85	0.98
F-test, number of observations							2067	2068

Notes: The value displayed for t-tests and F-tests are p-values. ***, **, and * - 1, 5, and 10 percent level.

Table A6: Balance Table - PIT Active

	(1) Control		(2) T1 Deterrence		(3) T2 FE		T-test P-value	
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)
Turnover 2018	2566	3.00	2548	2.76	2551	2.77	0.05**	0.06*
		(0.09)		(0.09)		(0.09)		
Tax 2018	2566	1.92	2548	1.51	2551	1.80	0.02**	0.48
		(0.12)		(0.12)		(0.12)		
Sole trader	2566	0.37	2548	0.37	2551	0.37	0.95	0.97
		(0.01)		(0.01)		(0.01)		
Hhohho	2566	0.42	2548	0.42	2551	0.42	0.97	0.95
		(0.01)		(0.01)		(0.01)		
Lubombo	2566	0.12	2548	0.12	2551	0.12	0.94	0.96
		(0.01)		(0.01)		(0.01)		
Manzini	2566	0.37	2548	0.37	2551	0.37	0.98	0.99
		(0.01)		(0.01)		(0.01)		
Shiselweni	2566	0.09	2548	0.09	2551	0.09	0.98	0.97
		(0.01)		(0.01)		(0.01)		
First year	2566	0.02	2548	0.02	2551	0.02	0.50	0.50
		(0.00)		(0.00)		(0.00)		
# returns	2566	4.10	2548	4.14	2551	4.09	0.32	0.74
		(0.03)		(0.03)		(0.03)		
Perpetual	2566	0.82	2548	0.81	2551	0.81	0.26	0.41
		(0.01)		(0.01)		(0.01)		
Trading	2566	0.21	2548	0.21	2551	0.20	0.84	0.31
		(0.01)		(0.01)		(0.01)		
F-test of joint significance (p-value)							0.17	0.58
F-test, number of observations							5114	5117

Notes: The value displayed for t-tests and F-tests are p-values. ***, **, and * - 1, 5, and 10 percent level.

Table A7: Types of Messages from SRA

Notice	Channel	Content	From	Timing
1. Annual notice for filling for income tax, VAT, PAYE	Newspaper, billboards	A call to all eligible taxpayers to submit income tax returns. The message is comprehensive and includes information on submission deadlines.	Domestic Tax (DT)	Approx. 5 days, but it varies. Billboards are placed throughout the country for a specific duration.
2. Topical announcements	Radio, TV	Topical matters such as VAT with a stress on submission deadlines	DT Customs	Radio once a week; TV - a certain number of slots are purchased and presentations are on an ad hoc basis. Mostly it depends when the need arises.
3. Notice of assessment	Letter	Advice taxpayer on the outcome of the tax assessed and also mention the right to object/dispute	DT	Immediately upon closure of the tax assessment. One NOA per year for each taxpayer.
4. Debt notice	Letter, phone call, physical visits	Advice taxpayer on the taxes due	DT	(i) Calls 7 days after deadline (ii) Letters of demand: 7 and 14 days after (iii) Final demand letter 14 days after
5. PAYE and VAT non filers	Letter, phone call, email, physical visits	Nudge on un-submitted PAYE and VAT returns	DT	Every month, after deadline. Site visit for defaulter more than 5 to 6 times consecutively.
6. Provisional Tax	letter	Pre-payment of taxes	DT	Twice a year.

A2 Main Results

Table A8: Non-filers - Impact on Deregistration Probability

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT	(7) PIT	(8) PIT
Deterrence	0.004	0.002	-0.005	-0.004	0.006	0.005	0.006	0.005
	(0.003)	(0.003)	(0.007)	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)
Compliance costs	0.005	0.003	0.004	0.005	-0.000	0.000	-0.000	0.000
	(0.003)	(0.003)	(0.007)	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)
De-registration	0.007^{**}	0.005	0.004	0.008	0.002	0.002	0.002	0.002
	(0.003)	(0.003)	(0.007)	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)
Fiscal Exchange							0.005	0.005
							(0.004)	(0.004)
Social Norm							0.009^{**}	0.009^{**}
							(0.004)	(0.004)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.017	0.017	0.026	0.026	0.017	0.017	0.017	0.017
R-sq.	0.000	0.004	0.001	0.018	0.000	0.003	0.000	0.005
F Joint Test	0.070	0.354	0.442	0.328	0.534	0.573	0.227	0.258
Observations	22550	22425	4231	4106	18319	18319	21058	21058

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having deregistered from the tax system. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A9: Non-filers - Impact on Log Tax Declared

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	CIT	CIT	PIT	PIT	PIT	PIT
Deterrence	0.013	0.020	-0.016	-0.012	0.013	0.013	0.013	0.013
	(0.024)	(0.025)	(0.049)	(0.050)	(0.031)	(0.030)	(0.031)	(0.030)
Compliance costs	0.010	0.016	-0.087**	-0.091**	0.063*	0.063*	0.063*	0.062*
	(0.024)	(0.025)	(0.043)	(0.044)	(0.036)	(0.036)	(0.036)	(0.036)
De-registration	0.022	0.028	-0.000	0.000	0.017	0.017	0.017	0.017
	(0.024)	(0.025)	(0.050)	(0.051)	(0.031)	(0.031)	(0.031)	(0.031)
Fiscal Exchange							-0.003	-0.001
							(0.029)	(0.029)
Social Norm							-0.012	-0.012
							(0.028)	(0.028)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.125	0.125	0.158	0.158	0.123	0.123	0.123	0.123
R-sq.	0.000	0.005	0.001	0.007	0.000	0.006	0.000	0.007
Observations	22399	22274	4210	4085	18189	18189	20906	20906

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is the log(tax+1) declared. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A10: Nil-filers - Impact on Deregistration Probability

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	CIT	CIT	PIT	PIT
Deterrence	0.005	0.005	0.005	0.004	0.006	0.005
	(0.004)	(0.004)	(0.009)	(0.009)	(0.004)	(0.004)
De-registration	0.004	0.004	0.000	0.000	0.007*	0.005
	(0.004)	(0.004)	(0.009)	(0.009)	(0.004)	(0.004)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.008	0.008	0.018	0.018	0.003	0.003
R-sq.	0.000	0.010	0.000	0.006	0.001	0.020
F Joint Test	0.439	0.439	0.844	0.844	0.231	0.231
Observations	3508	3508	1352	1352	2156	2156

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having deregistered from the tax system. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A11: VAT discrepancy - Impact on Higher Tax Declared

	(1) All	(2) All	(3) Non-Active	(4) Non-Active	(5) Active	(6) Active
Deterrence	-0.022	-0.005	-0.146	-0.111	-0.016	0.003
	(0.040)	(0.041)	(0.131)	(0.133)	(0.041)	(0.043)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.381	0.381	0.263	0.263	0.390	0.390
R-sq.	0.001	0.011	0.034	0.117	0.000	0.011
Observations	600	600	36	36	564	564

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having increased the income tax remitted. All coefficients are OLS estimates from a LPM. *All* refers to the total of non-active and active taxpayers, *Non-active* indicates non- and nil-filers, while *Active* stands for positive filing taxpayers.

A3 LATE estimates

Table A12: Nil-filers - Impact on Active Filing Probability - LATE

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT
Deterrence	-0.004	-0.003	-0.011	-0.013	0.010	-0.004
	(0.027)	(0.027)	(0.018)	(0.018)	(0.064)	(0.065)
De-registration	0.008	0.005	-0.007	-0.005	0.038	0.021
	(0.027)	(0.027)	(0.018)	(0.017)	(0.069)	(0.068)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.168	0.168	0.075	0.075	0.225	0.225
R-sq.	.	0.011	.	0.029	0.002	0.017
F-stat	785.73	1840.09	3917.22	3924.23	181.52	452.76
Observations	3508	3508	1352	1352	2156	2156

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed a non-zero income tax return. All coefficients are LATE estimates where actual exposure to the treatment is instrumented by the random assignment to it. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

A4 Additional Results

Table A13: Active - Impact on the Probability to Increase Tax Declared - LATE

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT
Deterrence	0.017	0.017	0.016	0.016	0.017	0.017
	(0.024)	(0.024)	(0.025)	(0.025)	(0.044)	(0.044)
Fiscal Exchange	-0.023	-0.023	-0.058**	-0.058**	0.020	0.020
	(0.024)	(0.024)	(0.024)	(0.024)	(0.044)	(0.043)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.300	0.300	0.348	0.348	0.280	0.280
R-sq.	0.001	0.014	0.003	0.012	.	0.017
F-stat	1435.30	3328.82	6483.64	6473.84	483.18	1167.01
Observations	8678	8678	2497	2497	6181	6181

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are LATE estimates where actual exposure to the treatment is instrumented by the random assignment to it. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A14: Active - Impact on VAT outcomes

	(1) Filing	(2) Filing	(3) IHS VAT	(4) IHS VAT
Deterrence	0.023	0.024	0.128	0.120
	(0.022)	(0.022)	(0.502)	(0.516)
Fiscal Exchange	0.013	0.017	0.269	0.158
	(0.023)	(0.022)	(0.493)	(0.507)
Controls	No	Yes	No	Yes
Mean of Y	0.843	0.843	6.745	6.678
R-sq.	0.001	0.034	0.000	0.012
Observations	1578	1578	1302	1192

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. All coefficients are OLS estimates. *Filing* refers to the probability of filing a VAT return while *PIT* stands for the VAT amount remitted.

Table A15: Non-filers - Impact on Number of Past Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	CIT	CIT	PIT	PIT	PIT	PIT
Deterrence	0.043***	0.023*	0.036	0.032	0.019	0.018	0.019	0.018
	(0.012)	(0.012)	(0.026)	(0.026)	(0.015)	(0.015)	(0.015)	(0.015)
Compliance costs	0.021*	0.003	0.011	0.009	-0.000	-0.000	-0.000	-0.000
	(0.011)	(0.011)	(0.025)	(0.026)	(0.012)	(0.012)	(0.012)	(0.012)
De-registration	0.038***	0.021*	0.048*	0.050*	0.000	0.001	0.000	0.001
	(0.012)	(0.012)	(0.027)	(0.028)	(0.013)	(0.013)	(0.013)	(0.013)
Fiscal Exchange							0.024*	0.024*
							(0.014)	(0.014)
Social Norm							0.028**	0.028**
							(0.013)	(0.013)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.065	0.065	0.107	0.107	0.062	0.062	0.062	0.062
R-sq.	0.001	0.005	0.001	0.002	0.000	0.003	0.000	0.004
F Joint Test	0.000	0.105	0.250	0.266	0.637	0.685	0.143	0.157
Observations	22550	22425	4231	4106	18319	18319	21058	21058

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is a continuous variable for the number of past (2013-2018) returns filed. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A16: Non-filers - Impact on E-tax Registration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	CIT	CIT	PIT	PIT	PIT	PIT
Deterrence	0.014***	0.010***	0.004	0.006	0.004	0.004	0.004	0.004
	(0.003)	(0.003)	(0.008)	(0.009)	(0.004)	(0.004)	(0.004)	(0.004)
Compliance costs	0.008***	0.005	-0.003	-0.003	0.001	0.001	0.001	0.001
	(0.003)	(0.003)	(0.008)	(0.008)	(0.003)	(0.003)	(0.003)	(0.003)
De-registration	0.009***	0.006*	-0.002	0.001	0.001	0.001	0.001	0.001
	(0.003)	(0.003)	(0.008)	(0.008)	(0.003)	(0.003)	(0.003)	(0.003)
Fiscal Exchange							0.007*	0.007*
							(0.004)	(0.004)
Social Norm							-0.003	-0.003
							(0.003)	(0.003)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.014	0.014	0.038	0.038	0.012	0.012	0.012	0.012
R-sq.	0.002	0.007	0.000	0.019	0.000	0.003	0.000	0.004
Observations	22550	22425	4231	4106	18319	18319	21058	21058

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having registered for the e-tax system. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A17: Active - Impact on Positive Filing Probability

