Working Paper 58

One Size Does Not Fit All:
A Field Experiment on the Drivers of Tax Compliance and Delivery Methods in Rwanda

Giulia Mascagni, Christopher Nell and Nara Monkam
January 2017

www.ataftax.org
www.ictd.ac

UKaid
from the British people

Norad

ICTD Working Paper 58

One Size Does Not Fit All: A Field Experiment on the Drivers of Tax Compliance and Delivery Methods in Rwanda

Giulia Mascagni, Christopher Nell and Nara Monkam

January 2017

One Size Does Not Fit All: A Field Experiment on the Drivers of Tax Compliance and Delivery Methods in Rwanda

Giulia Mascagni, Christopher Nell, Nara Monkam

Summary

Although field experiments in tax compliance represent a growing area of research, the literature has so far focussed exclusively on high and middle-income countries. This paper starts to fill this gap by reporting the results of a tax field experiment in Rwanda, while also highlighting some characteristics that may be common to other low-income countries. We evaluate an intervention carried out by the Rwanda Revenue Authority (RRA), which involved sending messages to taxpayers to nudge their declaration behaviour during the filing period of January-March 2016. Focussing particularly on business profits tax, our study is designed to address two interrelated questions. First, what are the key drivers of compliance in Rwanda? Second, what is the best delivery method to reach taxpayers with messages designed to improve compliance? Although other studies have explored delivery methods in the context of taxpayer communication, our study is the first one to interact these methods with different message contents. As a result, we evaluate a set of nine treatments that combine three message contents (deterrence, fiscal exchange, reminders) and three delivery methods (letters, SMS, emails) – as compared to a control group that received no message. We find that friendly approaches to taxpayers are generally more effective than deterrence. However, small taxpayers are still quite responsive to the possibility of being fined and prosecuted (deterrence). We also show that low-cost delivery methods like SMS and emails can be highly effective as compared to letters.

Giulia Mascagni is Research Fellow at the Institute of Development Studies (IDS) and Research Director of the International Centre for Tax and Development (ICTD).

Christopher Nell was Research Officer at the IDS and the ICTD, and is now Integrated Expert at the Rwanda Revenue Authority (RRA) and the Centre for International Migration and Development (CIM).

Nara Monkam is Research Director of the African Tax Administration Forum (ATAF).

Contents

Figures

Acknowledgements

Acronyms

ATAF African Tax Administration Forum
ATRN African Tax Research Network
CIT   Corporate income Tax
ICTD International Centre for Tax and Development
ITT   Intention-To-Treat
IV   Instrumental Variable
LATE Local Average Treatment Effect
OLS   Ordinary Least Squares
PAYE Pay As You Earn
PIT   Personal income tax
RWF Rwandan Franc
TE   Tax experiments
VAT   Value Added Tax

Introduction

Although increasing compliance is a challenge everywhere in the world, in low-income countries it often presents further layers of complexity related to accountability, institutional capacity, and good governance. While governments need to have the necessary capacity to enforce tax laws, citizens may be willing to comply even when they are not forced to do so. In low-income countries, both these aspects – enforcement and voluntary compliance – are severely constrained. On the one hand, administrative constraints affecting government institutions range from the difficulty of recruiting and retaining skilled staff to fewer financial resources than high-income counterparts, limiting the extent and effectiveness of tax enforcement. On the other hand, low level and quality of public services, imperfect democratic processes, and limited access to information (e.g. tax advisers, internet coverage) all represent obstacles to voluntary compliance. The experimental literature on tax compliance has largely focused on high and middle-income countries. This paper starts filling this gap by reporting the results of a tax field experiment in Rwanda, while also highlighting some characteristics that may be common to other low-income countries. It represents the first wave of such work ever carried out in the African continent.[1]

Our field experiment evaluates an intervention carried out in collaboration with the Rwanda Revenue Authority (RRA), which involved sending messages to taxpayers to nudge their declaration behaviour during the filing period of January-March 2016. Our study focuses particularly on declarations for corporate income tax (CIT) and personal income tax (PIT), although these taxpayers are usually subject to other taxes as well (e.g. Value Added Tax (VAT)). Building on the existing theoretical and empirical literature,[2] our study is designed to address two interrelated questions. First, what are the key drivers of compliance in Rwanda? In answering this question, we compare our results with the findings available in the literature, while also recognising that more evidence is needed from other low-income countries before making further generalisations. Second, what is the best delivery method for reaching taxpayers when sending messages to improve compliance? Although other studies have explored delivery methods in the context of taxpayer communication (Ortega and Scartascini 2016; Doerrenberg and Schmitz 2015), our study is the first one to interact them with different message contents. As a result, we evaluate a set of nine treatments that combine three message contents (deterrence, fiscal exchange, reminders) and three delivery methods (letters, SMS, emails), and compare these with a control group that received no message. By doing this, we introduce some significant innovations into the literature, while relying on a research design that is largely comparable with similar studies (see Section 0).[3]

As far as the drivers of compliance are concerned, most studies are based on the seminal model of Allingham and Sandmo (1972) that postulates that taxpayer behaviour is the result of an economic calculation of the costs and benefits of evasion, similar to models in the tradition of the economics of crime (Becker 1968). Allingham and Sandmo's model was later developed to include other factors, beyond the monetary ones related to deterrence (Erard and Feinstein 1994; Myles and Naylor 1996; Traxler 2010). These non-monetary factors are often referred to as 'tax morale', which is a general term encompassing peer effects, intrinsic motivations, culture, and fiscal exchange, amongst others.[4] A growing experimental literature has tested the validity of such theoretical predictions in practice. On the one hand, field experiments have shown that deterrence can be highly effective in increasing compliance

---

[1] We are aware of similar studies carried out in other African countries that, to the best of our knowledge, have not yet been published.

[2] Comprehensive reviews of theoretical and empirical studies on tax compliance are available in Hashimzade, Myles, and Tran-Nam (2013) and Mascagni (2016), respectively.

[3] More details on the design and key findings of tax experiments in the existing literature can be found in Mascagni (2016).

[4] For a more detailed discussion of tax morale and its components, see Luttmer and Singhal (2014).

(Slemrod, Blumenthal and Christian 2001; Castro and Scartascini 2015; Fellner, Sausgruber, and Traxler 2013; Ariel 2012; Bott, Cappelen and Sørensen 2014; Dwenger, Kleven, Raul and Rincke 2015). However, they also highlighted its limitations in low enforcement environments (Carrillo, Pomeranz and Singhal 2016), where 'threats of audit' cannot be followed up with actual checks by the revenue authority. In those cases, it may even be counterproductive to emphasise deterrence, because it may reinforce distrust and taxpayer resistance (see for example Fjeldstad and Semboja (2001)), as well as further undermining the credibility of the revenue administration. On the other hand, existing evidence on factors related to tax morale is more mixed, as some studies find that they matter for compliance (Hallsworth, List, Metcalfe and Vlaev 2014; Bott et al. 2014), while others fail to find an effect (Dwenger et al. 2015; Blumenthal, Christian and Slemrod 2001; Castro and Scartascini 2015; Fellner et al. 2013). The impact of tax morale also seems to be highly dependent on what specific aspect is highlighted in any given study, as well as prior motivations and beliefs on compliance (see for example Dwenger et al. 2015; Del Carpio 2014; Ali, Fjeldstad, and Sjursen 2015).

In addition to deterrence, our experiment focuses in particular on fiscal exchange, as we believe it is particularly suitable and relevant in the Rwandan context (see Section 1). Moreover, we also include a simple reminder of the deadline for filing tax returns, which serves as a control message. The reminder allows us to disentangle responses to the contents of the message from the effect of simply receiving any message. Several studies have found that receiving any message from the tax authority can generate taxpayer responses (Del Carpio 2014; Ortega and Scartascini 2016). Therefore, failing to include a control message can potentially confound the results. Our experimental design allows us to test if deterrence is more effective than fiscal exchange to nudge taxpayers to increase compliance. In the case of Rwanda, we find that a friendly approach to taxpayers, including both information about how tax revenues are spent (fiscal exchange) and gentle reminders of deadlines (control message), is generally more effective than deterrence. However, and consistent with our expectations (see Section 1.2), our sub-group analysis shows that small taxpayers and PIT taxpayers, who are generally smaller than those subject to CIT, are still quite responsive to the possibility of being fined and prosecuted (deterrence).

As regards delivery methods, the tax experiments (TE) literature has overwhelmingly used physical letters to communicate with taxpayers. However this method may not be efficient in low or middle-income countries, where communications by post are uncommon.[5] For this reason, many field TE that use letters in these countries require the assistance of tax officials, who often deliver them personally to taxpayers (Mascagni et al. 2016b; Ortega and Scartascini 2016). Although we adopt the same solution in this study, we recognise that it has at least three drawbacks. First, it represents a large burden on an already overstretched tax administration. Second, and partly as a result of the previous point, the success of the delivery process may be weakened by constraints as to how much time tax officials are able or willing to spend in trying to deliver letters. These two issues become even more relevant when considering the possibility of scaling up the intervention beyond a (randomly) selected sample. Third, the interaction with tax officials may confound and affect the impact of letters on taxpayers' behaviour. While we can partly address this issue (see Section 1.5), it is very difficult to fully control these interactions. Based on these practical considerations, we decided to experiment with other delivery methods, namely emails and SMS, which are cheaper and more manageable in a low-income context. However, besides these practical motivations, varying delivery methods also allows us to test some theoretical predictions regarding taxpayers' perceived probability of detection. In particular, the model developed by Ortega and Scartascini (2016) postulates that taxpayers may update their perceptions on the probability of being caught by the tax authority, also based on the delivery method used to

---

[5] For more details on the challenges and lessons learned regarding implementing tax experiments in low and middle-income countries, see Mascagni et al. (2016b) and Castro and Scartascini (2014).

contact them. More costly and selective methods such as, in our case,[6] the letters delivered by tax officials, may reveal higher enforcement effort than the cheaper emails or SMS. In turn, taxpayers may believe that such greater efforts are targeted at those who are suspected of evasion and who are therefore more likely to be subject to actual checks. Our results partly confirm this theoretical prediction. Based on information on the actual delivery status of our messages, we show that the effect of our treatment on those who received the messages is larger for letters than for SMS – though still smaller than for emails.[7] Cheaper methods like SMS and emails still proved to be highly effective and they resulted in significant revenue gains for the RRA. Overall, we estimate that our experiment generated almost US$9 million as extra revenue for the RRA.

The results of this study have both practical and theoretical relevance. On a practical level, we provide concrete policy recommendations to the RRA, and to revenue administrations in similar countries, on effective communication strategies to increase taxpayer compliance. These include both the key messages that should be highlighted and the way in which they are delivered. On a theoretical level, we empirically test the validity of predictions both on drivers of compliance and on delivery methods, in a low-income country. In doing so, we make two original contributions to the tax experiments literature. The first one is to expand this field of research to the African continent and to low-income countries more generally. Secondly, we combine message contents and delivery methods, in a research design that allows us to provide a complex but nuanced set of results. In addition, we include both corporations (liable to CIT) and individual companies and the self-employed (liable to PIT), which is one of the directions for future research encouraged by a recent review (Hallsworth 2014).[8] By including these taxpayer types, we can test whether individuals respond to our nudges differently than more structured organisations.

---

6   Note that Ortega and Scartascini (2016) used personal visits by tax officials as a delivery method. In our case, this was not possible due to constraints in administrative capacity and resources. However, we can still test their predictions by comparing letters to SMS or emails.

7   However, the delivery status of emails is more uncertain due to a glitch in the RRA system that prevented us from tracing many of them (see Section 1.5).

8   Other studies in the literature, such as Ortega and Scartascini (2016), have also included both corporations and individuals in their sample.