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	CIT	CIT	PIT	PIT
Deterrence	0.013^{**}	0.013^{**}	0.023^{**}	0.027^{***}	0.008	0.008
	(0.006)	(0.006)	(0.010)	(0.010)	(0.007)	(0.007)
Fiscal Exchange	0.017^{***}	0.017^{***}	0.026^{***}	0.028^{***}	0.013^{*}	0.012^{*}
	(0.006)	(0.006)	(0.009)	(0.010)	(0.007)	(0.007)
Baseline income		0.001^{***}		0.000		0.002^{***}
		(0.000)		(0.000)		(0.000)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.943	0.943	0.948	0.948	0.940	0.940
R-sq.	0.001	0.010	0.004	0.006	0.001	0.013
Observations	8678	8500	2497	2361	6181	6139

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed a non-zero income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A18: Active - Impact on IHS Tax Amounts Declared

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	CIT	CIT	PIT	PIT
Deterrence	-0.172	0.046	0.065	0.124	-0.260	0.030
	(0.167)	(0.135)	(0.352)	(0.321)	(0.187)	(0.139)
Fiscal Exchange	-0.178	-0.100	-0.266	-0.229	-0.143	-0.062
	(0.165)	(0.133)	(0.346)	(0.313)	(0.186)	(0.138)
Baseline income		0.582***		0.402***		0.673***
		(0.012)		(0.024)		(0.012)
Controls	No	Yes	No	Yes	No	Yes
Control Mean Baseline	1.87	1.87	1.88	1.88	1.86	1.86
Control Mean Endline	2.01	2.01	1.51	1.51	2.19	2.19
R-sq.	0.000	0.346	0.000	0.171	0.000	0.458
Observations	8504	8422	2370	2308	6134	6114

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is the tax amount remitted. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

A5 Discussion of results

Table A19: Non-filers - Impact on Filing Probability

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	CIT	CIT	PIT	PIT	PIT	PIT
Deterrence	0.036***	0.026***	0.038***	0.039***	0.017**	0.017**	0.017**	0.017**
	(0.007)	(0.007)	(0.014)	(0.014)	(0.008)	(0.008)	(0.008)	(0.008)
Cluster Town	0.036***	0.026**	0.038***	0.039***	0.017*	0.017*	0.017*	0.017*
	(0.004)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Drop weeks 5-6	0.032***	0.024***	0.034***	0.034***	0.016**	0.016**	0.016**	0.016**
	(0.006)	(0.006)	(0.013)	(0.013)	(0.007)	(0.007)	(0.007)	(0.007)
Compliance costs	0.024***	0.013**	0.007	0.006	0.018**	0.018**	0.018**	0.018**
	(0.006)	(0.006)	(0.013)	(0.013)	(0.008)	(0.008)	(0.008)	(0.008)
Cluster Town	0.024***	0.013***	0.007	0.006	0.018*	0.018**	0.018*	0.018**
	(0.002)	(0.002)	(0.008)	(0.008)	(0.007)	(0.006)	(0.007)	(0.006)
Drop weeks 5-6	0.022***	0.014**	0.005	0.004	0.020***	0.020***	0.020***	0.020***
	(0.006)	(0.006)	(0.012)	(0.012)	(0.007)	(0.007)	(0.007)	(0.007)
De-registration	0.024***	0.014**	0.021	0.021	0.009	0.009	0.009	0.009
	(0.006)	(0.006)	(0.013)	(0.014)	(0.007)	(0.007)	(0.007)	(0.007)
Cluster Town	0.024**	0.014*	0.021**	0.021**	0.009	0.009	0.009	0.009
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Drop weeks 5-6	0.025***	0.016***	0.028**	0.028**	0.007	0.007	0.007	0.007
	(0.006)	(0.006)	(0.012)	(0.013)	(0.007)	(0.007)	(0.007)	(0.007)
Fiscal Exchange							0.012	0.013*
							(0.008)	(0.008)
Cluster Town							0.012	0.013
							(0.007)	(0.007)
Drop weeks 5-6							0.012*	0.012*
							(0.007)	(0.007)
Social Norm							0.013*	0.014*
							(0.008)	(0.008)
Cluster Town							0.013	0.014*
							(0.007)	(0.006)
Drop weeks 5-6							0.018**	0.019***
							(0.007)	(0.007)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
C. Mean	0.069	0.069	0.095	0.095	0.067	0.067	0.067	0.067
C. Mean drop week 5-6	0.054	0.054	0.075	0.075	0.052	0.052	0.052	0.052
R-sq.	0.003	0.016	0.002	0.009	0.001	0.015	0.001	0.015
F Joint Test	0.000	0.000	0.036	0.027	0.015	0.015	0.017	0.014
Observations	22550	22425	4231	4106	18319	18319	21058	21058

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed an income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A20: Nil-filers - Impact on Positive Filing Probability

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	CIT	CIT	PIT	PIT
Deterrence	-0.002	-0.002	-0.010	-0.012	0.003	-0.002
	(0.015)	(0.015)	(0.017)	(0.017)	(0.022)	(0.022)
Cluster Town	-0.002	-0.001	-0.010	-0.012	0.003	-0.003
	(0.014)	(0.013)	(0.009)	(0.010)	(0.024)	(0.023)
Drop weeks 5-6	0.007	0.005	-0.020	-0.014	0.028	0.022
	(0.030)	(0.029)	(0.035)	(0.033)	(0.042)	(0.042)
De-registration	0.004	0.003	-0.006	-0.005	0.012	0.007
	(0.016)	(0.015)	(0.017)	(0.017)	(0.022)	(0.022)
Cluster Town	0.004	0.006	-0.006	-0.005	0.012	0.006
	(0.012)	(0.010)	(0.015)	(0.015)	(0.023)	(0.025)
Drop weeks 5-6	0.022	0.017	-0.019	-0.016	0.050	0.043
	(0.030)	(0.030)	(0.035)	(0.034)	(0.042)	(0.042)
Controls	No	Yes	No	Yes	No	Yes
C. Mean	0.168	0.168	0.075	0.075	0.225	0.225
C. Mean drop week 5-6	0.298	0.298	0.153	0.153	0.404	0.404
R-sq.	0.000	0.011	0.000	0.030	0.000	0.017
F Joint Test	0.912	0.958	0.830	0.773	0.856	0.924
Observations	3508	3508	1352	1352	2156	2156

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed a non-zero income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A21: Active - Impact on the Probability to Increase Tax

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	CIT	CIT	PIT	PIT
Deterrence	0.008	0.008	0.015	0.015	0.005	0.006
	(0.012)	(0.012)	(0.023)	(0.023)	(0.014)	(0.014)
Cluster Town	0.008	0.008	0.015	0.015	0.005	0.006
	(0.013)	(0.013)	(0.019)	(0.019)	(0.014)	(0.014)
Drop weeks 5-6	0.018	0.018	0.023	0.025	0.016	0.016
	(0.015)	(0.015)	(0.028)	(0.027)	(0.018)	(0.017)
Fiscal Exchange	-0.011	-0.012	-0.056**	-0.055**	0.006	0.007
	(0.012)	(0.012)	(0.023)	(0.023)	(0.014)	(0.014)
Cluster Town	-0.011	-0.012	-0.056***	-0.055***	0.006	0.007
	(0.011)	(0.011)	(0.018)	(0.018)	(0.011)	(0.010)
Drop weeks 5-6	-0.008	-0.009	-0.057**	-0.056**	0.014	0.013
	(0.015)	(0.015)	(0.027)	(0.027)	(0.018)	(0.017)
Controls	No	Yes	No	Yes	No	Yes
C. Mean	0.300	0.300	0.348	0.348	0.280	0.280
C. Mean drop weeks 5-6	0.323	0.323	0.358	0.358	0.308	0.308
R-sq.	0.000	0.014	0.004	0.013	0.000	0.017
F Joint Test	0.000	0.000	0.000	0.000	0.000	0.000
Observations	8678	8678	2497	2497	6181	6181

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A22: Nudge Impact by Receiving any T

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	CIT	CIT	PIT	PIT
	Panel A: Nil-filers - Prob. to Positively File					
Any nudge	0.001	0.001	-0.008	-0.008	0.008	0.003
	(0.013)	(0.013)	(0.015)	(0.015)	(0.019)	(0.019)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.168	0.168	0.075	0.075	0.225	0.225
R-sq.	0.000	0.011	0.000	0.030	0.000	0.017
Observations	3508	3508	1352	1352	2156	2156
	Panel B: Actives - Prob. to Increase Tax Liability					
Any nudge	-0.002	-0.002	-0.020	-0.020	0.006	0.006
	(0.010)	(0.010)	(0.020)	(0.020)	(0.012)	(0.012)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.300	0.300	0.348	0.348	0.280	0.280
R-sq.	0.000	0.013	0.000	0.010	0.000	0.017
Observations	8678	8678	2497	2497	6181	6181

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. In panel A, the dependent variable is an indicator variable for having filed a non-zero income tax return. In panel B, the dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A23: Non-filers - Impact on On-Time Filing Probability

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	CIT	CIT	PIT	PIT	PIT	PIT
Deterrence	0.013^{***}	0.008	0.026^{**}	0.028^{***}	0.002	0.002	0.002	0.002
	(0.005)	(0.005)	(0.010)	(0.010)	(0.006)	(0.006)	(0.006)	(0.006)
Compliance costs	0.013^{**}	0.007	0.012	0.012	0.012^{*}	0.012^{*}	0.012^{*}	0.012^{*}
	(0.005)	(0.005)	(0.010)	(0.010)	(0.006)	(0.006)	(0.006)	(0.006)
De-registration	0.014^{***}	0.008	0.017^{*}	0.017^{*}	0.011^{*}	0.011^{*}	0.011^{*}	0.011^{*}
	(0.005)	(0.005)	(0.010)	(0.010)	(0.006)	(0.006)	(0.006)	(0.006)
Fiscal Exchange							0.004	0.004
							(0.006)	(0.006)
Social Norm							0.002	0.002
							(0.006)	(0.006)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.043	0.043	0.045	0.045	0.043	0.043	0.043	0.043
R-sq.	0.001	0.007	0.002	0.005	0.000	0.009	0.000	0.009
F Joint Test	0.000	0.147	0.066	0.054	0.121	0.111	0.310	0.288
Observations	22544	22419	4226	4101	18318	18318	21057	21057

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed an income tax return on time (by the date of the deadline). All coefficients are OLS estimates from a LPM. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A24: Active - Impact on Log Income Declared

	(1) All	(2) All	(3) CIT	(4) CIT	(5) PIT	(6) PIT
Deterrence	0.320**	0.380**	0.274	0.356*	0.307	0.335*
	(0.161)	(0.153)	(0.258)	(0.199)	(0.190)	(0.188)
Fiscal Exchange	0.120	0.143	0.262	0.093	0.060	0.090
	(0.161)	(0.152)	(0.259)	(0.200)	(0.188)	(0.187)
Baseline income		0.384***		0.908***		0.160***
		(0.017)		(0.024)		(0.021)
Controls	No	Yes	No	Yes	No	Yes
C. Mean Baseline	5.59	5.59	12.24	12.24	3.00	3.00
C. Mean Endline	9.46	9.46	11.80	11.80	8.55	8.55
Observations	8591	8571	2410	2390	6181	6181

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is log income declared. All coefficients are OLS estimates. *All* refers to the total of corporate and individual taxpayers, *CIT* indicates companies only while *PIT* stands for individual income taxpayers.

Table A25: Active - Impact on Log Expenses and Log Deductions Declared (CIT only)

	Expenses	Expenses	Deductions	Deductions
Deterrence	0.148	0.253*	0.142	0.261
	(0.212)	(0.152)	(0.284)	(0.263)
Fiscal Exchange	0.228	0.149	0.166	0.138
	(0.212)	(0.151)	(0.285)	(0.261)
Baseline expenses		1.118***		0.522***
		(0.044)		(0.022)
Controls	No	Yes	No	Yes
Control Mean Baseline	12.85	12.85	3.80	3.80
Control Mean Endline	11.68	11.68	3.69	3.69
Observations	2472	2357	2411	2383

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.*Expenses* indicate the amount of costs declared, while *Expenses* stand for the amount of deductions. All coefficients are OLS estimates from a LPM.

Table A26: Active - Impact on the Probability to Increase Tax Declared

	(1)	(2)	(3)	(4)
	All	All	Non-targeters	Non-targeters
Deterrence	0.008	0.008	0.034*	0.035**
	(0.012)	(0.012)	(0.017)	(0.017)
Fiscal Exchange	-0.011	-0.012	0.008	0.008
	(0.012)	(0.012)	(0.018)	(0.018)
Controls	No	Yes	No	Yes
Control Mean	0.300	0.300	0.525	0.525
R-sq.	0.000	0.014	0.001	0.006
Observations	8678	8678	4813	4813

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are OLS estimates from a LPM. *All* refers to the total of targeting and non-targeting taxpayers, *Non-targeters* refers to the subsample of actives who are not targeting, after I removed the 44% of the actives who are remitting the same tax amount as at baseline.

Table A27: Active - Impact on the Probability to Increase Tax Declared

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Perpetual	Perpetual	Non-perp.	Non-perp.
	Panel A: Actives - Prob. to Increase Tax Liability					
Deterrence	0.008	0.008	-0.001	-0.001	0.056**	0.054**
	(0.012)	(0.012)	(0.014)	(0.014)	(0.025)	(0.025)
Fiscal Exchange	-0.011	-0.012	-0.019	-0.019	0.025	0.026
	(0.012)	(0.012)	(0.013)	(0.013)	(0.025)	(0.025)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.300	0.300	0.323	0.323	0.194	0.194
R-sq.	0.000	0.014	0.000	0.013	0.003	0.026
Observations	8678	8678	7063	7063	1615	1615
	Panel B: Nil-filers - Prob. to Positively File					
Deterrence	-0.002	-0.002	-0.015	-0.015	-0.002	0.001
	(0.015)	(0.015)	(0.016)	(0.015)	(0.029)	(0.029)
De-registration	0.004	0.003	-0.001	0.001	0.003	0.003
	(0.016)	(0.015)	(0.016)	(0.016)	(0.030)	(0.030)
Controls	No	Yes	No	Yes	No	Yes
Mean of Y	0.168	0.168	0.097	0.097	0.276	0.276
R-sq.	0.000	0.011	0.001	0.032	0.000	0.010
Observations	3508	3508	2115	2115	1393	1393

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. In panel A, the dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. In panel B, the dependent variable is an indicator variable for having filed a non-zero income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of perpetual and non-perpetual taxpayers, *Perpetual* indicates taxpayers who have been actively filing (panel A) or nil-filing (panel B) since registration with the authority, while *Non-perp.* stands for taxpayers who have alternatively opted for the 3 filing behaviours.

Table A28: Nudge Impact by Registration Year

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	First	First	2nd+	2nd+
Panel A: Non-filers - Prob. to File						
Deterrence	0.036***	0.024***	0.029	-0.003	0.037***	0.026***
	(0.007)	(0.007)	(0.034)	(0.035)	(0.007)	(0.007)
Compliance costs	0.024***	0.011*	0.013	-0.021	0.024***	0.014**
	(0.006)	(0.007)	(0.032)	(0.034)	(0.006)	(0.007)
De-registration	0.024***	0.012*	0.024	-0.009	0.024***	0.014**
	(0.006)	(0.006)	(0.033)	(0.034)	(0.006)	(0.007)
Fiscal Exchange	0.010	0.013*	-0.072**	-0.064**	0.016**	0.018**
	(0.008)	(0.008)	(0.032)	(0.032)	(0.008)	(0.008)
Social Norm	0.012	0.014*	0.109**	0.115**	0.007	0.009
	(0.008)	(0.008)	(0.054)	(0.053)	(0.007)	(0.007)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.069	0.069	0.148	0.148	0.064	0.064
R-sq.	0.002	0.016	0.007	0.020	0.002	0.013
Observations	25289	25164	1494	1493	23795	23671
Panel B: Nil-filers - Prob. to Positively File						
Deterrence	-0.002	-0.002	-0.090	-0.092*	0.006	0.007
	(0.015)	(0.015)	(0.056)	(0.055)	(0.016)	(0.016)
De-registration	0.004	0.004	-0.029	-0.012	0.008	0.006
	(0.016)	(0.016)	(0.059)	(0.059)	(0.016)	(0.016)
Controls	No	No	No	Yes	No	Yes
Control Mean	0.168	0.168	0.233	0.233	0.161	0.161
R-sq.	0.000	0.000	0.009	0.056	0.000	0.012
Observations	3508	3508	287	287	3221	3221
Panel C: Actives - Prob. to Increase Tax Liability						
Deterrence	0.008	0.008	-0.042	-0.053	0.010	0.011
	(0.012)	(0.012)	(0.067)	(0.067)	(0.012)	(0.012)
Fiscal Exchange	-0.011	-0.012	-0.038	-0.053	-0.010	-0.010
	(0.012)	(0.012)	(0.064)	(0.065)	(0.012)	(0.012)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.300	0.300	0.324	0.324	0.299	0.299
R-sq.	0.000	0.014	0.002	0.020	0.000	0.014
Observations	8678	8678	291	291	8387	8387

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. In panel A, the dependent variable is an indicator variable for having filed an income tax return. In panel B, the dependent variable is an indicator variable for having filed a non-zero income tax return. In panel C, the dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are OLS estimates from a LPM. *All* refers to the total of young and old taxpayers, *First year* indicates taxpayers who filed for the first time in the baseline year 2018, while *2nd Year and more* stands for taxpayers who were at their second year of more when they filed in in the baseline year 2018.

Table A29: Nudge Impact by Urban/Rural Location

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Urban	Urban	Rural	Rural
Panel A: Nil-filers - Prob. to Positively File						
Deterrence	-0.002	-0.002	0.009	0.009	-0.022	-0.019
	(0.015)	(0.015)	(0.019)	(0.019)	(0.026)	(0.026)
De-registration	0.004	0.004	0.008	0.004	-0.003	0.001
	(0.016)	(0.016)	(0.019)	(0.019)	(0.027)	(0.027)
Controls	No	No	No	Yes	No	Yes
Control Mean	0.168	0.168	0.233	0.233	0.161	0.161
R-sq.	0.000	0.000	0.000	0.019	0.001	0.009
Observations	3508	3508	2167	2167	1341	1341
Panel B: Actives - Prob. to Increase Tax Liability						
Deterrence	0.008	0.008	0.010	0.009	0.005	0.007
	(0.012)	(0.012)	(0.016)	(0.016)	(0.019)	(0.019)
Fiscal Exchange	-0.011	-0.012	-0.002	-0.004	-0.027	-0.023
	(0.012)	(0.012)	(0.016)	(0.016)	(0.018)	(0.018)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.300	0.300	0.318	0.318	0.272	0.272
R-sq.	0.000	0.014	0.000	0.011	0.001	0.020
Observations	8678	8678	5251	5251	3427	3427

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. In panel A, the dependent variable is an indicator variable for having filed a non-zero income tax return. In panel B, the dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are OLS estimates from a LPM. *All* refers to the total of urban and rural taxpayers, *Urban* indicates non-filers located in urban areas, while *Rural* stands for taxpayers are located in rural areas.

Table A30: Active - Impact on the Probability to Increase Tax

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Dec. 1	Dec. 1	Dec. 2-10	Dec. 2-10
Deterrence	0.008	0.008	-0.018	-0.018	0.012	0.012
	(0.012)	(0.012)	(0.038)	(0.038)	(0.013)	(0.013)
Fiscal Exchange	-0.011	-0.012	-0.062*	-0.070*	-0.005	-0.005
	(0.012)	(0.012)	(0.038)	(0.038)	(0.013)	(0.013)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.300	0.300	0.384	0.384	0.289	0.289
R-sq.	0.000	0.014	0.003	0.015	0.000	0.014
Observations	8678	8678	971	971	7707	7707

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are OLS estimates from a LPM. *All* refers to the total of small and large taxpayers, *Dec. 1* indicates taxpayers for the top income decile, while *Dec. 2-10* stands taxpayers in all remaining deciles.

Table A31: Nudge Impact by Age (PIT only)

	(1) All PIT	(2) All PIT	(3) Young	(4) Young	(5) Old	(6) Old
Panel A: Non-filers - Prob. to File						
Deterrence	0.017**	0.017**	0.013	0.013	0.030	0.028
	(0.006)	(0.006)	(0.014)	(0.013)	(0.023)	(0.022)
Compliance costs	0.018*	0.018**	-0.004	-0.003	0.039	0.041
	(0.007)	(0.006)	(0.007)	(0.008)	(0.019)	(0.018)
De-registration	0.009	0.009	0.001	-0.000	0.014	0.016
	(0.006)	(0.006)	(0.004)	(0.003)	(0.010)	(0.012)
Fiscal Exchange	0.012	0.013*	0.002	0.002	0.020*	0.020*
	(0.007)	(0.007)	(0.010)	(0.010)	(0.008)	(0.008)
Social Norm	0.013*	0.014*	0.013*	0.013*	0.018	0.018
	(0.007)	(0.006)	(0.005)	(0.004)	(0.013)	(0.011)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.067	0.067	0.069	0.069	0.071	0.071
R-sq.	0.001	0.015	0.000	0.017	0.002	0.030
Observations	21058	21058	9909	9909	9362	9362
Panel B: Nil-filers - Prob. to Positively File						
Deterrence	0.003	0.003	0.007	0.006	0.009	-0.000
	(0.022)	(0.022)	(0.031)	(0.031)	(0.033)	(0.033)
De-registration	0.012	0.012	-0.004	-0.008	0.029	0.024
	(0.022)	(0.022)	(0.031)	(0.031)	(0.033)	(0.033)
Controls	No	No	No	Yes	No	Yes
Control Mean	0.225	0.225	0.217	0.217	0.241	0.241
R-sq.	0.000	0.000	0.000	0.016	0.001	0.035
Observations	2156	2156	1065	1065	1011	1011
Panel C: Actives - Prob. to Increase Tax Liability						
Deterrence	0.005	0.006	0.023	0.022	-0.013	-0.010
	(0.014)	(0.014)	(0.020)	(0.020)	(0.020)	(0.020)
Fiscal Exchange	0.006	0.007	0.006	0.008	0.008	0.007
	(0.014)	(0.014)	(0.020)	(0.019)	(0.020)	(0.020)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.280	0.280	0.277	0.277	0.283	0.283
R-sq.	0.000	0.017	0.000	0.028	0.000	0.013
Observations	6181	6181	3124	3124	3021	3021

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. In panel A, the dependent variable is an indicator variable for having filed an income tax return. In panel B, the dependent variable is an indicator variable for having filed a non-zero income tax return. In panel C, the dependent variable is an indicator variable for having increased the tax declared compared to the baseline year. All coefficients are OLS estimates from a LPM. *All* refers to the total of young and old taxpayers, *Young* indicates taxpayers who are younger than the median age in the sample (50), while *2nd Year and more* stands for taxpayers who are older than the median age.