1 Context, experimental design, and empirical strategy

1.1 The Rwandan context

Rwanda is a small landlocked country in Eastern Africa, with a population of about 12 million people. After experiencing a civil war and genocide in the early 1990s, the country embarked on a path of development that resulted in sustained economic growth, progress on human and social development, and a large expansion of public services. From the tax perspective, Rwanda is often seen as a success story. The RRA was established in 1997 with substantial support from foreign donors. Since its establishment, the RRA has been collecting increasing amounts of tax revenue for the government's coffers, representing about 15 per cent of GDP in 2014 (ATAF 2016). This figure is in line with other African and low-income countries, despite the absence of significant natural resources in the country. Moreover, as compared to other African countries, Rwanda collects proportionally more income taxes – about a third of total revenue (Mascagni et al. 2016a; ATAF 2016). However, descriptive evidence (Mascagni et al. 2016a) suggests that in Rwanda non-compliance is a larger issue for income taxes than for Value Added Tax (VAT) or Pay As You Earn (PAYE) taxes, thus making it relevant to focus on those. Income taxes are typically seen as hard to collect and often represent a minor share of total revenue in many low-income countries, as opposed to the major role they play in high-income countries' budgets (Besley and Persson 2013).

In spite of its success, the RRA faces many challenges that are common to revenue administrations in other low-income countries as well. The informal sector represents about 40 per cent of national income,[9] meaning that many operate completely outside the tax net. By using administrative data, these fully informal actors are excluded from our analysis. Amongst those who are registered with the RRA, the level of evasion and avoidance is still expected to be relatively high, although these phenomena are typically hard to measure. For example, a sizeable proportion of taxpayers file tax declarations indicating zero turnover. Although this behaviour seems puzzling, there are at least two possible explanations. The first one is that this is a way to avoid fines for non-filing while still managing to avoid paying any income tax (Mascagni et al. 2016a; Mascagni and Mengistu 2016). The second one is that processes for de-registration are so complex that taxpayers prefer to remain registered and keep filing nil returns. Like some other countries in Africa, Rwanda still experiences some shortcomings related to good governance and democracy, although public institutions in the country present a high degree of internal accountability and are seen to be more efficient than in other countries (Curtis 2015). The RRA is a good example of this, as it is widely regarded as one of the most professional revenue authorities on the continent. In addition to a strong revenue performance, the RRA has also fully embraced the principles of modern tax administration – recognising the importance of a customer-oriented approach based on professionalism and the sensitisation and education of taxpayers, rather than on aggressive enforcement alone. In this context, the RRA adopts a number of measures to facilitate filing and compliance more generally. Amongst others, since 2012 taxpayers can file their returns online thanks to a new e-tax system, while micro taxpayers can even file using a mobile platform. In addition, presumptive regimes ('flat amount' and 'lump sum') are available for small taxpayers to facilitate filing and record-keeping, which would be more complex under the 'real regime'.[10]

These characteristics represent an advantage for our study, because tax field experiments require a high degree of commitment and engagement by the local revenue administration.

---

[9]   This figure is reported in Schneider and Williams 2013 and it represents an average value between 1999 and 2006.
[10]   For more information on these regimes, see Mascagni et al. 2016a.

The Rwandan context also provides an ideal setting to test the effect of friendly approaches to tax compliance, which are fully in line with RRA's current strategy. At the same time, the RRA may be special in many respects and may not be fully representative of the average tax administration in Africa. This calls for caution when trying to generalise our results to other low-income countries. Still, the RRA is not a unique case of a successful revenue authority in the continent (ATAF 2016).

In terms of public services, Rwanda has achieved almost full coverage of health and education at highly subsidised prices. For example, health insurance covers 91 per cent of the population, which constitutes the highest share in Africa (The Economist 2016).[11] Although the majority of the government budget is financed by taxes, foreign aid still contributes substantially to the public purse: about 30-40 per cent of the budget.[12] A recent dispute with donors in 2010 reinforced the government's rhetoric of self-reliance that highlights the importance of the country being independent from foreign donors and of being able to 'stand on its own feet'. In this context, the Government of Rwanda has been highlighting its recognition of taxpayers as key contributors to the country's sustainable development, for example through taxpayer appreciation days, and it has even created a fund for voluntary contributions that all citizens can make to finance public investment and the development strategy more generally.

1.2 Conceptual framework and hypotheses

As highlighted in the introduction, our analysis explores two broad hypotheses related to the drivers of compliance and delivery methods. As far as the former is concerned, motivations to comply can be broadly divided in monetary (often related to deterrence) and non-monetary (related to tax morale). Although we build on an extensive theoretical literature on the drivers of tax compliance, we do not review it here in any detail, because it would be outside the scope of this paper (for a good review, see for example Hashimzade, Myles and Tran-Nam 2013). We are particularly interested in testing whether the empirical finding that deterrence is highly effective in high and middle-income countries is also valid in low-income countries, where enforcement levels and state capacity is generally lower. The literature has shown that lower institutional capacity can severely limit the effectiveness of standard prescriptions to increase tax compliance. For example, Carrillo, Pomeranz, and Singhal (2016) show that, although the use of third-party information to monitor firms' turnover results in an increase in declared income, such increases are accompanied by higher reported costs. By doing this, firms can maintain a similar level of profit and therefore of tax. In these cases, real change in compliance behaviour can only be obtained by following up on the discrepancies highlighted by third-party information with traditional audits, which can allow the revenue authority to monitor less observable margins. In the case of Rwanda, like many other low-income countries, third-party information[13] is rarely used and traditional enforcement is severely limited, with very few audits being carried out each year.[14] In this context, taxpayers may take deterrence messages less seriously and therefore respond less or not at all. On the other hand, they may be more responsive to non-monetary drivers of tax compliance. Although we would have ideally tested more than just one element of tax morale, this was not possible due to an already large number of treatments and a limited sample size (see Section 1.4).[15] Therefore, we decided to focus on fiscal exchange, because it seems to fit the

---

[11]   Rwanda has been successful in tackling diseases that are frequent and deadly but cheap to treat (e.g. diarrhoea). However, the country still faces considerable challenges to treat chronic diseases (e.g. heart disease or cancer), which occur more often as life expectancy is rising (The Economist 2016).

[12]   As reported by the World Bank: http://www.worldbank.org/en/country/rwanda/overview (accessed 5 January 2017).

[13]   Here we refer to third-party information coming, for example, from banks or other institutions and not to the third-party reporting that occurs in the case of PAYE.

[14]   The RRA only carries out a very limited number of formal audits each year. For example, in 2015-16 it audited less than 300 taxpayers (Mascagni et al. 2016b).

[15]   For example, several studies in the literature have looked at peer effects, using messages highlighting that most pay their taxes on time and/or in full (Hallsworth et al. 2014; Blumenthal et al. 2001; Castro and Scartascini 2015). By doing this, they are meant to encourage taxpayers to join the compliant majority and perhaps elicit feelings of shame for being

Rwandan context particularly well, including the current rhetoric on self-reliance and the government's efforts to expand service delivery (see Section 1.1). Using messages on deterrence and fiscal exchange, as well as a simple reminder, we therefore test the following hypothesis:

HP1: Fiscal exchange is the most effective content, amongst the ones considered, to nudge taxpayers in Rwanda to comply more.

To explore this hypothesis (and others) in more detail, we test it separately for relevant sub-groups (see Section 1.6). We are especially interested to explore differences in taxpayers' reactions to various contents across the size distribution. In particular, small taxpayers in Rwanda differ from large ones in at least two dimensions that are relevant to our study. First, they have much less access to information on the tax system, due to the shortage of good tax advisors and to limited resources they can dedicate to tasks related to taxpaying. In low-income countries, this often results in higher compliance costs and even larger tax burdens than large taxpayers (Coolidge 2012; Yesegat, Vorontsov, Coolidge and Corthay 2015; Mascagni and Mengistu 2016). Second, it is very unlikely for the RRA to get in touch with small taxpayers with formal audits or even informal checks (Mascagni et al. 2016a). This is likely to result in a low perceived probability of detection, which may be corrected upwards (with a compliance effect) as a response to any personalised message from the tax authority, and particularly to letters. Deterrence messages may be particularly threatening to small taxpayers, who may feel they cannot afford to pay fines or engage in burdensome audit procedures. Instead, large taxpayers are less likely to respond to those messages, as they already know how to plan their tax affairs to their advantage. In fact deterrence messages in these cases can even backfire (Slemrod et al. 2001; Ariel 2012). Therefore, we would expect a larger response to deterrence in the sub-groups of small and PIT taxpayers (the latter being smaller than those subject to CIT).

Furthermore, the three message contents (see Table 1) were interacted with three delivery methods: physical letters, emails and SMS. In other words, each message content was delivered using each method, making a total of nine treatments. Therefore, we can test a second hypothesis that is related to delivery methods and how they affect the perceived probability of detection. Based on the theoretical predictions of Ortega and Scartascini (2016), we would expect taxpayers to react more to the letters than to other delivery methods. This hypothesis can be expressed as follows.

HP2: Physical letters generate the largest increase in compliance, amongst the delivery methods tested in Rwanda.

The reason for this expectation is that letters reveal a higher effort by the revenue authority, primarily because of the engagement of tax officials in delivering letters. Since this delivery method is more expensive than others, rational taxpayers may expect the revenue authority to use it particularly in cases where it expects large returns from uncovering previously undeclared income. As a result, their perceived probability of detection would increase and they would be more likely to respond to the letter. This theoretical framework is indeed in line with the practice of revenue administration in Rwanda. For example, audits are based on risk rather than allocated randomly, as the limited resources available need to be used in a targeted manner. Due to limited resources, the probability of an audit is indeed related to the fact of being visited. Moreover, anecdotal evidence shows that many taxpayers do not realise they are on the RRA's radar, even if they are registered (Mukama, Karangwa, and Hakizimana 2016). In this context, receiving a letter from a tax official can be quite an exceptional event. Therefore, we hypothesise that letters would trigger a greater reaction

part of a small minority of evaders. However, non-compliance rates in low-income countries are quite high, and thus communicating them may backfire and generate the opposite effect. If the compliance rate revealed in the message is lower than previously thought, then taxpayers may actually comply less than before.

than SMS, which are used more routinely by the RRA for communication campaigns. Similarly to the previous hypothesis, we also test the validity of HP2 on various relevant sub-groups (see Section 1.6).

Finally, by combining contents and methods, we can also test more detailed hypotheses about the interaction between the two. For example, the deterrence message paired with the letter delivery method may generate a particularly large response if the channel is the increased probability of audit, since both aspects would contribute to increasing such probability. Although we test all the nine specific combinations, we do not postulate specific theoretical predictions *a priori* on such interactions.

1.3 Research design

Based on the conceptual framework set out in the previous section, we use nine treatments to test our two broad hypotheses. These treatments interact three message contents (deterrence, fiscal exchange, reminder) and three delivery methods (letter, email, SMS), and they are compared to a control group that received no message. It is worth noting at the outset that our treatments do not vary actual parameters of the tax system, such as the actual probability of being audited or the level of public services. Instead, they aim to provide information and affect perceptions about deterrence and fiscal exchange. Table 1 summarises the details of message contents. As far as emails and letters are concerned, the treatment messages (with related pictures[16] and subject lines) were added to otherwise identical letters/emails that started with a reminder of deadlines (as in row 1, column 3 of Table 1) and ended with a line to thank taxpayers and wish them well. The text messages necessarily had to be slightly different. They did not include pictures and were shorter due to character limits.[17] However, to preserve some comparability, the body of the text messages aimed to concisely combine the body and subject line of the emails and letters (columns 2 and 3 in Table 1). The complete text of emails, letters and text messages is reported in Appendix 1.

---

[16]   The picture of the gavel used for the deterrence message was also used in Castro and Scartascini (2015).
[17]   The main body of the SMS was limited to 140 characters and the name of the taxpayer was limited to 15 characters. Taxpayers with longer names were abbreviated automatically after the 14th character.

Table 1 Summary of treatment messages' contents

Treatment (1)	Subject line (2)	Message (3)	Image (4)
Reminder	Tax filing period open until 31st March 2016	*RRA would like to inform you that you can file your tax return until 31st March 2016. For more information about the filing process and payment methods, contact the call centre (3004) or visit the RRA website (www.rra.gov.rw).*	No image
Deterrence	Pay your taxes on time and avoid fines and penalties	Reminder as above, plus: *Do you know that if you do not declare and pay your taxes on time, RRA can fine and possibly prosecute you? Pay your taxes on time and avoid fines and penalties.*	
Fiscal exchange	Pay taxes. Build Rwanda. Be proud.	Reminder as above, plus: *By paying your taxes you make it possible to educate our children, fund our healthcare, and keep us safe. Pay taxes. Build Rwanda. Be proud.*	
Control group	No message	No message	No message

In line with the best practice in behavioural economics (BIT 2012) and in the empirical tax experiments literature (Hallsworth 2014; Mascagni 2016), all messages were addressed to taxpayers using their names and were simple and concise – five short sentences at most. To maximise reach and facilitate understanding, our messages were translated in two languages: English and Kinyarwanda.[18]

The contents, including the general message and the specific phrasing, were developed in close collaboration with the RRA to make sure that they were both relevant to the Rwandan context and in line with existing laws, regulations, and practices. The RRA implemented the whole intervention in full accordance with their normal procedures. All letters and emails were signed by the Deputy Commissioner General and letters were validated by the official RRA stamp and letterhead. While SMS and emails were sent through the official RRA platform, letters were delivered personally by tax officials. In this way we kept our study as aligned as possible with a standard taxpaying situation where the RRA is the main interface for taxpayers, who were not informed of the existence of this study. Even within the RRA, we kept the research project highly confidential with information restricted to a few key people, to minimise the risk of spreading the information about the study. If taxpayers had thought that letters were simply part of a study, they might have been less encouraged to respond. Instead, this intervention was a real pilot communication strategy of the RRA, which was then evaluated as part of this study. Although we cannot know with certainty whether the information spread beyond the core RRA team, we received several reassurances on confidentiality and we did not receive any report of leakage of information about the study.

---

[18]   In emails, the two translations appeared one right after the other, while in the case of letters they were displayed on the same page. For SMS, the RRA sent two separate messages in the two languages right after each other. In the pilot experiment (Mascagni et al. 2016b), we translated messages into three languages: English, French and Kinyarwanda. However, in consultation with the RRA, we have decided to drop French in this experiment.

Emails and letters are fully comparable, based on the contents summarised in Table 1, except for the delivery method. They were both sent at the same time in early February 2016. On the other hand, SMS are different from emails and letters in more than one respect. In addition to being shorter and without images, we also sent the same SMS twice to taxpayers allocated in the relevant groups: once in early February and once in mid-March. The main rationale is that we expected SMS to be less tangible than emails and letters in particular, so that sending them twice may maximise their impact.[19]

1.4 Data and randomisation

Our administrative dataset contains anonymised taxpayer-level information that the RRA collects when it receives CIT and PIT declarations. It is worth noting that in Rwanda both these tax types are applicable to businesses, although the businesses may be of different types. While CIT declarations are filed by incorporated firms, PIT taxpayers are individual businesses or the self-employed. The latter are usually smaller, have a less institutionalised structure (or none at all), and are typically highly reliant on one individual.[20] Our sample therefore excludes those taxpayers who only receive employment income, which is subject to Pay As You Earn taxes (PAYE) that are usually withheld by the employer. As opposed to PAYE, incomes subject to PIT and CIT have to be self-declared by taxpayers – therefore offering a greater opportunity for non-compliance (Slemrod et al. 2001; Kleven, Knudsen, Kreiner, Pederson and Saez 2011). Consequently, taxpayers in our sample are likely to have some margin to increase their compliance and therefore to respond to our nudges. A companion study based on the same administrative data from Rwanda shows that the compliance gap is indeed larger for these tax types than for VAT and PAYE (Mascagni et al. 2016a), thus justifying our focus on these tax types.

The RRA administrative dataset available to us spans from the fiscal years 2012 to 2015 and includes the full population of taxpayers. We used 2014 as the baseline year to randomly allocate them to treatment groups and 2015 to observe outcomes after the intervention.[21] The dataset includes primarily financial variables, such as turnover, gross profits, and tax liability. In addition, we know the tax centre where the taxpayer is registered, the geographical area where the business operates, and the sector of activity – although the latter variable is considered unreliable.[22] Our outcome variable is the tax liability (henceforth, just 'tax') that results from the declaration. Tax, defined as such, includes the whole liability regardless of whether the actual payment happens through withholding procedures, tax credits, or actual payments. Using tax instead of income as an outcome variable allows us to look at the overall response to our treatments, including both potential changes in declared income and costs (see Carrillo et al. 2016 and Section 1.2).

The information needed to contact taxpayers is included in the RRA's taxpayer registry. However, like in many other low-income countries, the taxpayer registry in Rwanda is often incomplete and out-dated, presenting an obvious challenge in reaching taxpayers based on the information contained there. Since this issue was identified as a key challenge in the pilot study for this experiment (Mascagni et al. 2016b),[23] we could take it into account at the research design stage. Our pragmatic solution to this challenge was to restrict the sample in

---

[19] In the remainder of this paper, we will use the short 'SMS' instead of 'SMS sent twice', even if all SMS were indeed sent twice during the filing period.

[20] Note that this means that, for PIT, the individual who owns the firm would normally be the person who is receiving the message. In contrast, for CIT, it could be either the owner/CEO (especially for small companies) or the accountant (for larger companies).

[21] Data for 2015 contains information from declarations filed between 1 January and 31 March 2016, when our intervention occurred, and refers to the tax year January-December 2015.

[22] At the time of this study, the RRA was in the process of reorganising the sector information in the taxpayer registry.

[23] Note that a small number of taxpayers who were part of the pilot are also included in this experiment. However, those who participated in the pilot experiment are evenly distributed across treatment groups. Moreover, the exclusion of these taxpayers from the econometric analysis did not influence our findings.

two ways. First, we only considered taxpayers who are registered in the Province of Kigali and who had recently filed a PIT or CIT tax return.[24] Second, we selected those taxpayers that are more likely to have updated contact information in the taxpayer registry. Therefore, to be eligible for inclusion in our experiment, taxpayers had to fulfil one of the following two conditions: 1) to have registered with the RRA in 2013 or 2014;[25] or 2) to be registered for the e-tax system. At the time of registration, both with the RRA and with the e-tax system, taxpayers are required to confirm their contact details. As a result, this restricted sample minimises the possibility of having absent or outdated contact details, as these would have been provided within a couple of years prior to our intervention. Still, even after restricting the sample according to these conditions, we could only include those taxpayers for whom contact details were available. Eventually, we included 79 per cent of the registered CIT and 14 per cent of the PIT filers in Kigali. Since only a relatively small number of PIT taxpayers had all the necessary contact information, we only used SMS for PIT taxpayers as the phone number seemed to be the information available most frequently. As far as CIT taxpayers are concerned, the sample size was instead sufficient to test all nine treatments.[26] We report power calculations in Appendix 4.

Since our sample was necessarily constrained by the availability of contact information, we can only consider a selected group of taxpayers. More specifically, taxpayers in our sample are relatively larger and, relatedly, pay relatively more tax than the broader population of taxpayers. This selection implies two limitations. The first one is that we cannot argue that we would obtain the same results if we were to scale up the intervention to the entire population of taxpayers, since ours is a selected sample. However, until the tax registry is fully updated (an effort to do so is underway at the RRA), a full scale-up will not be feasible. In this study we are more interested in testing the concept[27] that nudges can work in a low-income context to increase compliance – and to find out how, using the hypotheses described in Section 1.2. The second limitation is that we can only test the full set of treatments for incorporated businesses that file CIT returns. For PIT taxpayers, we are unable to test the hypothesis on delivery methods (HP2).

Based on this sample, each taxpayer was assigned to one of the ten groups (nine treatments plus one control) based on stratified randomisation. In line with the best practice on randomised evaluations,[28] we decided to stratify to achieve balance, maximise statistical power, and allow sub-group analysis. After consultation with the RRA, we chose two variables as strata, both of which are discrete and expected to be highly related to the outcome. The first stratum divides taxpayers by regime type. About 55 per cent of the taxpayers in the CIT sample were subject to the stricter reporting requirements of the real regime (see Section 1.1), while the same figure is 22 per cent for PIT. The second stratum identifies 'zero-tax' filers (see Section 1.1), which captures respectively 53 per cent and 28 per cent of CIT and PIT taxpayers.[29] Although we did not use size as a stratum, we check the balance between treatment groups before performing sub-group analysis based on size in Section 2. Our randomisation was successful in statistical terms (p-values always higher than 5 per cent), as confirmed in the balance tests reported in Appendix 2.

---

[24] Most of the taxpayers included in our sample had filed in 2014, the year prior to our intervention. However, to increase sample size, we also included some CIT taxpayers who filed in 2013 (about 5 per cent of the CIT sample), as well as PIT taxpayers who had last filed in 2012 (2 per cent of the PIT sample) and 2013 (12 per cent of the PIT sample) respectively. Note that although we include these taxpayers in the randomisation, some of them did not file in 2015, or in 2014, and are therefore excluded from the analysis (see Section 1.5). Regarding the restriction of the sample to registrations in Kigali, see Mascagni et al. 2016b.

[25] Taxpayers who registered at the RRA for the years 2013 and 2014 may have already been active taxpayers before these years, but they updated their registration details in 2013 or 2014.

[26] To make sure all our treatment cells were comparable, we selected the sample so that everyone needed to have all contact information (address, phone, email) to be eligible, regardless of what delivery method they would be assigned to.

[27] See Glennester and Takavarasha 2013: 91.

[28] For example, see Glennester and Takavarasha 2013.

[29] The lower number of zero-tax filers in the PIT sample is due to the higher proportion of taxpayers in the lump-sum and flat amount regimes, which require taxpayers to make a positive tax payment (Mascagni et al. 2016a).

1.5 Implementation

Before the implementation of the intervention we briefed all tax officials who were involved in the delivery of letters, as well as the call centre, and the heads of key departments such as audit and taxpayer services. For those who knew about the study, the briefing stressed the importance of keeping the research confidential and of directing all potential taxpayer enquiries to the standard RRA channel, the call centre. Officials working at the call centre were informed that the letters aimed to provide information to taxpayers and did not imply an upcoming audit. We also encouraged call centre officials to collect information about taxpayer enquiries regarding the messages, although this proved quite challenging.[30] Officials involved in the delivery of messages were instructed to collect information about the actual delivery of the messages, in the form of a return slip for letters or a delivery report for SMS and emails.

The RRA sent messages to all selected taxpayers during the filing period between 1 January and 31 March 2016, when they had to file their declarations for the fiscal year 2015. Initially, we planned to send out the emails, letters, and the first round of text messages at the beginning of January 2016. However, due to unforeseen delays, all messages were eventually delivered in the first week of February 2016. The second round of text messages was sent in mid-March, two weeks before the deadline.

We received delivery reports for most SMS and letters, showing delivery rates of about 97 per cent for SMS[31] and 53 per cent for letters.[32] Although the compliance rate for the latter is relatively low, it generates a sufficient difference in exposure between our treatment and control groups. As far as emails are concerned, a glitch in the RRA server means that we cannot be sure about the delivery to many taxpayers in these treatments. Although the RRA received a failed delivery message from roughly 39 per cent of the email sample, at least five of these taxpayers responded to the messages. This could indicate that at least some of the messages were delivered at a later stage, potentially with some delay. In addition, we do not have information about the delivery status of another 50 per cent of the email sample.