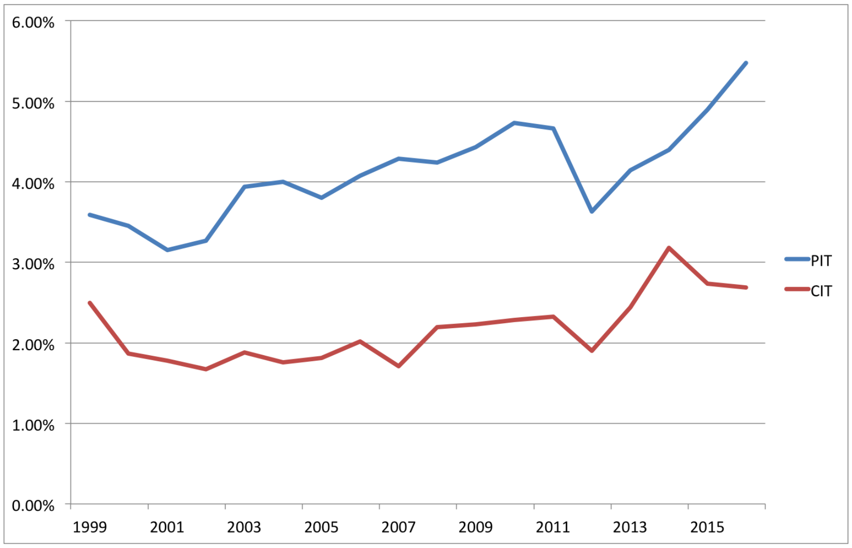
Table A32: Nudge Impact by Civil Status (PIT only)

	(1) All PIT	(2) All PIT	(3) Married	(4) Married	(5) Non-married	(6) Non-married
Panel A: Non-filers - Prob. to File						
Deterrence	0.017**	0.017**	0.020	0.017	-0.001	0.002
	(0.008)	(0.008)	(0.016)	(0.016)	(0.013)	(0.013)
Compliance costs	0.018**	0.018**	-0.010	-0.009	0.006	0.009
	(0.008)	(0.008)	(0.013)	(0.013)	(0.014)	(0.014)
De-registration	0.009	0.009	0.022	0.021	0.012	0.011
	(0.007)	(0.007)	(0.016)	(0.016)	(0.014)	(0.014)
Fiscal Exchange	0.012	0.013*	-0.004	-0.007	0.026*	0.026*
	(0.008)	(0.008)	(0.014)	(0.014)	(0.015)	(0.015)
Social Norm	0.013*	0.014*	0.023	0.024	0.018	0.019
	(0.008)	(0.008)	(0.016)	(0.016)	(0.014)	(0.014)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.067	0.067	0.084	0.084	0.058	0.058
R-sq.	0.001	0.015	0.001	0.025	0.001	0.025
Observations	21058	21058	6166	6166	5555	5555
Panel B: Actives - Prob. to Increase Tax Liability						
Deterrence	0.005	0.006	-0.029	-0.027	0.012	0.009
	(0.014)	(0.014)	(0.023)	(0.023)	(0.034)	(0.033)
Fiscal Exchange	0.006	0.007	0.027	0.027	-0.045	-0.046
	(0.014)	(0.014)	(0.024)	(0.024)	(0.033)	(0.033)
Controls	No	Yes	No	Yes	No	Yes
Control Mean	0.280	0.280	0.296	0.296	0.279	0.279
R-sq.	0.000	0.017	0.003	0.018	0.003	0.045
Observations	6181	6181	2236	2236	1051	1051

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. The dependent variable is an indicator variable for having filed an income tax return. All coefficients are OLS estimates from a LPM. *All* refers to the total of married and non-married taxpayers, *Married* indicates non-filers who are married, while *Non-married* stands for taxpayers who are not married.

A6 Figures

Figure A1: PIT vs CIT shares over GDP



Source: ICTD Government Revenue Dataset.

Figure A2: Deregistrations from the Tax System

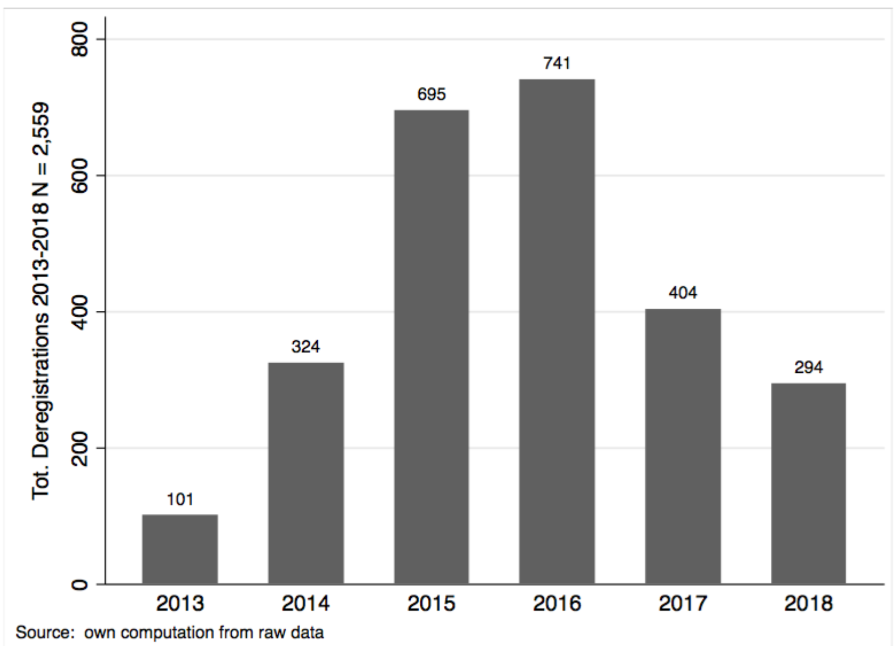


Figure A3: Experiment Letter

Eswatini Revenue Authority

HEADQUARTERS
Portion 419 of Farm 50
Ezulwini
Along MR 103 (Mvutshini-Gables Road)
GPS Coordinates: S 26 25.120 E 31 10.623

PO Box 5628, Mbabane, Eswatini
Tel: +268-2406 4000 Fax: +268-2406 4001
E-mail: info@sra.org.sz
Website: www.sra.org.sz

Our Ref: SRA11/6/«TIN»/ZN | Date: **23/10/2019**

«Taxpayername»
«POBox_2»
«Town_2»

Dear Sir / Madam,

RE: COMPLY TO AVOID PENALTIES AND FINES FOR FAILURE TO FILE

According to the records of the Eswatini Revenue Authority (SRA), you did not submit your income tax return for the tax year 2018. For this reason, SRA wants to inform you that:

- SECTION 66 of the INCOME TAX ORDER (1975): a taxpayer who fails to submit a return within the stipulated period commits an offence and may be liable on conviction to a fine of **E 10,000**, or imprisonment for a period of up to **one year**, or both.
- SECTION 40 of the INCOME TAX ORDER (1975): a taxpayer who defaults in submitting a return for any year of assessment is liable to pay additional tax of an amount equal to **twice the tax chargeable** in respect to his taxable income for such year of assessment

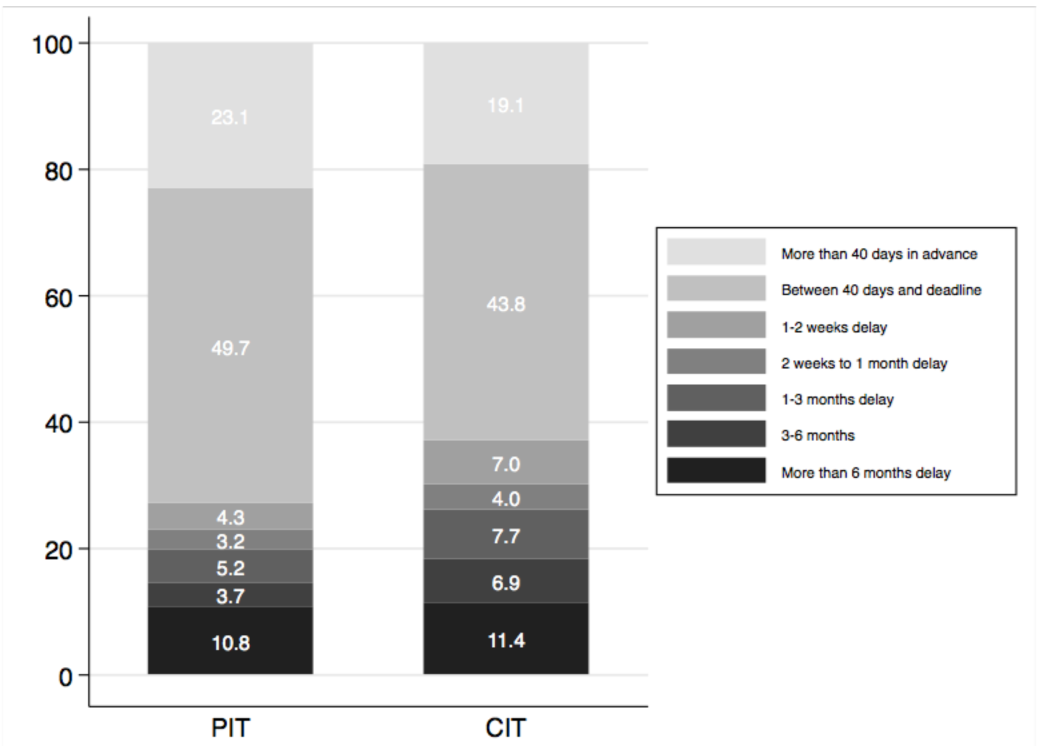
Please, comply with your tax obligations and to ensure your declarations are correct to avoid fines and penalties.

You can get further information or assistance by contacting the SRA Service Center at: Tel +268 2406 4000; Fax +268 2406 4001. You can also email us at: info@sra.org.sz. Please, always mention your taxpayer identification number (TIN) in all correspondence with the SRA.

Yours faithfully,

Director of Compliance
Domestic Taxes

Figure A4: Timing of Filing of Income Tax Return 2013-2018



Chapter 5

Conclusions

5.1 Main findings

The main contribution of this thesis consists in a more concrete understanding of which factors explain tax compliance and how they could be realistically leveraged by resource-constrained revenue authorities in sub-Saharan Africa. The quest for more knowledge on taxpayers' behaviour and for more successful, evidence-based compliance strategies for the tax administrators to implement is the common thread linking the three studies presented in this dissertation.

The first study (Chapter 2) explores the impact of the provision of tax education in building income tax compliance. The context, Rwanda, is one of a low-income country, where income tax is a key source of revenue but compliance is far from being optimal. Especially newly registered taxpayers find it difficult to comply with a presumably obscure tax law – even in terms of filing a tax return. In collaboration with the Rwanda Revenue Authority, who implemented the Taxpayer Training Programme countrywide, I produce causal evidence on the impact of such trainings on attendees' behaviour. By means of an encouragement design and the combination of survey and administrative data, I show that attending the training increases the probability to file a return by 43% over the control group average. While impacts at the extensive margin are positive, non-significant results are found in terms of tax remitted. As key mechanisms of the impact on filing, improvements in tax knowledge and complexity perceptions seem to play a major role.

Also, an alternative educational initiative, a one-to-one coaching service, produces null results – mostly due to poor implementation from the authority.

The second study (Chapter 3), run in Eswatini, explores the drivers of the decision to file an income tax return by surveying one thousand entrepreneurs, evenly split between non-filers and active filers. Both (neoclassical) pecuniary and (morale-based) non-pecuniary theoretical formulations are tested. In this chapter too, survey data is matched with administrative data. Dataset matching also allows for the comparison of survey-based, self-reported measures of compliance with actual filing behaviour. This study shows that different sets of factors are at work when taxpayers decide whether to file or not: (i) the threat of audits, (ii) compliance costs, proxied by perceptions on the ease to file and tax knowledge, (iii) non-pecuniary factors such as peer pressure and intrinsic motivation. Also, the role of these factors is strengthened when persistent non-filers and persistent active filers are compared. This study also shows that self-reported compliance is not necessarily correlated with factors (i)–(iii) above: when regressing the self-reported measure of compliance, deterrence is no longer important, while the quality of communication with the authority, as well as not expecting anything in return (somehow in contrast with the concept of fiscal reciprocity) are significantly affecting compliance.

Lastly, the third study (Chapter 4), always set in Eswatini, represents a methodologically more robust way of understanding taxpayers' behaviour. With a randomised-controlled trial, I test different theoretical motivations behind evasion by nudging 20,000 income taxpayers. Three different filing categories are drawn from tax records – non-, nil- and active filers – and nudged with behaviourally-informed mailings from the tax authority. Results indicate that non-filers are significantly responsive to nudges, irrespective of whether they stress pecuniary or non-pecuniary motives. Nil and positive filers do not respond to mailings and, in some cases, as with large active companies, they even reduce their tax liability. If anything, nudges targeting positive filers are effective in impeding them to fall into nil-filing. Importantly, lots of variation is observed between firms and individuals and across additional dimensions such as past filing behaviour, location, size and age of the business.

5.2 Research and policy implications

This thesis attempted to contribute to the growing literature on public finance and development and tried to do so both theoretically and methodologically.

In terms of the theoretical contributions, this work is one of the first in considering the multidimensional nature of tax compliance, by considering three filing categories – those of non-, nil- and active filers. These are agents who only in few instances report positive tax to the fiscus. In the majority of cases, they either fail to file or, when filing, report zero income and zero tax due. The phenomenon of non- and nil-filers is widespread in SSA but still insufficient evidence on them has been produced in the literature, which has historically focused on positive filers. Second, this thesis explores some factors of the taxpayer evasion decision problem which are still under-researched. Above all, the role of tax knowledge seems to be pivotal both for newly registered taxpayers (Chapter 2) and the population of taxpayers – old and new – as a whole (Chapter 3-4), as well as in encouraging both self-reported and actual compliance (Chapter 3). Relevantly, results are consistent across countries. Third and relatedly, additional aspects of the taxpayer's profile are taken into account. For example, the role of past filing behaviour, largely neglected in the literature, is explicitly recognised in Eswatini and has important implications for the Rwandan study on newly registered taxpayers as well, since the training program's goal was to build a solid taxpaying habit from the moment in which a new taxpayer enters the tax system.