Although most taxpayers filed in the second half of the taxpaying period, about 3 per cent of our CIT sample[33] had already filed by the time they received our treatment message on 1 February. In addition, another 13 per cent of our CIT sample failed to file by the deadline of 31 March. These taxpayers may have declared later or may not declare at all. The comparable figures for PIT are 10.5 per cent for early filing and 21 per cent for late or non-filing.[34] As a result, around 84 and 69 per cent respectively of our CIT and PIT samples declared for 2015 between 1 February 2016 and 1 April 2016.[35] This is relevant to our

---

[30] For more details on this and other challenges we encountered in the implementation of the experiment, see the report on the pilot to this study (Mascagni et al. 2016b) and the summary of taxpayer feedback (Mukama et al. 2017).

[31] This refers to the numbers of SMS successfully delivered in the second round, although we cannot infer the messages were actually read. Due to a glitch in the RRA system, we did not receive delivery reports for the first round of SMS.

[32] The share of delivered letters using the control and the public service message are very similar (delivery rate of 56 and 57 per cent, respectively). However, the number of delivered deterrence messages is considerably lower (success rate 47 per cent). These differences were unexpected and are puzzling to us and to the RRA, because the auditors, who were responsible for the letter delivery, did not know whether they delivered a deterrence message or not (since the letters were sealed in envelopes). Order effects in the delivery process may partially explain the different success rates in the letter delivery process (e.g. the letters could have been handed over to auditors in the order they were printed and the auditors may have stopped all delivery attempts after some days).

[33] These firms are mainly subject to the flat and lump-sum regimes (65 per cent out of these 291), which face less stringent bookkeeping requirements.

[34] The relatively higher number of early PIT declarations is largely due to the relatively higher share of PIT filers in the flat and lump-sum regimes (94 per cent of the 293 early filers are in the flat or lump-sum regimes).

[35] The official filing period ended on 31 March 2016. However, there were long queues at the RRA headquarters and computer servers were temporarily down on 31 March 2016. As a result, several taxpayers were only able to file on the next day, 1 April 2016 (indeed 59 taxpayers filed on 1 April, while only three more filed on 2 April). While this paper includes filings on 1 April 2016 in its analysis, the results are almost identical and the main findings do not change if we

analysis because these taxpayers constitute the relevant sample for our statistical estimations. Importantly, using Fisher's tests, we do not find statistical differences in the likelihood of having declared after our experimental intervention across the CIT and PIT, respectively, control and treatments groups. Therefore, the delayed implementation did not have an impact on the balance of our randomisation.

1.6 Empirical strategy

To estimate the treatment effects, we use the following equation:

$$Tax_i = \alpha + \sum_{j=1}^{9} \beta_j Treatment_{ji} + \gamma X_i + u_i,$$

where *i* indicates the taxpayer and *j* is an index for the treatments, nine for the CIT sample and only three for the PIT sample. *X* is a vector of taxpayer characteristics, such as a binary variable indicating whether a firm is large (according to registration of the taxpayer at the RRA's large or top-medium taxpayers' office), binary variables for the geographical location,[36] a binary variable identifying zero-tax filers in the previous year, lagged gross profit, and an interaction variable between the latter two variables.[37] The dependent variable is the CIT or PIT tax liability, depending on the sample used, as reported in tax declarations relative to the financial year 2015. All monetary variables are expressed in Rwandan francs (RWF).[38] For a description of the variables used in our regressions, see Table 8 in Appendix 3.

As mentioned in Sections 1.1 and 1.4, many taxpayers in our sample have a tax liability equal to zero, either because they declare a turnover of zero or because they declare losses instead of profits. Therefore, our dependent variable is effectively censored at zero. Intuitively, one could think about the 'willingness to pay tax' as a latent variable (which can be negative), while we can only observe the actual tax liability that is strictly equal to or greater than zero.[39] Regardless of how much taxpayers may dislike paying taxes, the minimum they can pay is zero. Similarly, regardless of how much they under-declare their income, the minimum tax liability they can achieve would always be zero. In this situation, using Ordinary Least Squares (OLS) on our sample would lead to inconsistent estimates. To solve this problem, we estimate the equation above with two models. The first one is a standard OLS, which however we estimate on a sample that excludes zero-tax filers in the baseline. As a consequence of using zero-tax filers in our stratification, this sub-sample is still balanced across treatments. Second, we estimate our equation with a tobit model, using the full sample. The tobit model is especially suited to censored data and it has been used in other studies in the tax experiments literature, for the same reason that motivates us to adopt it here (Slemrod and Weber 2012; Alm and McClellan 2012; Alm, Cherry, Jones and McKee 2010; Coricelli, Joffily, Montmarquette and Villeval 2007). However, it is particularly sensitive to deviations from the assumptions of normality of the error. Whenever this assumption is not

---

satisfied, we also present a two-part tobit model, which maintains consistency of the estimator even when normality is violated (Cameron and Trivedi 2010).

Particularly in a small country like Rwanda, it is possible that taxpayers talk to others in their network about our treatment messages, therefore potentially generating spillover effects. Although most tax experiments in the literature do not take spillovers explicitly into account (Mascagni 2016), recent studies have shown that they can be sizeable (Drago, Mengel, and Traxler 2015; Carrillo, Castro, and Scartascini 2016). Although we do not formally consider spillovers across firms, we use cluster-corrected standard errors based on districts.[40] As such, our errors are robust to heteroscedasticity and allow for correlation of observations within districts. However, this is only a very imperfect and partial solution to the problem of spillovers.[41] The major drawback of using cluster-standard errors is that they have a tendency to over-reject the null hypothesis if the number of clusters is relatively small (i.e. the actual size of the test is higher than suggested by the standard t and F statistics). Consequently, there could be a tendency to find too many significant treatment effects, even if the effects actually only occur by chance. To tackle this issue, we use a small cluster correction based on a T distribution with G-L degrees of freedom, where G is the number of clusters and L the number of coefficients (Bell and McCaffrey 2002; Cameron and Miller 2015).

We estimate the treatment effects both based on the original assignment to the treatments (intention-to-treat, or ITT) and on the information described in Section 1.5 on whether taxpayers actually received our letters (local average treatment effects, or LATE). On the one hand, ITT estimates are particularly relevant when providing policy recommendations, since they reflect the fact that treatments cannot be mandated (Bloom 2008) – that is, the RRA cannot force taxpayers to receive and read our messages. ITT provides the expected effect of an intervention taking into account that some messages will not be delivered. On the other land, LATE shows the estimated treatment effect on the compliers (the group of taxpayers who received the messages), using the treatment assignment as an instrumental variable (IV) for the actual treatment. It therefore provides an indication of the effectiveness of the intervention on those who effectively received the treatment. In addition, it also allows us to compare treatments that may differ in terms of delivery rates. This is particularly relevant to making comparisons across delivery methods, as we had a considerably higher share of delivered messages for SMS than for letter and email. With LATE, we can infer whether any difference in the effects of these delivery methods is due to the different delivery rates or to an actual difference in effectiveness. In our LATE analysis, we conservatively assume that both the emails that got lost from the server and the emails with returned delivery failure notifications were delivered.[42]

Finally, we perform the analysis across three sets of sub-groups of our overall sample. The first one distinguishes individual businesses that pay PIT from corporations that are subject to CIT. The second one looks more specifically at zero-tax filers, who are a sizeable proportion of our sample. The third one divides taxpayers into two groups based on size: the largest 10 per cent and the remaining 90 per cent. This distinction is relevant because the

---

[40] While most of the taxpayers are from three districts in Kigali (Gasabo, Kicukiro, and Nyarugenge), our sample includes CIT filers from 26 different districts. We cluster on districts rather than sector location (i.e. the administrative level below districts in Rwanda: e.g. the sector Kiyovu is part of the district Gasabo), because the literature on cluster-robust standard errors generally recommends to cluster on the highest available cluster level (see Cameron and Miller 2015 for an overview). While the main conclusions of this study are generally only mildly affected if we cluster on sector location instead of districts, the control letter loses its statistical significance if we cluster on sectors (compare to results of Table 2: ITT results for CIT sample in Section 2.1). Results using clustering on sectors are available from the authors upon request.

[41] Ideally, our design would have considered spillovers more explicitly by randomising both across and within clusters (e.g. including entire districts in which everyone is part of a certain treatment group and otherwise). In our case, we would have faced several difficulties in trying to implement such a design, including a limited sample size, the concentration of taxpayers in urban areas, and imprecise information on addresses.

[42] Therefore, we consider that 87 per cent of taxpayers received the public service email, 99 per cent the deterrence email and 83 per cent the reminder email.

largest 10 per cent of taxpayers in Rwanda, as in most low income countries, contribute the largest share of tax revenues (Mascagni et al. 2016a). The first two sets of sub-groups were taken into account in our randomisation, and are therefore balanced by design. The third one was not a stratum, but the resulting sub-groups are balanced nonetheless,[43] therefore allowing the sub-group analysis of treatment effects. Due to the larger sample size, we report results for the zero-tax and size sub-groups only for the CIT sample.

2 Results

2.1 CIT results

Table 2 shows the intention-to-treat estimation including all taxpayers who declared for 2015 after 1 February 2016 and by 1 April 2016. All treatment effects in the table are to be interpreted relative to the control group. In addition to the effects of all nine treatments, the table also reports results for pooled treatments by content and by delivery methods. Column 1 estimates a tobit model including all CIT filers in our sample. Since the assumption of normality of the error term for this model is rejected, we also estimate a two-part tobit model in columns 2 and 3. Column 2 reports the first part, a probit model including the full sample and explaining the binary decision of being a 'zero-tax' filer. Column 3 reports the second part, an OLS restricted to those taxpayers who have a positive tax liability in 2015. Finally, column 4 reports the results for the sub-group of taxpayers with zero CIT due in the baseline and is based on OLS estimates, since censoring in this sub-sample is less severe than with the full sample of taxpayers (see Section 1.6).[44] Columns 1 and 2 report marginal effects evaluated at the mean – therefore making the figures relatively more comparable with the OLS coefficients of columns 3 and 4.

---

[43]   The two-sided Fisher's tests to check the balance between the control and treatment group is not significant for any treatment. The null hypothesis is that the classification in the top turnover decile is not associated with the assignment into the treatment or control group.

[44]   Most zero-tax filers in the baseline still report zero tax in 2015. However, some of those with a positive tax in 2015 reported zero tax in the baseline.

Table 2 ITT results for CIT sample

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
SMS public service	1,407,199.26*** (153,442)	-0.04*** (0)	3,544,368.63** (1,292,235)	4,550,480.08*** (1,036,797)
SMS deterrence	379,518.08 (500,337)	-0.03 (0)	-245,033.48 (1,860,993)	324,137.34 (1,781,739)
SMS reminder	-15,902.13 (240,477)	0.00 (0)	1,241,134.24 (2,763,019)	2,331,515.56 (2,840,454)
Letter public service	707,583.03 (1,266,081)	0.00 (0)	3,796,213.90 (3,355,908)	4,388,817.55 (3,113,822)
Letter deterrence	634,482.54 (739,065)	-0.03* (0)	1,231,126.13 (1,959,400)	903,638.99 (2,053,014)
Letter reminder	1,119,430.64** (426,378)	-0.02 (0)	5,809,435.63* (2,817,386)	5,602,792.51* (3,089,071)
Email public service	345,458.48 (1,126,076)	-0.01 (0)	1,967,733.51 (1,723,669)	-783,095.80 (2,332,100)
Email deterrence	430,401.07 (485,345)	-0.00 (0)	2,993,798.13** (1,339,807)	3,697,592.20*** (1,208,896)
Email reminder	2,664,015.28** (898,269)	-0.04*** (0)	10,639,216.85** (4,964,584)	9,308,465.76 (5,432,401)
Public service	823,187.26 (625,114)	-0.02 (0)	3,096,329.43 (1,810,455)	2,714,466.70 (1,962,408)
Deterrence	481,613.28 (529,090)	-0.02 (0)	1,300,026.49 (843,973)	1,643,254.83* (928,069)
Reminder	1,273,011.31** (457,644)	-0.02** (0)	5,967,428.97** (2,234,761)	5,726,421.86** (2,101,499)
Email	1,166,352.28* (676,493)	-0.02* (0)	5,288,674.58* (2,623,881)	4,072,925.37 (2,855,417)
SMS	594,526.72*** (176,020)	-0.02*** (0)	1,521,214.71 (1,341,336)	2,400,533.44* (1,324,915)
Letter	823,157.22 (681,655)	-0.02 (0)	3,608,689.17* (1,997,274)	3,645,078.22** (1,662,965)
Observations	9096	9096	4053	4002

Notes: Cluster robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1, using a T distribution with G–L degrees of freedom. Columns 1 and 2 report marginal effects evaluated at the mean. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for size, lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit. All controls are highly significant but are omitted for the sake of clarity. Reported coefficients for pooled treatments by methods (email, SMS, letter) and contents (public service, deterrence, reminder) result from two separate regressions where the disaggregated treatments are not included.