For what concerns the methodological contributions, this thesis aimed at proposing a new solution to the intricate problem of measuring tax compliance. All three studies relied on the combination of survey and a wealth of detailed administrative data - merged with unique identifiers - a novel approach rarely found in the literature. Also, this approach shed some light on the inconsistencies inherently involved with self-reports (Chapter 3), which often are not consistent with actual records on behaviour. As a second important contribution, this effort aimed at showing a successful way of doing tax research in the South, by setting up a long-term research collaboration with local partners (tax authorities), based on mutual learning and the in-depth understanding of the local realities.

Last but not least, this thesis has important policy implications. First, it highlighted the role of scientific research in terms of designing better tax compliance strategies and generated a genuine and enthusiastic interest towards research within the partnering tax administrations. Most importantly, my work with revenue authorities increased my partners' awareness of the huge potential that can be untapped by an adequately sophisticated quantitative analysis of the wealth of tax data routinely produced by their IT systems. Second, the results of this thesis are able to show that revenue authorities should complement their traditional enforcement strategies, usually based on audits and fines, with a variety of alternative, non-pecuniary, solutions that seem to be equally effective. Among these solutions, improving taxpayers' knowledge through dedicated educational initiatives seems to be the way forward (Chapter 2, 3, 4). Likewise, modernising the communication strategies is also recommended, especially so in Eswatini, as it may improve both the intrinsic willingness to comply (Chapter 3) as well as compliance at the extensive margin, pushing non-filers to submit tax returns (Chapter 4).

5.3 Limitations and next steps

Despite consisting of completed projects, this thesis has also prompted many plans for additional analysis and improvements. This is mostly due to the fact that, due to time and budget constraints, not everything I dreamed to do actually happened in the field. The fact that the following limitations exist represents an exciting opportunity for developing more research work in the future:

- **Long-term impacts.** It is crucial to understand if an intervention fostering compliance has long-lived effects. Needless to say, it takes time to produce evidence on this. While I found the time to show only some initial and inconclusive evidence on mid-term effects of the tax trainings in Rwanda (Chapter 2), it is unknown whether the tax nudge impacts I found in Eswatini (Chapter 4) will persist over time. For this reason, I am committed to track the experimental sample in the next tax years and realise whether a one-time input from the authority is enough to generate sustained effects. The same will be done with the Rwandan sample in Chapter 2. The long-term research collaboration with the two revenue authorities ensures that the relevant datasets will be shared in the future.

- **Alternative communication channels.** Chapter 4 speaks of the effectiveness of mailings, while it cannot say much about other delivery methods. Ongoing discussions with the SRA are exploring the possibility of using alternative channels, such as emails and SMSs, to nudge taxpayers. While emails are available for a restricted, non-representative subsample of the taxpayer population, phone numbers can be derived from all taxpayers. On the one hand, emails can be used to better target nil- and active filers (they are more likely to have e-mail addresses) to shed more light on the null/negative impacts of mailings from Chapter 4. On the other hand, SMSs could be cheaply implemented when targeting non-filers, who are the vast majority and seem to respond to any nudge from the authority.

- **More on nil-filers.** Nil-filers still remain a puzzle in the African tax systems. The limited evidence on them from few countries - mostly in Rwanda and Eswatini, thanks to the work I contributed to at the International Centre for Tax and Development - is inconclusive as well. It may seem that they are just legitimately reporting zero turnover, mostly because they have not been operating since registration. Hence, nil-filers are not affected by deterrence and do not need to de-register from the system (Chapter 4). If this is true, the tax administrations should revise their registration policies, which are often aggressive and aimed at registering as many new taxpayers as possible even if they are not generating any income. To understand better what lies behind this behaviour, I plan to run additional qualitative work on this group, especially through focus group discussions and open interviews. Extra survey data, to be compared to what gathered in Chapter 3, will be beneficial as well.

- **Explore the linkages between registered non-filers and informals.** While I have been able to discriminate between non-filers and active taxpayers in Chapter 3, the open question remains of whether non-filers are more similar to small, subsistence-level, informal traders (such as street vendors) than to active agents. Survey evidence from Chapter 3 seems to confirm this conjecture - non-filers perform poorly in terms of tax knowledge and awareness, direct involvement in the tax system and relationship with the revenue authority. Also, they have more traditional and rudimentary business practices than active payers. Additional survey data can be gathered from totally informal traders, not registered with the authority, to see how comparable they are with their registered, formal counterparts. Robust evidence on these linkages could inform the revenue authorities on how they should register new taxpayers. If including informal-like traders into the system does not bring any benefit, not only in terms of revenue raised but also in terms of filing habits, the authority should reconsider its registration strategy. Ideally, this could potentially lead to the institution of a new, simplified tax regime for all informal taxpayers, with only a fixed, presumptive tax payment required. Much more work should be carried on in this direction.

- **Better understand the differences in compliance between companies and individuals.** As documented in Chapter 4, companies and sole traders differently react to behavioural nudges. More specifically, companies are influenced by the enforcement approach only, while individual businesses react to both hard- and soft-toned messages. Furthermore, even within incorporated entities, differential impacts are found depending on the business' size - top-income decile companies are showing perverse responses, reducing tax declared. This indicates that the determinants of the compliance decision of companies and individuals are partially different. Most tax morale studies are still focused on individual taxpayers, with findings on tax morale often extrapolated to businesses as well (Prichard et al., 2019). However, businesses, which are more likely to be larger and more operative than sole traders, may face the compliance decision more self-interestedly rather than driven by absolute ethical concerns. Companies are likely to be more focused on factors such as the predictability of business environment, profit maximisation, government investment and competition. On top of that, companies consist of groups of individuals, often organised within a well structured and bureaucratic setting, where the intrinsic motivation to comply of an individual component may significantly diverge from the company's values and behaviours. Chapter 4 only produced some preliminary evidence on this understudied topic and much more work is needed in order to understand what drives compliance in different kinds of taxpayers.

I trust that I will devote my future work to these key aspects of the tax compliance conundrum and ultimately improve my understanding of the complexities, challenges and potentials of the tax systems in sub-Saharan Africa.

Bibliography

Acharya, A., M. Blackwell, and M. Sen (2016). Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects. *American Political Science Review 110*(3), 512–529.

Advani, A., W. Elming, and J. Shaw (2019, May). The Dynamic Effects of Tax Audits. Working Paper Series 414, Centre for Competitive Advantage in the Global Economy.

Aiko, R. and C. Logan (2014). Africa's willing taxpayers thwarted by opaque tax systems, corruption. Policy Paper 7, Afrobarometer.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes 50*(2), 179–211.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control 19*(6), 716–723.

Akitoby, B., J. Honda, H. Miyamoto, K. Primus, and M. Sy (2019). Case Studies in Tax Revenue Mobilization in Low-Income Countries. *IMF Working Paper 19*(104).

Akpan, I. and K. Sempere (2019, May). Hidden Inequalities: Tax Challenges of Market Women in Enugu and Kaduna States, Nigeria. Working Paper Series 97, International Centre for Taxation and Development.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003). Fractionalization. Working Paper 9411, NBER.

Ali, M. (2018). Regulatory Burdens in Tax Administration and Firms' Compliance Costs in Africa. Working Paper 78, International Centre for Tax and Development.

Ali, M., O.-H. Fjeldstad, and I. H. Sjursen (2015). To pay or not to pay? citizens' attitudes towards taxation in Kenya, Tanzania, Uganda and South Africa. *World Development 64*, 828–842.

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

Allcott, H. and T. Rogers (2014). The Short-Run and Long-Run Effects of Behav- ioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review 104*(10), 3003–37.

Allingham, M. G. and A. Sandmo (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics 1*(3-4), 323–338.

Alm, J. (1988). Uncertain tax policies, individual behavior, and welfare. *The American Economic Review 78*(1), 237–245.

Alm, J. (2012, 02). Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies. *International Tax and Public Finance 19*(1), 54–77.

Alm, J., K. Bloomquist, and M. McKee (2015). On the external validity of laboratory tax compliance experiments. *Economic Inquiry 53*(2), 1170–1186.

Alm, J., T. Cherry, M. Jones, and M. McKee (2010). Tax Information Assistance Services and Tax Compliance Behavior. *Journal of Economic Psychology 31*(4), 577–586.

Alm, J., B. Jackson, and M. McKee (1992). Institutional uncertainty and taxpayer compliance. *American Economic Review 82*(4).

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

Alm, J., G. H. McClelland, and W. D. Schulze (1992). Why do people pay taxes? *Journal of Public Economics 21*(38).

Alm, J. and M. McKee (2003). Tax compliance as a coordination game. Working Paper 06-32, Andrew Young School of Policy Studies.

Alm, J. and B. Torgler (2011). Do ethics matter? Tax compliance and morality. *Journal of Business Ethics 101*(4), 635–651.

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

Alvaredo, F. and A. Atkinson (2010). Colonial Rule, Apartheid and Natural Resources: Top Incomes in South Africa, 1903–2007. CEPR Discussion Paper 8155, Centre for Economic Policy Research.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2006). Elicitation Using Multiple Price Lists. *Experimental Economics 9*(4), 384–405.

Anderson, S., C. Lazicky, and B. Zia (2019). Measuring the Unmeasured: Combining Technology and Behavioral Insights to Improve Measurement of Business Outcomes. Policy Research Working Paper 8836, World Bank Group.

Andreoni, J., B. Erard, and J. Feinstein (1998, June). Tax Compliance. *Journal of Economic Literature 36*(2), 818–860.

Andrisani, P. J. (1978). *Work attitudes and labor market experience.* New York, NY: Praeger Pub- lishers.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 ed.). Princeton University Press.

Ariel, B. (2012). Deterrence and moral persuasion effects on corporate tax compliance: Findings from a randomized controlled trial. *Criminology 50*(1), 27–69.

ATAF (2017). African Tax Outlook. Ataf publication, African Tax Administration Forum.

Atinyan, A. and Z. Asatryan (2019, December). Nudging for Tax Compliance: A Meta-Analysis. Discussion Paper 19-055, ZEW.

Attanasio, O., A. Barr, J. C. Cardenas, G. Genicot, and C. Meghir (2012). Risk pooling, risk preferences, and social networks. *American Economic Journal: Applied Economics 4*(2), 134–167.

Baldry, J. C. (1986). Tax evasion is not a gamble. *Economics Letters* (22), 333–335.

Banerjee, A. (1992, August). A Simple Model of Herd Behavior. *The Quarterly Journal of Economics 107*(3), 797–817.

Banerjee, A., R. Hanna, J. Kyle, B. A. Olken, and S. Sumarto (2018). Tangible information and citizen empowerment: Identification cards and food subsidy programs in indonesia. *Journal of Political Economy 126*(2), 451–491.

Banerjee, A. V., E. Duflo, R. Glennerster, and C. Kinnan (2015). The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics 7*(1), 22–53.

Banerjee, A. V. and A. F. Newman (1994). Poverty, incentives and development. *American Economic Review 84*(2), 211–15.

Bardach, E. (1989). Moral Suasion and Taxpayer Compliance. *Law and Policy 11*(49-69).

Bayart, J.-F. (1993). *The State in Africa: The Politics of the Belly.* London and New York, Longman.

Beck, P. J., J. S. Davis, and W.-O. Jung (1991). Experimental evidence on taxpayer reporting under uncertainty. *The Accounting review 66*(3).

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*, 169.

Benzarti, Y. (2015). How taxing is tax filing? Leaving money on the table because of hassle costs. mimeo, University fo California, Berkeley.

Bergolo, M., R. Ceni, G. Cruces, M. Giaccobasso, and R. Perez-Truglia (2019). Tax Audits as Scarecrows. Evidence from a Large-Scale Field Experiment. IZA Discussion Papers 12335, Institute of Labor Economics (IZA).