The first row of Table 2 shows that the public service SMS treatment was effective both to reduce the number of zero-tax filers and to increase tax revenues from those with a positive liability. The probability of declaring zero tax decreased by four percentage points as a result of this SMS (column 2). Conditional on having a positive tax liability, a taxpayer in this treatment group declared on average RWF 3.5 million more than a taxpayer in the control

group (column 3). This effect is economically large, being about 44 per cent of the average tax due for those with positive tax liability in the control group. The public service SMS, concisely emphasising nation building and pride in paying taxes, was generally more effective than the other SMS treatments. While this message is significantly different from the SMS reminder in the tobit model (*p*-value of 0.00; column 1), it does not statistically differ in the OLS estimation (column 3). As far as letters are concerned, the control treatment, providing a simple reminder, increased reported tax by over RWF 1.1 million on average for the full sample (column 1) and by more than RWF 5.8 million amongst those with positive tax liability (column 3) – representing over 70 per cent of the average liability of non-zero tax taxpayers in the control group. This effect size is similar, RWF 5.6 million, when we only consider taxpayers with positive tax due in the baseline (column 4). Consistent with our first hypothesis (HP1), neither the deterrence SMS nor letter is significant.

The results on emails are largely in line with those on letters, with the control email having the largest ITT effect of all nine treatments: almost RWF 2.7 million in the full sample (column 1) and over RWF 10 million amongst those with positive tax liability (column 3). For email, the deterrence message seems to be effective, particularly amongst the sample of taxpayers with a positive (non-zero) tax liability – both in 2015 and in the baseline (columns 3 and 4). However, the coefficients for deterrence are always substantially smaller than those for the reminder,[45] indicating that the deterrence content does not have any additional effect to increase compliance compared to the content of the reminder. This result is likely to be related more to the credibility of deterrence in a low-enforcement environment than to the effectiveness of actual deterrence, which has been proven highly effective in other contexts.

The pooled treatments by contents show that the simple reminder is the most effective content.[46] Deterrence is significant only in column 4 at the 10 per cent level, largely as a result of the significance of the deterrence email. However, its effect is smaller and statistically different (smaller) to the reminder. The pooled treatments by delivery method seem to suggest a ranking, in order of effectiveness, of emails, letters, and SMS. However, these coefficients are not statistically different from each other.

The results of the probit model in column 2 show that, although the messages had some statistically significant effect in making taxpayers switch to a positive tax liability, these effects are relatively small. In other words, our treatments do not seem to be particularly effective in encouraging taxpayers to switch from a zero-tax status to a positive tax liability. Most of the effects we observe in the full sample come from taxpayers with a positive tax liability.

Table 3 presents similar results to Table 2, but shows the treatment effects on compliers – those who effectively received the message. The LATE estimates are obtained by using an instrumental variable model where the original treatment assignment is an instrument for the actual treatment. By doing this, the LATE regression intends to estimate the causal effect on taxpayers who actually received the RRA notification (SMS, letter, or email) and represents the effect of the treatment on the compliers (Angrist and Pischke 2009). As such, we expect treatments to have larger effects on compliers – which is indeed the case in Table 3.

Not surprisingly, the letter interventions appear much more effective in the LATE estimations, which is a result of the relatively high number of undelivered letters (see Section 1.5). For taxpayers with positive tax liability (columns 3 and 4), the letters increased declared taxes significantly more than the SMS, but their effect is not statistically different from that of emails. These findings suggest that while SMS currently constitute an effective tool to increase compliance, letters and emails yield higher revenue. The main challenge is related

---

[45]   This difference is statistically weakly significant in column 3 but not in column 4.
[46]   Note that the results on pooled treatments by content and delivery method are obtained using two separate regressions where the disaggregated nine treatments are not included.

to scaling up letters, as the delivery process is particularly complex, uncertain, and burdensome. In the current circumstances, it seems that non-traditional delivery methods are more feasible and cost-effective.

The results based on OLS models (column 3 and 4) can be directly compared between tables 2 and 3 and show a largely consistent picture. In contrast, the LATE estimates for tobit and probit, presented in columns 1 and 2, are not marginal effects and are therefore not directly comparable to the ITT coefficients, at least in terms of magnitude.[47] However, they largely support the previous results, both in terms of the sign and statistical significance of the treatment variables. In particular, the public service SMS and the letter and email reminders remain very effective especially for taxpayers with a positive tax liability. The coefficients for pooled deterrence and reminder treatments are statistically significant, while the one for deterrence is significantly lower than the reminder one (columns 1, 3, 4) – thus largely confirming previous results.

Finally, we can calculate the overall revenue gain obtained by the RRA thanks to our intervention. To do so, we use the ITT marginal effects on the full sample (column 1 of Table 2). We first calculate the average effect of the treatments (i.e. the average value of the nine marginal effects). This number shows the additionally declared tax due of a taxpayer in a treatment group relative to a taxpayer in the control group. Then, we multiply the average additional tax due per taxpayer by the total number of taxpayers in our CIT treatment groups who declared between 1 February and 1 April (8,028 taxpayers). The goal of this calculation is to obtain an estimate on how much less tax would have been declared if no firm in the treatment group had received a message (equivalently, if all firms were in the control group). Given this calculation, the CIT experiment led to additional due taxes of more than RWF 6.8 billion, which is around US$8.3 million. The additional due taxes are still around RWF 4.6 billion (US$5.6 million) if we conservatively use only the statistically significant treatments of column 1 of Table 2. In relative terms, our intervention resulted in an increase in compliance in the range of 16.0 to 23.7 per cent.[48] These revenue gains can be compared with a very modest cost of sending messages. A rough calculation including all costs, including personnel cost to coordinate message delivery, results in an estimated cost per message of £0.90 for letters, £0.60 for SMS, and £0.20 for emails.

---

[47]   The LATE coefficients do not refer to marginal effects, because we used a two-step tobit estimator, as the maximum likelihood tobit had difficulties in converging due to the presence of multiple endogenous variables (i.e. multiple treatment variables).

[48]   The additional revenue is more than RWF 12 billion if we use the OLS estimation displayed in column 3 of Table 2 as our baseline (41.9 per cent increase in compliance among those who declared a positive tax due).

Table 3 LATE results for CIT sample

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
SMS public service	5,363,942.83***	-0.10***	3,569,631.66***	4,623,205.09***
	(322,324)	(0)	(858,448)	(970,625)
SMS deterrence	1,461,515.06	-0.07	-248,942.20	316,401.84
	(1,299,268)	(0)	(2,653,145)	(2,114,499)
SMS reminder	-94,180.98	0.00	1,228,229.79	2,319,468.48
	(1,092,264)	(0)	(1,764,621)	(2,400,879)
Letter public service	3,466,679.15	0.01	5,724,642.38	6,929,320.36
	(10,212,368)	(0)	(4,958,903)	(4,939,256)
Letter deterrence	4,834,456.11	-0.16	2,276,954.08	1,675,666.59
	(5,730,935)	(0)	(3,449,282)	(3,436,177)
Letter reminder	7,421,150.33***	-0.06	9,371,466.70***	9,326,503.16***
	(453,880)	(0)	(1,937,233)	(2,284,360)
Email public service	1,449,387.58	-0.04	2,218,042.05	-961,520.75
	(2,484,610)	(0)	(1,820,440)	(2,670,881)
Email deterrence	1,648,146.34	-0.00	3,027,511.72*	3,727,109.56***
	(2,113,767)	(0)	(1,642,372)	(1,198,950)
Email reminder	12,183,058.93***	-0.13***	12,695,332.35*	11,088,858.82*
	(4,278,392)	(0)	(6,954,886)	(6,303,194)
Public service	3,592,740.45	-0.05	3,649,319.31	3,249,755.25
	(3,019,863)	(0)	(2,116,150)	(2,562,327)
Deterrence	2,231,539.12	-0.06	1,553,364.09**	1,954,857.40***
	(2,504,690)	(0)	(552,263)	(672,286)
Reminder	6,060,000.12***	-0.06***	7,306,421.44**	7,046,330.51**
	(2,131,204)	(0)	(2,825,984)	(2,619,055)
SMS	2,267,910.69***	-0.06***	1,527,920.10***	2,414,571.87***
	(670,466)	(0)	(300,191)	(583,228)
Letter	5,376,106.00	-0.07	5,941,759.64**	6,173,069.90**
	(5,409,949)	(0)	(2,245,117)	(2,267,034)
Email	4,937,561.22	-0.06*	5,868,725.26*	4,488,731.19
	(2,892,354)	(0)	(3,353,987)	(3,197,734)
Observations	9096	9096	4053	4002

Notes: Cluster robust standard errors using a jackknife variance estimation are in parentheses. *** p<0.01, ** p<0.05, * p<0.1, using a T distribution with G–L degrees of freedom. All columns estimate instrumental variable (IV) models. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for size, lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit. Reported coefficients for pooled treatments by methods (email, SMS, letter) and contents (public service, deterrence, reminder) result from two separate regressions where the disaggregated treatments are not included.

2.1.1 Size sub-groups for CIT sample

To get into further detail on our results, we explore sub-groups based on size. More specifically, we divide the CIT sample in two groups, including respectively the bottom 90 per cent according to turnover in the baseline, and taxpayers with the highest turnover (top 10

per cent). Reporting results on these sub-samples allows us to investigate whether our results differ amongst taxpayers of different size (see Section 1.2). Table 9 and Table 10 (Appendix 1) report the results on the pooled treatments respectively for large taxpayers and for small and medium ones, as defined above. Not surprisingly, the treatment effects are much larger in monetary terms for large taxpayers, as they have higher turnovers and therefore pay more tax in absolute terms. Interestingly, however, the effect of the reminder is also larger as a proportion of tax due in the control group (with positive tax). For example, the additional tax due resulting from any reminder message (pooled letter, email, and SMS reminder treatments) is 79 per cent and 6 per cent of the average in the control group, respectively for large and smaller taxpayers.

However, consistent with our expectations (see Section 1.2), small taxpayers responded more to deterrence than to other treatments. The coefficient on the pooled deterrence treatments is highly significant amongst smaller taxpayers, while it is only significant in the LATE estimation for larger ones. More importantly, it is larger and significantly different from the reminder (column 1), indicating that deterrence has an additional effect to the reminder for small taxpayers. Amongst large taxpayers, in contrast, the effect of deterrence is always smaller and statistically not different to a simple reminder. This suggests that smaller taxpayers are more likely to react to threats than larger companies, who may already be aware of possible fines and sanctions due to better knowledge of tax law, or access to tax advisers. Indeed, as a result of the deterrence messages, several taxpayers approached the RRA to find out more about potential fines and sanctions.