Bernasconi, M. (1998). Tax evasion and orders of risk aversion. *Journal of Public Economics 67*, 123–134.

Besley, T. and T. Persson (2013). Taxation and Development. *Handbook of Public Economics 5*.

Besley, T. and T. Persson (2014). Why do developing countries tax so little? *Journal of Economic Perspectives 28*(4), 99–120.

Best, M., A. Brockmeyer, H. Kleven, J. Spinnewijn, and M. Waseem (2015). Production vs Revenue Efficiency with Low Tax Compliance: Theory and Evidence from Pakistan. *Journal of Political Economy 123*(6), 1311–1355.

Biddle, N., K. Fels, and M. Sinning (2017). Behavioral Insights and Business Taxation: Evidence from Two Randomized Controlled Trials. DISCUSSION PAPER SERIES 10795, IZA.

Binswanger, H. P. (1980). Attitudes toward risk, experimental measurement in rural india. *American Journal of Agricultural Economics 62*, 395–407.

Bird, R. (2015). Improving Tax Administration in Developing Countries. *Journal of Tax Administration 1*, 23–45.

Bird, R. and P.-P. Gendron (2007). *The VAT in Developing and Transitional Countries*. Cambridge: Cambridge University Press.

Bird, R. M., J. Martinez-Vazquez, and B. Torgler (2008). Tax effort in developing countries and high income countries: The impact of corruption, voice and accountability. *Economic Analysis and Policy 38*(1), 55 – 71.

BIT (2012). Applying behavioural insights to reduce fraud, error and debt. Technical report, Behavioural Insights Team, Cabinet Office, UK.

Blimpo, M., J. T. Mensah, K. Ochieng' Opalo, and R. Shi (2018, April). Electricity provision and tax mobilization in africa. Working Paper 179, Afrobarometer.

Blumenthal, M., C. Christian, and J. Slemrod (2001). Do normative appeals affect tax compliance? evidence from a controlled experiment in Minnesota. *National Tax Journal 54*(1), 125–138.

Bodea, C. and A. Lebas (2016). The Origins of Voluntary Compliance: Attitudes toward Taxation in Urban Nigeria. *British Journal of Political Science 46*(1), 215–238.

Bordignon, M. (1993). A fairness approach to income tax evasion. *Journal of Public Economics 52*, 345–362.

Bosco, L. and L. Mittone (1997). Tax Evasion and Moral Constraints: Some Experimental Evidence. *Kyklos 50*(3), 297–324.

Bott, K., A. W. Cappelen, and E. Ø. Sørensen (2014). You've got mail: A randomised field experiment on tax evasion. Discussion paper 26/2014, Department of Economics, Norwegian School of Economics.

Braithwaite, V. (2009). *Defiance in Taxation and Governance – Resisting and Dismissing Authority in a Democracy*. Edward Elgar Publishing.

Bratton, M. and E. Gyimah-Boadi (2016). Do trustworthy institutions matter for development? Corruption, trust, and government performance in Africa. Dispatch 112, Afrobarometer.

Brautigam, D., O.-H. Fjeldstad, and M. Moore (2008). *Taxation and state building in developing countries: capacity and consent*. Cambridge University Press.

Brockmeyer, A. and M. A. Hernandez (2016). Taxation, information, and withholding : evidence from Costa Rica. Policy Research working paper WPS 7600, World Bank Group, Washington, D.C.

Brockmeyer, A., S. Smith, M. Hernandez, and S. Kettle (2019, August). Casting a wider tax net: Experimental evidence from costa rica. *American Economic Journal: Economic Policy 11*(3), 55–87.

Bruhn, M. and D. McKenzie (2009). In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics 1*(4), 200–232.

Bryan, G., S. Chowdhury, and A. M. Mobarak (2014). Underinvestment in a profitable technology: The case of seasonal migration in bangladesh. *Econometrica 82*(5), 1671–1748.

Burbidge, J. B., L. Magee, and A. L. Robb (1988). Alternative Transformations to Handle Extreme Values of the Dependent Variable. *Journal of the American Statistical Association 83*(401), 123–127.

Cameron, A. C. and D. L. Miller (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resouces 50*(2), 317–72.

Cardenas, J. C. and J. Carpenter (2008). Behavioural development economics: Lessons from field labs in the developing world. *The Journal of Development Studies 44*, 311–338.

Carpena, F., S. A. Cole, J. Shapiro, and B. H. Zia (2015). The ABCs of financial education : experimental evidence on attitudes, behavior, and cognitive biases. Policy Research Working Paper WPS 7413, World Bank Group.

Carrillo, P., D. Pomeranz, and M. Singhal (2017). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics 9*(2), 144–64.

Casanegra de Jantscher, M. (1990). *Value Added Taxation in Developing Countries*, Chapter Administering the VAT. Washington: World Bank.

Castro, L. and C. Scartascini (2013). Tax compliance and enforcement in the Pampas: Evidence from a field experiment. *Journal of Economic Behavior and Organization 116*, 65–82.

Chabal, P. and J.-P. Daloz (1999). *Africa Works: Disorder as Political Instrument*. James Currey Publishers.

Chen, J. and Z. Chen (2008). Extended bayesian information criteria for model selection with large model spaces. *Biometrika 95*(3), 759–771.

Chetty, R., M. Mobarak, and M. Singhal (2014). Increasing Tax Compliance through Social Recognition. Policy Brief 31101, IGC.

Chetty, R. and E. Saez (2013). Teaching the tax code: Earnings responses to an experiment with eitc recipients. *American Economic Journal: Applied Economics 5*(1), 1–31.

Chirico, M., R. P. Inman, C. Loeffler, J. MacDonald, and H. Sieg (2016). An Experimental Evaluation of Notification Strategies to Increase Property Tax Compliance: Free-Riding in the City of Brotherly Love. *Tax Policy and the Economy 30*, 129–161.

Chorvat, T. (2007). Tax ompliance and the neuroeconomics of intertemporal substitution. *National Tax Journal 60*, 577–588.

Christian, C., S. Gupta, and S. Lin (1993). Determinants of tax preparer usage: Evidence from panel data. *National Tax Journal 46*(4), 487–503.

Cialdini, R. B. and N. J. Goldstein (1991). Social influence: compliance and conformity. *Annual Revenue of Psychology 55*, 591–622.

Cicchetti, C. J. and J. A. Dubin (1994). A microeconometric analysis of risk aversion and the decision to self-insure. *Journal of Political Economy 102*(1), 169–86.

Clotfelter, C. (1983). Tax evasion and tax rates: an analysis of individual returns. *Revenue of Economic Statistics 65*(3), 363–73.

Cole, S., X. Giné, and V. J. (2017). How does risk management influence production decisions? evidence from a field experiment. *The Review of Financial Studies 30*(6), 1935–1970.

Coolidge, J. (2012). Findings of tax compliance cost surveys in developing countries. *eJournal of Tax Research 10*(2), 250–287.

Coolidge, J. and D. Ilic (2009). Tax compliance perceptions and formalization of small businesses in south africa. Policy Research Working Paper 4992, World Bank Group.

Copeland, P. and A. Cuccia (2002). Multiple determinants of framing referents in tax reporting and compliance. *Organizational Behavior and Human Decision Processes 88*, 499–526.

Coricelli, G., M. Joffily, C. Montmarquette, and M. Villeval (2010). Cheating, emotions, and rationality: An experiment on tax evasion. *Experimental Economics 13*(2), 226–247.

Cottarelli, C. (2011). Revenue mobilization in developing countries. Technical report, International Monetary Fund.

Cowell, F. and J. Gordon (1988). Unwillingness to pay. *Journal of Public Economics 36*, 305–321.

Cronqvist, H., R. Thaler, and F. Yu (2018). When Nudges Are Forever: Inertia in the Swedish Premium Pension Plan. *American Economic Review: Papers Proceedings 108*, 153–158.

Cummings, R. G., J. Martinez-Vazquez, M. McKee, and B. Torgler (2009, June). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior & Organization 70*(3), 447–457.

Cyan, M. R., A. M. Koumpias, and J. Martinez-Vazquez (2016). The determinants of tax morale in pakistan. *Journal of Asian Economics 47*, 23–34.

Dabla-Norris, E., F. Misch, D. Cleary, and M. Khwaja (2017). Tax Administration and Firm Performance: New Data and Evidence for Emerging Market and Developing Economies. Working Paper 17/95, IMF.

D'Arcy, M. (2011). Why Do Citizens Assent to Pay Tax? Legitimacy, Taxation and the African State. *Afrobarometer Working Papers* (126).

De Neve, J.-E., C. Imbert, J. Spinnewijn, T. Tsankova, and M. Luts (2019). How to Improve Tax Compliance? Evidence from Population-wide Experiments in Belgium. Working paper, University of Warwick.

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

DellaVigna, S. and E. Linos (2020, March). RCTs to Scale: Comprehensive Evidence from Two Nudge Units. Working Paper.

Dhami, S. and A. al Nowaihi (2007). Why Do People Pay Taxes? Prospect Theory versus Expected Utility Theory. *Journal of Economic Behavior and Organization 64*(1), 171–92.

Doerrenberg, P. and J. Schmitz (2015). Tax compliance and information provision: A field experiment with small firms. Discussion Paper 15-028, ZEW, ZEW - Center for European Economic Research.

Drago, F., F. Mengel, and C. Traxler (2015). Compliance behavior in networks: Evidence from a field experiment. Discussion Paper 9443, IZA.

Duflo, E. (2017, May). Richard T. Ely Lecture: The Economist as Plumber. *American Economic Review 107*(5), 1–26.

Duflo, E., W. Gale, J. Liebman, P. Orszag, and E. Saez (2006). Saving Incentives for Low- and Middle-income Families: Evidnce from a Field Experiment with HR Block. *The Quarterly Journal of Economics 121*(4), 1311–1346.

Dunning, T., F. Monestier, R. Pineiro, F. Rosenblatt, and G. Tunón (2017). Is paying taxes habit forming? theory and evidence from Uruguay. Technical report, mimeo.

Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (forthcoming, 2015). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. *American Economic Journal: Economic Policy*.

Ebrahim, A., R. Gcabo, L. Khumalo, and J. Pirttilä (2019). Tax research in South Africa. WIDER Working Paper Series wp-2019-9, World Institute for Development Economic Research (UNU-WIDER).

Eisenhauer, J. G. (2006). Conscience as a deterrent to free riding. *International Journal of Social Economics 33*(8), 534–546.

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

Elffers, H., R. H. Weigel, and D. J. Hessing (1987). The consequences of different strategies for measuring tax evasion behavior. *Journal of Economic Psychology 8*(3), 311–337.

Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives 3*, 203–232.

Erard, B. (1997). Self-selection with measurement errors: A microeconometric analysis of the decision to seek tax assistance and its implications for tax compliance. *Journal of Econometrics 81*(2), 319–356.

Erard, B. and J. S. Feinstein (1994). Honesty and evasion in the tax compliance game. *RAND Journal of Economics (RAND Journal of Economics) 25*(1), 1 – 19.

Erard, B. and C.-C. Ho (2001). Searching for ghosts: who are the nonfilers and how much tax do they owe? *Journal of Public Economics 81*, 25–50.

Erard, B., P. Langetieg, M. Payne, and A. Plumley (2018). Flying Under the Radar: Ghosts and the Income Tax. Technical report, CESifo Economic Studies Conference.

Eriksen, K. and L. Fallan (1996). Tax knowledge and attitudes towards taxation; a report on a quasi-experiment. *Journal of Economic Psychology 17*, 387–402.

Falkinger, J. (1988). Tax evasion and equity: a theoretical analysis. *Public Finance 43*, 388–395.

Feige, E. L. (1990). Defining and estimating underground and informal economies: The new institutional economics approach. *World Development 18*(7), 989–1002.

Feldman, N., P. Katuščák, and L. Kawano (2016). Taxpayer confusion: Evidence from the child tax credit. *American Economic Review 106*(3), 807–835.

Fellner, G., R. Sausgruber, and C. Traxler (2013). Testing enforcement strategies in the field: Legal threat, moral appeal and social information. *Journal of the European Economic Association 11*(3).

Fjeldstad, O.-H. (2004). What's Trust Got To Do With It? Non-payment of Service Charges in Local Authorities in South Africa. *The Journal of Modern African Studies 42*(4), 539–562.

Fjeldstad, O.-H., C. Kagoma, E. Mdee, I. H. Sjursen, and V. Somville (2020, April). The Customer is King: Evidence on VAT Compliance in Tanzania. *World Development 128*.

Fjeldstad, O.-H., C. Schulz-Herzenberg, and I. H. Sjursen (2012). People's views of taxation in Africa: A review of research on determinants of tax compliance. Working Paper 8, International Centre for Taxation and Development.

Fjeldstad, O.-H. and J. Semboja (2001). Why people pay taxes: The case of the development levy in Tanzania. *World Development 29*(2059-2074).

Fortin, B., G. Lacroix, and M.-C. Villeval (2007). Tax Evasion and Social Interactions. *Journal of Public Economics 91*(8), 2089–2112.

Frey, B. S. and B. Torgler (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics 35*, 136–159.

Frey, S. B. and L. P. Feld (2002, August). Deterrence and morale in taxation: An empirical analysis. *CESIFO WORKING PAPER* (760).

Friedland, N., S. Maital, and A. Rutenberg (1978). A simulation study of income tax evasion. *Journal of Public Economics 10*(1), 107–116.

Friedman, B. (1974). Risk aversion and the consumer choice of health insurance option. *The Review of Economics and Statistics*, 209–214.

Friedman, J., T. Hastie, and R. Tibshirani (2001). *The elements of statistical learning*, Volume 1. Springer.

Gangl, K., B. Torgler, E. Kirchler, and E. Hofmann (2014). Effects of supervision on tax compliance: Evidence from a field experiment in Austria. *Economics Letters 123*(3), 378 – 382.

Gauthier, B. and R. Reinikka (2001). Shifting tax burdens through exemptions and evasion: An empirical investigation of uganda. Policy Research Working Paper 2735, World Bank Group.

Genschel, P. and L. Seelkopf (2016). Did they learn to tax? Taxation trends outside the OECD. *Review of International Political Economy 23*(2), 316–433.

Giertz, S. and J. Mortenson (2014). Policy Uncertainty and Rent Seeking by Firms and CEOs: Implications for Efficiency and Optimal Tax Rates. Technical Report 107, Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association.

Gilligan, M. J., B. J. Pasquale, and C. Samii (2014). Civil War and Social Cohesion: Lab-in-the-Field Evidence from Nepal. *American Journal of Political Science 58*(3), 604–619.

Glennerster, R. and K. Takavarasha (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton University Press.

Gordon, J. (1989). Individual Morality and Reputation costs as deterrents to tax evasion. *European economic revue 33*(4), 797–805.

Guyton, J., P. Langetieg, D. Manoli, M. Payne, B. Schafer, and M. Sebastiani (2017). Reminders and Recidivism: Using Administrative Data to Characterize Non lers and Conduct EITC Outreach. *American Economic Review: Papers Proceedings 107*(5), 471–475.

Halla, M. (2010, March). The Link between the Intrinsic Motivation to Comply and Compliance Behavior: A Critical Appraisal of Existing Evidence. IZA Discussion Paper 4843, Institute of Labor Economics (IZA).

Hallsworth, M. (2014). The use of field experiments to increase tax compliance. *Oxford Review of Economic Policy 30*(4), 658–679.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics 148*, 14–31.

Harju, J., T. Matikka, and T. Rauhanen (2019). Compliance costs vs. tax incentives: Why do entrepreneurs respond to size-based regulations? *Journal of Public Economics 173*, 139–164.

Harrison, G. W., M. Lau, and E. Rustrom (2007). Estimating Risk Attitudes in Denmark: A Field Experiment. *Scandinavian Journal of Economics 109*(2), 341–368.

Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic Literature 42*(4), 1009 – 1055.

Hasseldine, J., P. A. Hite, S. James, and M. Toumi (2007). Persuasive Communications: Tax Compliance Enforcement Strategies for Sole Proprietors. *Contemporary Accounting Research 24*(1).

Heckman, J., H. Ichimura, and P. E. Todd (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies 64*(4), 605–654.

Hernandez, M., J. Jamison, E. Korczyc, N. Mazar, and R. Sormani (2017). Applying Behavioral Insights to Improve Tax Collection: Experimental Evidence from Poland. Technical report, World Bank Group.

Hessing, D. J., H. Elffers, and R. H. Weigel (1988). Exploring the limits of self-reports and reasoned action: An investigation of the psychology of tax evasion behavior. *Journal of Personality and Social Psychology 54*(3), 405–413.

Higgins, S. and N. Lustig (2013). Commitment to Equity Assessment (CEQ): Estimating the Incidence of Social Spending, Subsidies and Taxes. CEQ Working Paper 1, Center for Inter-American Policy and Research and Department of Economics, Tulane University and Washington DC: Inter-American Dialogue.

Hofmann, E., M. Voracek, C. Bock, and E. Kirchler (2017, June). Tax compliance across sociodemographic categories: Meta- analyses of survey studies in 111 countries. *Journal of Economic Psychology 62*, 63–71.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *The American Economic Review 92*(5), 1644–1655.

Hurvich, C. and C.-L. Tsai (1989). Regression and time series model selection in small samples. *Biometrika 76*(2), 297–307.

ICTD/UNU-WIDER (2020). Government Revenue Dataset. Technical report, ICTD/UNU-WIDER.

Imbens, G. and D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences.* Cambridge University Press.

IMF (2019). Fiscal policy and development : Human, social, and physical investments for the sdgs. Technical report, IMF Policy Paper, Washington, DC.

Isbell, T. (2017). Tax compliance: Africans affirm civic duty but lack trust in tax department. Policy Paper 43, Afrobarometer.

Jakiela, P. and O. Ozier (2016). Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies. *Review of Economic Studies 1*, 231–268.

James, G., D. Witten, T. Hastie, and R. Tibshirani (2013). *An Introduction to Statistical Learning: with Applications in R*, Volume 1. New York: Springer.

Joshi, A., W. Prichard, and C. Heady (2012). Taxing the Informal Economy: Challenges, Possibilities and Remaining Questions. Working Paper 4, International Centre for Tax and Development.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*(2), 263–92.

Kahnemann, D., A. B. Krueger, D. A. Schkade, N. Schwarz, and A. Stone (2004). A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science 306*(5702), 1776–80.

Kaldor, N. (1963). Will Underdeveloped Countries Learn to Tax? *Foreign Affairs 41*, 410–19.

Kanbur, R., T. Paukkeri, J. Pirttilä, and M. Tuomala (2018). Optimal taxation and public provision for poverty reduction. *International Tax and Public Finance 25*(1), 64–98.

Kangave, J., S. Nakato, R. Waiswa, and P. L. Zzimbe (2016). Boosting Revenue Collection through Taxing High Net Worth Individuals: The Case of Uganda. Working paper 45, International Centre for Tax and Development.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics 129*(2), 597–652.

Kastlunger, B., S. Dressler, E. Kirchler, L. Mittone, and M. Voracek (2010). Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization. *Journal of Economic Psychology 31*(4), 542–552.

Keen, M. (2012). Taxation and Development – Again. Working Paper 12/220, IMF.

Kettle, S., M. A. Hernandez, S. Ruda, and M. Sanders (2016). Behavioral interventions in tax compliance : evidence from guatemala. Policy Research working paper WPS 7690, World Bank Group, Washington, D.C.

Kim, Y. (2003). Income Distribution and Equilibrium Multiplicity in a Stigma-based Model of Tax Evasion. *Journal of Public Economics 87*(9), 1591–1616.

Kinsey, K. and H. G. Gramsick (1993). Did the tax reform act of 1986 improve compliance? Three studies of pre- and post-TRA compliance attitudes. *Law and Policy 15*(4), 239–325.

Kira, A. R. (2017). An Evaluation of Governments' Initiatives in Enhancing Small Taxpayers' Voluntary Tax Compliance in Developing Countries. *International Journal of Academic Research in Accounting, Finance and Management Sciences 7*(1), 253–267.

Kirchler, E., E. Hoelzl, and I. Wahl (2008). Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology 29*, 210–225.

Kirchler, E., S. Muehlbacher, B. Kastlunger, and I. Wahl (2007). Why pay taxes? a review of tax compliance decisions. Working paper 07-30, Andrew Young School of Policy Studies.

Klepper, S. and D. Nagin (1989). The role of tax practitioners in tax compliance. *Policy Sciences 22*(2).

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica 79*(3), 651–692.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*(1), 83–119.

Kosonen, T. and O. Ropponen (2015). The role of information in tax compliance: Evidence from a natural field experiment. *Economics Letters 129*(18-21).

Lefebvre, M., P. Pestieau, A. Riedl, and M. C. Villeval (2015). Tax evasion and social information: an experiment in belgium, france, and the netherlands. *International Tax and Public Finance 22*(3), 401–425.

Levi, M. (1989). *Of rule and revenue*. University of California Press.

Levi, M., A. Sacks, and T. Tyler (2009). Conceptualizing legitimacy, measuring legitimating beliefs. *American Behavioral Scientist* (53), 354–375.

Levitt, S. and J. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives 21*(2), 153–174.

Ligomeka, W. (2019a, September). Assessing the Performance of African Tax Administrations: A Malawian Puzzle. African Tax Administration Paper 14, International Centre for Taxation and Development.

Ligomeka, W. (2019b, March). Expensive To Be a Female Trader: the Reality of Taxation of Flea Market Traders in Zimbabwe. Working Paper Series 93, International Centre for Taxation and Development.

Loewenstein, G., E. U. Weber, C. K. Hsee, and N. Welch (2001). Risk as feelings. *Psychological Bulletin 127*(2), 267–86.

Long, C. and M. Miller (2017, April). Taxation and the Sustainable Development Goals. Do good things come to those who tax more? Briefing note, Overseas Development Institute.

Lubua, E. W. (2014). Influencing Tax Compliance in SMEs through the Use of ICTs. *International Journal of Learning, Teaching and Educational Research 2*(1), 80–90.

Luttmer, E. and M. Singhal (2014). Tax morale. *Journal of Economic Perspectives 28*(4), 149–68.

MacKinnon, J. G. and L. Magee (1990, May). Transforming the Dependent Variable in Regression Models. *International Economic Review 31*(2), 315–39.

Maertens, A., A. V. Chari, and D. R. Just (2014). Why farmers sometimes love risks: evidence from india. *Economic Development and Cultural Change 62*(2), 239–274.

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

Mascagni, G. (2018). From the lab tot he field: A review of tax experimemts. *Journal of Economic Surveys 32*(2), 273–301.

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

Mascagni, G., A. T. Mengistu, and F. B. Wodleyes (2018). Can icts increase tax? experimental evidence from ethiopia. Working Paper 82, International Centre for Tax and Development.

Mascagni, G., N. Monkam, and C. Nell (2016). Unlocking the potential of administrative data in Africa: Tax compliance and progressivity in Rwanda. Working Paper 56, International Centre for Taxation and Development.

Mascagni, G., D. Mukama, and F. Santoro (2019). An Analysis of Discrepancies in Taxpayers' VAT Declarations in Rwanda. Working Paper 92, International Centre for Tax and Development.

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

Mascagni, G., C. Nell, N. Monkam, and D. Mukama (2016). The carrot and the stick: Evidence on voluntary tax compliance from a pilot field experiment in Rwanda. mimeo.

Mascagni, G. and F. Santoro (2018). What is the role of taxpayer education in Africa? African Tax Administration Paper 1, International Centre for Tax and Development.

Mascagni, G., F. Santoro, and D. Mukama (2019, February). Teach to Comply? Evidence from a Taxpayer Education Programme in Rwanda. Technical Report 91, International Centre for Taxation and Development.

Mascagni, G., F. Santoro, D. Mukama, J. Karangwa, and N. Hakizimana (2020). Acrive ghosts: Nil-filing in Rwanda. Working paper, International Centre for Tax and Development.

McCulloch, N., T. Moerenhout, and J. Yang (2020). Fuel Subsidy Reform and the Social Contract in Nigeria: a Micro-economic Analysis. Working Paper 104, International Centre for Tax and Development, Brighton: Institute of Development Studies.