As far as delivery methods are concerned, emails seem to be particularly effective for small taxpayers as they always yield a larger effect than other delivery methods, including letters. This difference is statistically significant amongst taxpayers with a positive tax liability (second part of the two-part tobit model even when we take into account the relatively lower delivery rate for letters in the LATE estimation). Although SMS yield significant increases amongst smaller taxpayers, they are smaller and statistically different from emails (all LATE estimations) and significantly different from letters (column 4 of Table 10). Similarly, letters (more so) and emails have generally significantly larger effects than SMS amongst large taxpayers.

This finding, combined with the results on deterrence, could suggest that for larger taxpayers being approached by an RRA officer through the letter delivery is relatively more important than the content of the message itself, while for smaller taxpayers the combined threat of the content and the delivery method might be crucial. The results on all disaggregated treatments,[49] although weaker due to a smaller sample size, confirm that the deterrence letter has a significant and larger effect than the letter reminder – the latter difference being however only significant in the tobit specification.

2.2 PIT results

Results for PIT are less complex than those for CIT, mostly as a result of a smaller sample. The number of treatments here is limited to the three message contents delivered by SMS (see Section 1.3). In contrast to the CIT analysis, for PIT we do not separate taxpayers based on turnover (top 10 per cent and bottom 90 per cent) partly because of the smaller sample and partly because all PIT taxpayers are relatively small, and therefore display a very similar behaviour across the distribution. Table 4 reports the intention-to-treat estimates for the PIT sample, while Table 5 reports the treatment effects on compliers based on LATE. Both tables report results from the tobit model on the full sample (column 1); the two parts of the tobit (columns 2 and 3); and those from the sub-sample of taxpayers with positive tax in the baseline (column 4). All regressions include control variables for lagged gross profit and

---

[49]   Available from the authors upon request.

district, a dummy variable indicating lagged zero tax due, and an interaction variable between lagged zero tax due and lagged gross profit.[50]

The results for the PIT sample are in line with the previous ones. The reminder SMS seems to be effective particularly in the decision to pay a positive tax amount, but also for the tax amount – more so in the LATE estimation (Table 5). The public service SMS increased the tax due on average by over RWF 350,000 in the full PIT sample and by over RWF 850,000 if we exclude taxpayers with zero tax liability in the baseline. These effects are large, representing respectively 28 and 70 per cent of the average tax due for those with a positive liability in the control group. Consistently with the previous results on small taxpayers, the effect of deterrence in the PIT sample is significant – although it is always smaller than the public service and the reminder SMS. However, the three treatment coefficients are only significantly different from each other in the tobit model, both in the ITT and the LATE estimations. Therefore, at least amongst those with a positive tax liability, the specific content of the message does not seem to have an effect in addition to the pure fact of receiving any message from the RRA (in this case, the reminder).

To estimate the additional PIT tax revenue resulting from our experimental intervention, we use the marginal treatment effects of the main ITT regression shown in column 1 of Table 4. As the three PIT treatments are jointly statistically significant, we first calculate the average effect of the three treatments, which gives the additionally declared tax due of a taxpayer in a treatment group relative to a taxpayer in the control group. We then multiply the average additional tax due per taxpayer by the total number of taxpayers in our PIT treatment groups who declared between 1 February and 1 April (1,445 taxpayers). This calculation suggests that our intervention led to additional due taxes of roughly RWF 316 million amongst PIT taxpayers, which is almost US$380,000. Taken together with the revenue gains from the CIT sample, the total gains from this experiment are about US$8.7 million. The additional PIT due is still more than RWF 292 million if we conservatively use only the two statistically significant treatments of column 1. On average, our SMS communication with PIT filers resulted in an increase in compliance of between 24.1 and 26.3 per cent.[51]

Table 4 ITT results for PIT sample

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
SMS public service	352,705.28**	-0.03	1,137,873.42**	863,833.22**
	(170,198)	(0)	(485,808)	(433,058)
SMS deterrence	55,755.24	0.02	536,636.41**	308,588.94
	(123,100)	(0)	(247,232)	(251,841)
SMS reminder	248,156.06*	-0.03	645,553.42*	687,512.79**
	(138,670)	(0)	(371,385)	(328,372)
Observations	1919	1919	1097	1268

Notes: Standard errors are in parentheses (robust for columns 3 and 4). *** p<0.01, ** p<0.05, * p<0.1, using a T distribution with G–L degrees of freedom. Columns 1 and 2 report marginal effects evaluated at the mean. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit.

---

[50] We do not include a dummy variable for size in the PIT regressions, as we did for the CIT sample, because less than 2 per cent of PIT payers are registered at the top-medium or large taxpayers' office. Due to the smaller sample, and hence lower number of districts, we do not use cluster standard errors in the PIT estimations, but White-heteroscedasticity corrected standard errors. In this case, using cluster standard errors could lead to considerably downward biased standard errors (Angrist and Pischke 2009; Cameron and Miller 2015).

[51] The additional revenue is more than RWF 639 million if we use the OLS estimation displayed in column 3 of Table 4 as our baseline (52.9 per cent increase in compliance among those who declared a positive tax due).

Table 5 LATE results for PIT sample

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
SMS public service	948,025.92**	-0.07	1,197,592.95**	903,796.42**
	(418,026)	(0)	(509,550)	(452,289)
SMS deterrence	153,283.21	0.05	562,570.33**	324,749.26
	(425,856)	(0)	(257,925)	(264,302)
SMS reminder	658,577.49	-0.07	659,427.10*	709,317.58**
	(413,522)	(0)	(377,755)	(337,715)
Observations	1919	1919	1097	1268

Notes: Standard errors are in parentheses (robust in column 3 and 4). *** p<0.01, ** p<0.05, * p<0.1. Columns 1 and 3 estimate tobit models. Column 1 and 2 do not report marginal effects. The dependent variable is tax due for fiscal year 2015 in all columns except column 2, where it is a dummy for zero-tax filers. All regressions include a constant, control variables for lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit.

2.3 Discussion of results

Taken together, the results presented in the previous sections allow us to draw some conclusions on the hypotheses we set out to explore in Section 1.2.

As far as message contents are concerned, our results partially confirm hypothesis HP1 (Section 1.2) that fiscal exchange is more effective than deterrence. Put more precisely, the broad result is that kind approaches to compliance (including both fiscal exchange and gentle reminders) can be highly effective compared to deterrence, which seems to work mainly on small taxpayers. Our results show that the control messages were particularly effective in increasing tax revenues and they generally increased tax due by a higher amount than the deterrence messages. Taxpayers may have perceived the control message as a polite, neutral, and gentle reminder of the filing period, which some may prefer over messages emphasising threats and even the emphasis on public services. This finding is largely driven by the relative success of the control message in the letter and the email treatments, and particularly amongst large taxpayers. For emails in particular, it is worth noting that the clear and straightforward subject line of the control message ('Tax filing period open until 31st March 2016') could have been more effective than the more general one of the public service email ('Pay taxes. Build Rwanda. Be proud.'). However, for taxpayers with low and medium turnover the deterrence message was the most effective one and led to larger revenue gains than the reminder. This result is partly confirmed in the PIT sample, where the deterrence SMS also displayed a significant coefficient.

Emails and letters highlighting public services seemed largely ineffective, compared to the public service SMS. It may be that our message had mixed effects in two different directions, which would contribute to the lack of significance: 1) positive for those who appreciate public services, and 2) negative for those who do not feel connected to nation building because of their political views, or who are not satisfied with the quality of public services. However, amongst the SMS treatments, the public service message was the most successful one. In the case of SMS, the public service message has a significantly higher coefficient than other treatments[52] – indicating that the effect is due to the content, rather than the pure fact of receiving an SMS from the RRA. In addition to being more concise than emails and letters, the public service SMS did not contain a graph showing how the government spends its tax revenues – which could provide a possible explanation for the discrepancy in results across

---

[52] This refers to the LATE estimates for CIT (Table 3). For ITT, the effect of the public service message is significantly higher than the control messages (*p*-value below 0.01), but (just) not higher than the deterrence message (*p*-value just above 0.10 except for in column 3 of Table 2).

delivery methods. While we expected this graph to have a positive impact on compliance, it could have also had an adverse effect if the displayed spending did not meet certain taxpayers' expectations, preferences or needs.

As far as delivery methods are concerned, our results suggest that emails and SMS can be highly effective to increase compliance in countries where the success of letter delivery cannot be ensured for many taxpayers. While the effects of pooled treatments by method in Table 2 show that email notifications led to higher tax than the letter treatments, the difference across methods is not statistically significant in the full sample. The sub-group analysis however suggested that emails and letters are significantly more effective than SMS for smaller taxpayers. Still, the SMS messages also brought about a substantial increase in declared taxes, especially driven by the public service SMS. Therefore, these less traditional delivery methods represent a cost-effective and efficient way to reach out to a large number of taxpayers. This result, which does not find evidence for hypothesis HP2 in Section 1.2, could be due to two factors. First, letters may actually have less effect than the theory might suggest, perhaps because of the weaker credibility of deterrence in low-enforcement environments. Second, the apparent weaker effect could be the result of a lower compliance with the letter treatments. Since we have information on the delivery status, we can disentangle these two effects thanks to the LATE estimation. As shown in Table 3, the coefficients on the email treatments are still larger than the corresponding one for letters. While the letters seem to increase declared taxes by a higher amount than the SMS, only the latter effect is statistically significant when using the full sample. In other words, the ranking of methods regarding the increase in the level of tax due is the same as under ITT: emails, letters, SMS (in descending order of effectiveness). Therefore, our results do not seem to support the validity of hypothesis HP2 for the case of Rwanda – and possibly low-income countries more generally. However, once lower compliance with the letters is taken into account, their effectiveness increases substantially (compare column 2 in Table 2 to column 2 in Table 3).

2.4 Robustness

We test the robustness of our results to three factors in particular. The first one is outliers, which is particularly relevant in low-income countries like Rwanda because tax revenues are highly concentrated in a small group of taxpayers. Table 11 and Table 12 in Appendix 6 report respectively the ITT and LATE results for CIT, with the top 1 per cent of observations in terms of tax payable (our dependent variable) capped at the 99th percentile. The results are largely in line with the main analysis and consistent with our conclusions. The treatment effects are naturally much smaller, since we are capping the largest tax payments to smaller amounts. Perhaps the most notable difference is that some of the deterrence messages, particularly emails and SMS, become significant in some regressions. However, they are not statistically different from the reminder. This result is consistent with the slightly larger effect of deterrence for smaller taxpayers in Rwanda that we showed in the sub-group analysis in Section 2.1.1. The corresponding results for the PIT sample are reported in Table 13 and are also largely consistent with the previous results.

Secondly, as mentioned in Section 1.6, we did not include the lagged tax liability in our regressions to avoid collinearity with lagged gross profit. However, to test the robustness of our analysis, we add a control variable for the lagged tax liability. Table 14 and Table 15 report respectively the ITT and LATE results for CIT, while Table 16 reports the corresponding results for PIT. Our results remain qualitatively similar, but some coefficients lose their significance. This is largely expected, as the inclusion of this variable could cause higher standard errors and lead to imprecise estimations of coefficients, potentially leading to type II errors (i.e. failure to find a treatment effect if there is an effect). In particular, the effect of the control letter is no longer significant in this specification. However, the control letter remains highly significant in the LATE analysis (see Table 15), even if we control for the

lagged tax due. Similarly, the PIT liability of taxpayers who received the control SMS is not significantly different from the control group if we include the lagged PIT liability (see Table 16). However, the qualitative implications of the previous section remain valid, particularly as regards the public service SMS that remains the most effective way to approach personal income taxpayers.

In our analysis we investigated taxpayers who declared by 1 April 2016, the deadline for declaring CIT and PIT for the financial year 2015. As only a small number of taxpayers filed income taxes between 2 April and mid-April (54 and 11 CIT and PIT filers, respectively), we do not expect that the inclusion of this additional period has influenced our findings. However, declarations in April may follow a different pattern than declarations which happen before the declaration deadline, for example as a result of outliers or due to different timing of filings of taxpayers in the treatment and control groups.[53] To test the robustness of our results to this potential issue, we include CIT and PIT declarations that were filed after our experimental intervention and by 15 April 2016. Comparing the coefficients using the alternative time window to the results in our main analysis, our results appear robust to the inclusion of late filers.[54]

Conclusions

This paper reports the results of a large-scale field experiment on tax compliance in Rwanda, representing the first wave of such work ever carried out in the African continent. By doing this, it contributes to the literature in two main ways. First, it expands the geographical scope of existing evidence that so far has been almost exclusively composed of studies from high and middle-income countries. Second, it introduces a richer set of treatments than previous studies, combining message contents and delivery methods. The results provide a complex and nuanced picture of tax compliance in Rwanda, differentiating between small and large businesses, corporations and individual firms, as well as taxpayers of 'zero-tax' status. Our results confirm some of the key findings of the literature, while challenging others in a low-income context. These results can be summarised in two main conclusions.

The first broad conclusion regards drivers of compliance and what revenue authorities can do to make taxpayers declare their incomes in full. We find that messages which simply remind taxpayers of deadlines are highly effective in increasing compliance. This is in line with other studies that find that the fact of receiving a letter is often responsible for a large portion of observed treatment effects (Del Carpio 2013; Ortega and Scartascini 2016). Fiscal exchange also seems to be a potentially important driver of compliance, although it is only significant when we use SMS. At the same time, we challenge some of the existing literature by showing that deterrence does not seem to be an effective way to achieve large increases in compliance in Rwanda. This result is likely due to the credibility of deterrence 'threats' in a low-enforcement environment, rather than to the lack of effectiveness of actual deterrence measures. Although small taxpayers are responsive to this type of message, their contribution to the public purse is relatively small, and so is the increased compliance deriving from their response to the messages. Therefore, a friendly approach to taxpayers, including both information about how tax is spent and gentle reminders of deadlines, seems generally more effective than deterrence.

---

[53] Among those taxpayers who filed CIT for 2015 after our experiment intervention, taxpayers who received the control messages filed nearly one day earlier than firms in the control group. A possible reason is the subject line of the control message, which is 'Tax filing period open until 31st March 2016' which explicitly emphasises the filing period. In contrast, there are no significant differences between the treatment and control group regarding the filing date in the PIT experiment.

[54] Results are available from the authors.

The second broad conclusion is related to delivery methods. Our results show that non-traditional ways to communicate to taxpayers, such as SMS and emails, can be highly effective and lead to substantial increases in declared taxes. In low-income contexts, this finding is particularly important as governments are constantly looking for cost-effective ways to achieve policy goals, in the context of very limited resources. Contrary to the theoretical predictions, we find that emails lead to greater revenue gains, followed by letters and, finally, SMS. The fact that letters are delivered by tax officials does not seem to have generated a higher response due to an increase in the probability of detection, in contrast to related results in the literature (Ortega and Scartascini 2016; Doerrenberg and Schmitz 2015).

Our study has clear policy implications for the Rwanda Revenue Authority and for other revenue administrations in low-income countries. First, our results confirm the validity of the RRA's customer-oriented approach, where enforcement is accompanied by several measures to educate and sensitise taxpayers, as well as to facilitate the taxpaying process. Kind approaches seem more effective than 'threats' related to deterrence in Rwanda and potentially in low-income countries more generally, where enforcement is severely limited by the lack of financial and human resources. Overall, our intervention, which involved a relatively cheap communication strategy, generated almost US$9 million as extra tax revenue for the RRA, proving highly cost effective. Second, this study highlights the importance of incorporating rigorous evaluation in the pilot and design of new policies, from the early stages of implementation. By doing this, governments can focus on the measures that work most, without wasting resources on those that are instead of dubious effectiveness.

Finally, this paper will hopefully encourage more researchers to engage in the field of tax experiments in low-income countries and, more broadly, in the rigorous evaluation of tax policies using administrative data. This is a brand new field of research in many countries, as such data had been largely unavailable for research purposes until recently. However, this study adds to a growing body of evidence showing that studies like this are feasible in low-income countries – in addition to being much needed to generate a body of evidence that is globally representative.

Appendices

Appendix 1 Treatment messages

<u>Letters and emails</u>

Kigali, on 8 January 2016

RE: Tax filing period open until 31st March 2016

To XXXXX,

RRA would like to inform you that your CIT tax return is due by 31st March 2016. For more information about the filing process and payment methods, contact the call centre (3004) or visit the RRA website (http://www.rra.gov.rw).

We would like to thank you for your collaboration and wish you a prosperous year 2016.

IMPAMVU: Itariki ntarengwa yo kumenyekanisha umusoro kunyungu ni 31 Werurwe 2016

Kuri XXXXX,

Ikigo cy' Imisoro n' Amahoro kirifuza kubamenyesha ko kumenyekanisha umisoro ku nyungu bikorwa bitarenze itariki ya 31 Werurwe 2016. Ku bisobanuro birambuye byerekeye uburyo bwo kumenyekanisha umisoro ku nyungu bikorwa n'uburyo wishyurwa, mwahamagara umurongo utishyurwa 3004 cyangwa mugasura urubuga rw' Ikigo cy' Imisoro n' Amahoro (RRA) (www.rra.gov.rw).

Tubashimiye ubufatanye mudahwema kutugaragariza, tuboneyeho no kubifuriza umwaka mushya muhire wa 2016, uzababere uw'uburumbuke.

Kind Regards/Murakoze

BIZIMA RUGANINTWALI Pascal
The Deputy Commissioner General & Commissioner for Corporate Services
Komiseri mukuru wungirije akaba na Komiseri w'imirimo

Kigali, on 8 January 2016

RE: Pay your taxes on time and avoid fines and penalties

To XXXXX,

RRA would like to inform you that your CIT tax return is due by 31st March 2016. For more information about the filing process and payment methods, contact the call centre (3004) or visit the RRA website (http://www.rra.gov.rw).

Do you know that if you do not declare and pay your taxes on time, RRA can fine and possibly prosecute you?

Pay your taxes on time and avoid fines and penalties.

We would like to thank you for your collaboration and wish you a prosperous year 2016.

IMPAMVU: Ishyura imisoro ku gihe, wirinde ibihano by'ubukererwe

Kuri XXXXX,

Ikigo cy' Imisoro n' Amahoro kirifuza kubamenyesha ko kumenyekanisha umisoro ku nyungu bikorwa bitarenze itariki ya 31 Werurwe 2016. Ku bisobanuro birambuye byerekeye uburyo bwo kumenyekanisha umisoro ku nyungu bikorwa n'uburyo wishyurwa, mwahamagara umurongo utishyurwa 3004 cyangwa mugasura urubuga rw' Ikigo cy' Imisoro n' Amahoro (RRA) (www.rra.gov.rw).

Wari uzi ko iyo utamenyekanishije ngo unishyure imisoro yawe ku gihe, RRA iguca ibihano by'ubukererwe ikanagukurikirana mu nkiko?

Ishyura imisoro yawe ku gihe, wirinde ibihano by'ubukererwe.

Tubashimiye ubufatanye mudahwema kutugaragariza, tuboneyeho no kubifuriza umwaka mushya muhire wa 2016, uzababere uw'uburumbuke.

Kind Regards/Murakoze

BIZIMA RUGANINTWALI Pascal
The Deputy Commissioner General & Commissioner for Corporate Services
Komiseri mukuru wungirije akaba na Komiseri w'imirimo rusange

Spending of tax of RWF 100
Uko Leta ikoresha amafaranga 100 y'umusoro

- Education/Uburezi 15
- Other public services/Ibindi bikorwa bya Leta 40
- Health/Ubuzima 11
- Economy/Ubukungu 19
- Defense, safety/Umutekano 15
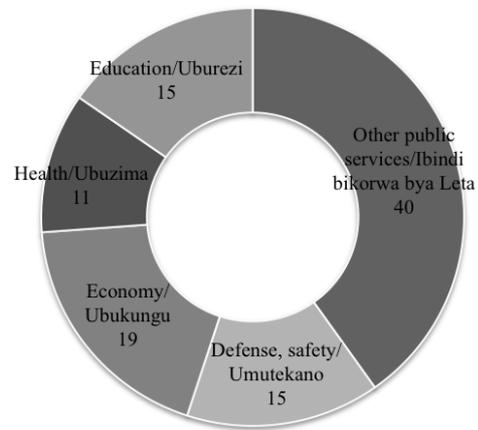
Kigali, on 8 January 2016

RE: Pay taxes. Build Rwanda. Be proud

To XXXXX,

RRA would like to inform you that your CIT tax return is due by 31st March 2016. For more information about the filing process and payment methods, contact the call centre (3004) or visit the RRA website (http://www.rra.gov.rw).

By paying your taxes you make it possible to educate our children, fund our healthcare, and keep us safe.

Pay taxes. Build Rwanda. Be proud.

We would like to thank you for your collaboration and wish you a prosperous year 2016.

IMPAMVU: Ishyura imisoro. Ubaka u Rwanda. Gira ishema

Kuri XXXXX,

Ikigo cy' Imisoro n' Amahoro kirifuza kubamenyesha ko kumenyekanisha umisoro ku nyungu bikorwa bitarenze itariki ya 31 Werurwe 2016. Ku bisobanuro birambuye byerekeye uburyo bwo kumenyekanisha umisoro ku nyungu bikorwa n'uburyo wishyurwa, mwahamagara umurongo utishyurwa 3004 cyangwa mugasura urubuga rw' Ikigo cy' Imisoro n' Amahoro (RRA) (http://www.rra.gov.rw)

Iyo wishyuye imisoro yawe, ushoboza abana kwiga, uba ushyigikiye ubuvuzi n'umutekano byacu.

Ishyura imisoro. Ubaka u Rwanda. Gira Ishema.

Tubashimiye ubufatanye mudahwema kutugaragariza, tuboneyeho no kubifuriza umwaka mushya muhire wa 2016, uzababere uw'uburumbuke.

Kind Regards/Murakoze

BIZIMA RUGANINTWALI Pascal
The Deputy Commissioner General & Commissioner for Corporate Services
Komiseri mukuru wungirije akaba na Komiseri w'imirimo rusange

SMS

Deterrence:

To ....

If you do not declare and pay your taxes on time, the RRA will fine and possibly prosecute you. Pay your taxes on time and avoid sanctions.

Kuri ....

Nutamenyekanisha ngo unishyure imisoro ku gihe, RRA izaguca amande, kandi yanagukurikirana munkiko! Ishyura imisoro kugihe wirinde ibihano.

. . . .

Public service:

To ....

By paying your taxes you make it possible to educate our children, fund our healthcare, and keep us safe. Pay taxes. Build Rwanda. Be proud.

Kuri ....

Iyo wishyuye imisoro, uba ushyigikiye uburezi bw'abana, ubuvuzi n'umutekano byacu. Ishyura imisoro. Ubuka u Rwanda. Gira ishema.

. . . .

Control:

To ....

Your CIT tax return is due by 31/3/16. For more information, contact the call centre (3004) or visit www.rra.gov.rw.

Kuri ....

Itariki ntarengwa yo kumenyekanisha umusoro k'unyungu ni 31/3/2016. Kubisobanuro birambuye hamagara 3004 cyangwa usure www.rra.gov.rw.