McKee, M., C. Siladke, and C. A. Vossler (2018). Behavioral Dynamics of Tax Compliance when Taxpayer Assistance Services Are Avail- able. *International Tax Public Finance 25*, 722–56.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics 99*(2), 210–221.

McKerchar, M. and C. Evans (2009). Sustaining growth in developing economies through improved taxpayer compliance: Challenges for policy makers and revenue authorities. *eJournal of Tax Research 7*(2), 171–201.

Meiselman, B. S. (2018). Ghostbusting in Detroit: Evidence on nonfilers from a controlled field experiment. *Journal of Public Economics 158*, 180–193.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*, 175–208.

Monteiro, J. and J. Assuncao (2012). Coming out of the shadows? estimating the impact of bureaucracy simplification and tax cut on formality in brazilian microenterprises. *Journal of Development Economics 99*(1), 105–115.

Moore, M. (2013). Revenue reform and statebuilding in anglophone Africa. Working paper, International Centre for Taxation and Development.

Moore, M. (2020, September). What is Wrong with African Tax Administration? ICTD Working Paper Series 111, International Centre for Tax and Development.

Moore, M., W. Prichard, and O.-H. Fjeldstad (2018). *Taxing Africa: Coercion, Reform and Development.* London: Zed Books.

Mossin, J. (1968). Taxation and risk-taking: An expected utility approach. *Economica 35*(137), 74 – 82.

Myles, G. D. and R. A. Naylor (1996, April). A model of tax evasion with group conformity and social customs. *European Journal of Political Economy 12*(1), 49–66.

Nalishebo, S. and A. Halwampa (2014). Uncovering the Unknown: An Analysis of Tax Evasion in Zambia. Working Paper 18, Zambia Institute for Policy Analysis and Research.

OECD (2015). Building tax culture, compliance and citizenship.

Onu, D. and L. Oats (2016). The role of social norms in tax compliance: theoretical overview and practical implications. *Journal of Tax Administration 1*(1).

Ortega, D. and P. Sanguineti (2013). Deterrence and reciprocity effects on tax compliance: experimental evidence from venezuela. CAF Working paper 2013/08, CAF Caracas.

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

Ortega, D. and C. Scartascini (2016b). Who's calling? The effect of phone calls as a deterrence mechanism. mimeo, Inter-American Development Bank.

Palil, M. (2010, 08). *TAX KNOWLEDGE AND TAX COMPLIANCE DETERMINANTS IN SELF ASSESSMENT SYSTEM IN MALAYSIA.* Ph. D. thesis, The University of Birmingham.

Pence, K. (2006). The Role of Wealth Transformations: An Application to Estimating the Effect of Tax Incentives on Saving. *Contributions to Economic Analysis and Policy 5.*

Perez-Truglia, R. and U. Troiano (2016). Shaming tax delinquents evidence from a field experiment in the United States. Technical report, mimeo.

Pirttilä, J. (2017). Building up Tax Systems: Lessons from the Nordic Countries. In *Conference on Public economics for Development*. Embassy of Finland and UNU-WIDER.

Pirttilä, J. and F. Tarp (2019, October). Public economics and development action: an introduction to a special issue in international tax and public finance. *International Tax and Public Finance 26*(967-971).

Pissarides, C. A. and G. Weber (1989). An Expenditure-Based Estimate of Britain's Black Economy. *Journal of Public Economics 39*(1), 17–32.

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

Pomeranz, D. and J. Vila-Belda (2019). Taking State-Capacity Research to the Field: Insights from Collaborations with Tax Authorities. *Annual Review of Economics 11*(1), 755–781.

Pommerehne, W. W. and H. Weck-Hannemann (1996). Tax rates, tax administration and income tax evasion in Switzerland. *Public Choice* (88), 161–170.

Porta, R. L. and A. Shleifer (2016). *African Successes*, Volume 1. National Bureau of Economic Research.

Prichard, W. (2015). *Taxation, Responsiveness and Accountability in Sub-Saharan Africa: The Dynamics of Tax Bargaining*. Cambridge University Press.

Prichard, W., J.-F. Brun, and O. Morrissey (2012). Donors, aid and taxation in developing countries: an overview. ICTD Working Paper Series 6, International Centre for Tax and Development.

Prichard, W., A. Custers, R. Dom, S. Davenport, and M. Roscitt (2019). Innovations in Tax Compliance - Conceptual Framework. Policy Research Working Paper 9032, World Bank Group.

Reckers, P., D. Sanders, and S. Roark (1994, December). The influence of ethical attitudes on taxpayer compliance. *National Tax Journal 47*(4), 825–36.

Richardson, G. (2006). Determinants of tax evasion: A cross-country investigation. *Journal of International Accounting, Auditing and Taxation 15*, 150–169.

Roberts, M. L., P. A. Hite, and C. F. Bradley (1994). Understanding attitudes toward progressive taxation. *Public Opinion Quarterly* (58), 165–190.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

RRA (2017). *Compilation of fiscal laws and regulations in use in Rwanda.* Rwanda Revenue Authority.

Sacks, A. (2012, August). Can Donors and Non-State Actors Undermine Citizens' Legitimating Beliefs? Policy Research Working Paper 6158, World Bank, Washington, D.C.

Sandmo, A. (2005). The Theory of Tax Evasion: A Retrospective View. *National Tax Journal LVIIl*(4).

Santoro, F. and W. Mdluli (2019). Nil-filing in Eswatini: Should the revenue administration be concerned? African Tax Administration Paper 6, International Centre for Tax and Development.

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

Schwartz, R. and S. Orleans (1967). On legal sanctions. *University of Chicago Law Review 34*(274–300).

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics 6*(2), 461–464.

Segal, U an Spivak, A. (1990). First order versus second order risk aversion. *Journal of Economic Theory 51*, 111–125.

Shaw, K. L. (1996). An empirical analysis of risk aversion and income growth. *Journal of Labor Economics 14*(4), 626–653.

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

Sianesi, B. (2004). An evaluation of the swedish system of active labour market programmes in the 1990s. *Review of Economics and Statistics 86*(1), 133–155.

Siebert, M. and A. Mbise (2018, November). Toilets Not Taxes: Gender Inequity in Dar es Salaam's City Markets. Technical Report 89, International Centre for Taxation and Development.

Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives 21*(1), 25–48.

Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature 57*(4), 904–954.

Slemrod, J., M. Blumenthal, and C. Christian (2001). Taxpayer response to an increased probability of audit: Evidence from a controlled experiment in Minnesota. *Journal of Public Economics 79*(3), 455–483.

Slemrod, J., B. Collins, J. L. Hoopes, D. Reck, and M. Sebastiani (2017). Does Credit-Card Information Reporting Improve Small-Business Tax Compliance? *Journal of Public Economics 149*(2017), 1–19.

Slemrod, J. and C. Weber (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.

Slemrod, J. and S. Yitzhaki (2002). *Tax Avoidance, Evasion, and Administration*, Volume 3. Amsterdam: North-Holland.

Snow, A. and R. S. Warrren (2005). Ambiguity about audit probability, tax compliance, and taxpayer welfare. *Economic Review 43*(4), 865–871.

Song, Y.-d. and T. E. Yarbrough (1978). Tax Ethics and Taxpayer Attitudes: A Survey. *Public Administration Review 38*(5), 442–452.

Sparks, P., D. Hedderley, and R. Shepherd (1992). An investigation into the relationship between perceived control, attitude variability and the consumption of two common foods. *Journal of Social Psychology 22*(1), 55–71.

Spicer, M. W. and S. B. Lundstedt (1976). Understanding tax evasion. *Public Finance 21*(2), 295–395.

Spicer, M. W. and J. E. Thomas (1982). Audit probabilities and the tax evasion decision: An experimental approach. *Journal of Economic Psychology 2*(3), 241–245.

SRA (2018). Integrated Annual Report. Technical report, Eswatini Revenue Authority.

Srinivasan, T. N. (1973). Tax evasion: a model. *Journal of Public Economics 2*(4), 1–33.

Sugiura, N. (1978). Further analysts of the data by akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods 7*(1), 13–26.

Tanui, P. (2016). Demystifying the key taxation concepts to improve compliance by small and medium enterprises in kenya. *European Journal of Accounting, Auditing and Finance Research 4*(6), 47–60.

Tanzi, V. (1980). The underground economy in the United States: Estimates and implications. *Banco Nazionale del Lavoro Quarterly Review 135*(2), 427–453.

Tanzi, V. (2012). Equity, Transparency, Cooperation and the Taxation of High Net Worth Individuals. *Asia-Pacific Tax Bulletin*, 299–307.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*(58), 267–288.

Torgler, B. (2003). Beyond Punishment: a tax compliance experiment with taxpayers in Costa Rica. *Revista de Analisis Economico - Economic Analysis Review 18*(1), 27–56.

Torgler, B. (2004a). Cross-culture comparison of tax morale and tax compliance: Evidence from Costa Rica and Switzerland. *International Journal of Comparative Sociology 45*(1-2), 17–43.

Torgler, B. (2004b). Moral suasion: An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. *Economics of Governance 5*(3), 235–253.

Torgler, B., M. Schaffner, and A. Macintyre (2010). Tax Compliance, Tax Morale and Governance Quality. In J. Alm, J. Martinez-Vazquez, and B. Torgler (Eds.), *Developing Alternative Frameworks for Explaining Tax Compliance*. Routledge.

Tourek, G. (2020, January). Targeting in Tax Behavior: Evidence from Rwandan Firms. Job Market Paper.

Traxler, C. (2010). Social norms and conditional cooperative taxpayers. *European Journal of Political Economy 26*(1), 89–103.

UN Millennium Project (2005). Investing in Development. Policy report, United Nations, New York.

van den Boogaard, V., W. Prichard, and S. Jibao (2018). Norms, networks, power, and control: Understanding informal payments and brokerage in cross-border trade in sierra leone. Working Paper 74, International Centre for Tax and Development.

Vansteelandt, S. (2009). Estimating Direct Effects in Cohort and Case–Control Studies. *Epidiemology 20*(6), 851–860.

Vogel, J. (1974). Taxation and public opinion in Sweden: An interpretation of recent data. *National Tax Journal 27*(4), 499–514.

Voors, M. J., E. E. M. Nillesen, P. Verwimp, E. H. Bulte, R. Lensink, and D. P. Van Soest (2012). Violent Conflict and Behavior: A Field Experiment in Burundi. *American Economic Review 102*(2), 941–64.

Vossler, C. A. and M. McKee (2017). Efficient Tax Reporting: The Effect of Taxpayer Information Services. *Economic Inquiry 55*, 920–40.

Ward, P. S. and V. Singh (2015). Using Field Experiments to Elicit Risk and Ambiguity Preferences: Behavioural Factors and the Adoption of New Agricultural Technologies in Rural India. *The Journal of Development Studies 51*(6), 707–724.

Watrin, C. and R. Ullmann (2008). Comparing direct and indirect taxation: The influence of framing on tax compliance. *European Journal of Comparative Economics 5*, 33–56.

Webley, P. (1991). *Tax Evasion: An Experimental Approach.* Cambridge University Press.

Weiss, Y. (1972). The risk element in occupational and educational choices. *Journal of Political Economy 80*(6), 1203–1213.

Wenzel, M. and N. Taylor (2004). An experimental evaluation of tax-reporting schedules: a case of evidence-based tax administration. *Journal of Public Economics 88*, 2785–2799.

World Bank (2018). Eswatini - overview.

Yaniv, G. (1999). Tax compliance and advanced tax payments: A prospect theory analysis. *National Tax Journal 52*(4), 753–764.

Yesegat, W., D. E. Vorontsov, J. Coolidge, and L. O. Corthay (2015). Tax compliance cost burden and tax perceptions survey in Ethiopia. Technical report, World Bank Group, Washington, D.C.

Yitzhaki, S. (1974). A note on income tax evasion: a theoretical analysis. *Journal of Public Economics 3*(2), 201–202.

Yücedoğru, R. and J. Hasseldine (2016, December). Understanding tax morale of smes: A qualitative study. *eJournal of Tax Research 14*(3).

Zwane, A. P., J. Zinman, E. Van Dusen, W. Pariente, C. Null, E. Miguel, M. Kremer, D. S. Karlan, R. Hornbeck, X. Giné, E. Duflo, F. Devoto, B. Crepon, and A. Banerjee (2011). Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences*.