Appendix 2 Balance checks

Table 6 Balance between CIT control and treatment groups

Variable	(1) Real regime	(2) Zero tax due lagged	(3) Large and top-medium	(4) Gross profit lagged (in RWF)	(5) Tax due lagged (in RWF)
Public service SMS, *n*=1059	582 (0.80)	612 (0.90)	50 (0.35)	100,513,876 (0.40)	3,116,447 (0.32)
Deterrence SMS, *n*=1060	584 (0.77)	611 (0.83)	60 (1.00)	82,328,246 (0.83)	1,113,613 (0.51)
Control SMS, *n*=1060	579 (0.93)	614 (0.93)	52 (0.46)	118,786,246 (0.24)	5,869,047 (0.61)
Public service letter, *n*=1061	575 (0.93)	611 (0.80)	64 (0.73)	156,547,882 (0.08)	13,753,107 (0.53)
Deterrence letter, *n*=1060	580 (0.90)	619 (0.90)	55 (0.65)	64,957,390 (0.56)	1,190,928 (0.88)
Control letter, *n*=1060	584 (0.77)	618 (0.93)	50 (0.35)	46,326,911 (0.42)	1,854,897 (0.71)
Public service email, *n*=1060	585 (0.74)	614 (0.93)	56 (0.72)	68,401,118 (0.81)	3,947,415 (0.62)
Deterrence email, *n*=1060	577 (1.00)	613 (0.90)	49 (0.26)	58,574,565 (0.47)	2775117 (0.57)
Control email, *n*=1060	578 (0.97)	610 (0.80)	60 (1.00)	92,270,919 (0.89)	4,275,325 (0.41)
Control group, *n*=1270	691	738	72	78,213,997	2,941,650

Notes: *n* indicates the number of taxpayers in a given treatment. *p*-values comparing the balance between the treatment and the control group are shown in parentheses. For the categorical variables (columns 1 to 3), the table displays the *p*-value of the two-sided Fisher's exact test. The null hypothesis is that the classification of taxpayers (e.g. being in the real regime) is not associated with assignment into the treatment or control group. For the continuous variables (columns 4 to 5), we show the *p*-value of the Mann-Whitney test. The null hypothesis is that the distribution of the continuous variable is the same in the treatment and control groups. The variable 'large and top-medium' refers to the RRA segmentation of taxpayers into different tax offices (SMTO and LTO) based on taxpayer size.

Table 7 Balance between PIT control and treatment groups

Variable	(1) Real regime	(2) Zero tax due lagged	(3) Large and top-medium	(4) Gross profit lagged (in RWF)	(5) Tax due lagged (in RWF)
Public service SMS, *n*=697	153 (0.90)	474 (0.91)	11 (0.69)	14,214,255 (0.79)	588,316 (0.71)
Deterrence SMS, *n*=698	151 (1.00)	221 (0.82)	14 (1.00)	14,292,145 (0.84)	449,512 (0.43)
Control SMS, *n*=699	151 (1.00)	212 (0.45)	16 (0.85)	22,685,007 (0.40)	813,877 (0.14)
Control group, *n*=699	151	226	14	18,565,437	497,053

Notes: *n* indicates the number of taxpayers in a given treatment. *p*-values comparing the balance between the treatment and the control group are shown in parentheses. For the categorical variables (columns 1 to 3), the table displays the *p*-value of the two-sided Fisher's exact test. The null hypothesis is that the classification of taxpayers (e.g. being in the real regime) is not associated with assignment into the treatment or control group. For the continuous variables (columns 4 to 5), we show the *p*-value of the Mann-Whitney test. The null hypothesis is that the distribution of the continuous variable is the same in the treatment and control groups. The variable 'large and top-medium' refers to the RRA segmentation of taxpayers into different tax offices (SMTO and LTO) based on taxpayer size.

Appendix 3 Variables

Table 8 Summary of variables used in the regression analysis

Variable	Description
Tax due *(in RWF)*	Total tax liability for CIT and PIT for fiscal year 2015
Treatment *(binary)*	Binary variables indicating whether the taxpayer was in a given treatment group or not
Large *(binary)*	Binary variable that takes the value 1 if the taxpayer is registered at RRA as a large or top-medium firm, and zero otherwise
Nyarugenge *(binary)*	Binary variable reflecting location of the taxpayer in the district of Nyarugenge
Kicukiro *(binary)*	Binary variable reflecting location of the taxpayer in the district of Kicukiro
Zero tax due *(binary)*	Binary variable that takes the value 1 for firms who did not declare a positive tax liability in their previous tax return
Gross profit (in RWF)	Variable that reflects the lagged gross profits (i.e. business income minus the cost of goods and services) of firms in the real regime. The variable takes the value zero for firms who are in the lump-sum and flat-amount regimes

Appendix 4 Power calculations

Figure 1 Power calculations for CIT

Estimated experimental-group mean for a two-sample means test

t test assuming $\sigma_1 = \sigma_2 = \sigma$

$H_0: \mu_2 = \mu_1$ versus $H_a: \mu_2 > \mu_1$

Parameters: $\alpha = .05$, N = 2330, $N_1 = 1270$, $N_2 = 1060$, $\mu_1 = 2.9$, $\sigma = 36$
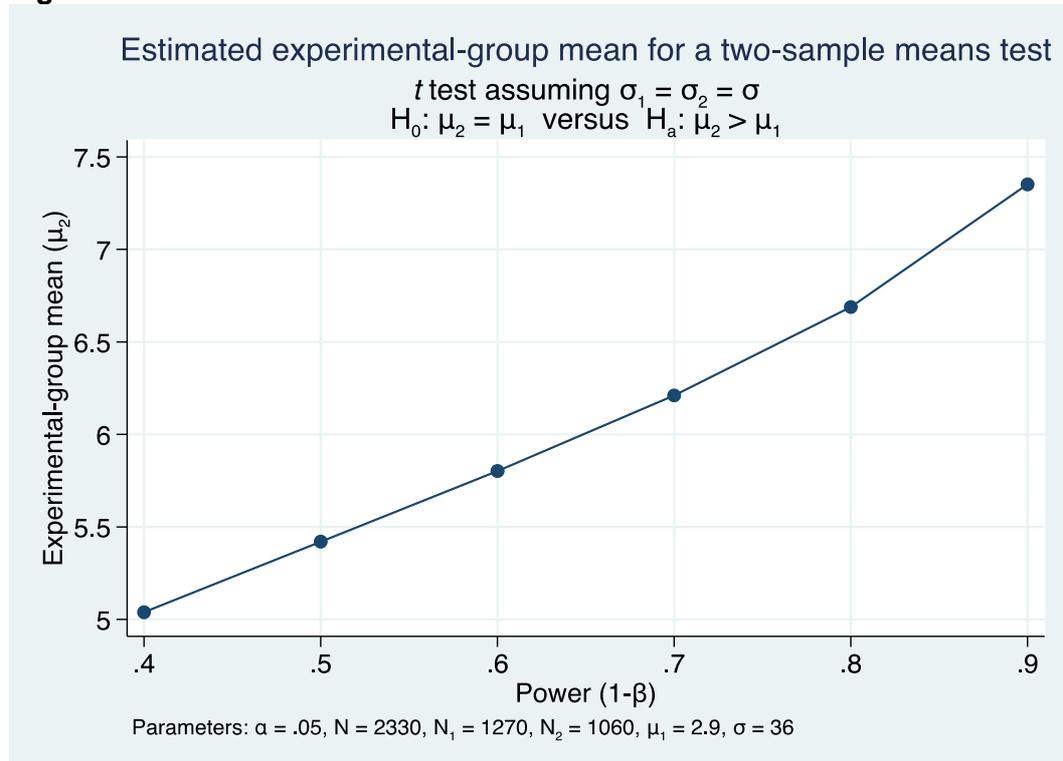
Note: The mean value used in the power calculations is the one reported in Table 6 for the control group. The mean values on the vertical axis are in millions of RWF.

Figure 2 Power calculations for PIT

Estimated experimental-group mean for a two-sample means test

t test assuming $\sigma_1 = \sigma_2 = \sigma$

$H_0: \mu_2 = \mu_1$ versus $H_a: \mu_2 > \mu_1$

Parameters: $\alpha = .05$, N = 1398, $N_1 = 699$, $N_2 = 699$, $\mu_1 = .5$, $\sigma = 2.9$
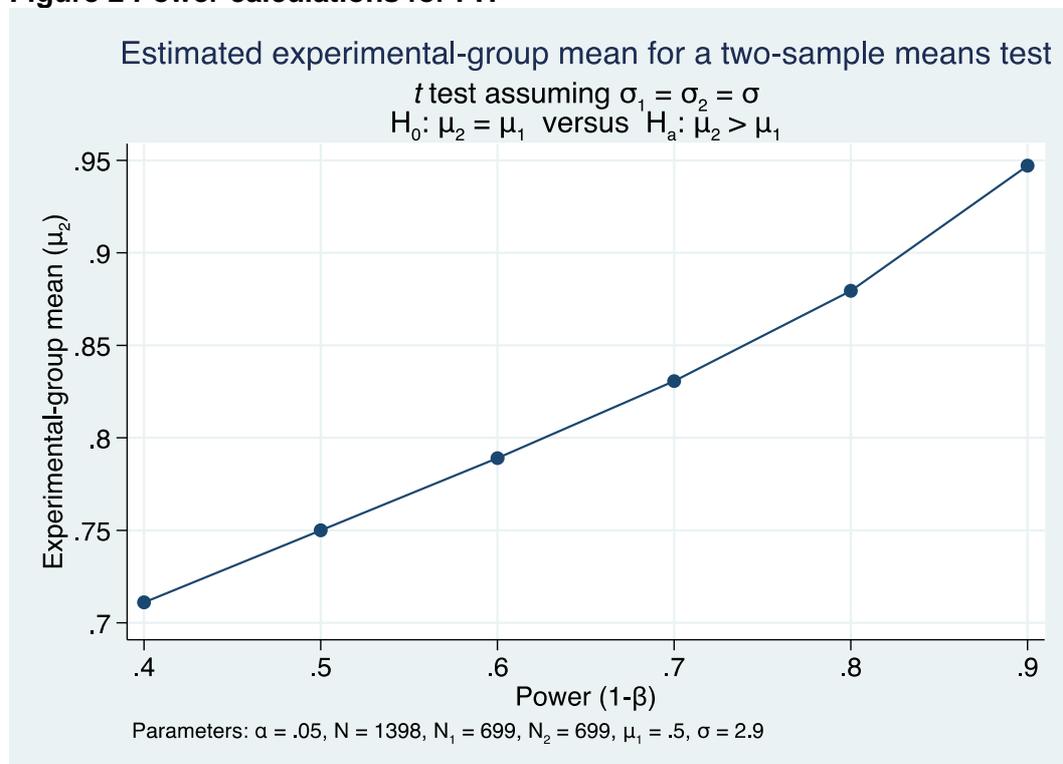
Note: The mean value used in the power calculations is the one reported in Table 7 for the control group. The mean values on the vertical axis are in millions of RWF.

Appendix 5 Results for size sub-groups

Table 9 Results on pooled treatments for largest 10%

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
		ITT		
Public service	3,749,856.48	0.01	12,578,287.45	11,410,608.19
	(2,930,759)	(0)	(7,739,032)	(8,512,742)
Deterrence	2,944,920.02*	-0.03	5,170,221.06	7,087,552.47
	(1,485,843)	(0)	(3,472,578)	(4,303,277)
Reminder	7,547,035.33**	0.01	26,220,209.94**	24,261,006.39**
	(3,033,921)	(0)	(10,744,791)	(10,095,785)
SMS	2,073,840.87	0.01	6,604,484.69	10,719,783.03
	(1,482,930)	(0)	(5,384,828)	(6,152,219)
Letter	4,499,963.34	-0.00	14,731,055.07	15,032,905.95**
	(3,409,729)	(0)	(8,687,894)	(6,912,178)
Email	7,722,691.94*	-0.01	22,151,693.86*	17,162,693.93
	(4,009,348)	(0)	(12,373,178)	(13,156,189)
		LATE		
Public service	9,653,244.56	0.03	14,620,783.06	13,443,256.04
	(6,137,954)	(0)	(8,238,727)	(10,010,063)
Deterrence	8,048,031.68**	-0.09***	5,943,613.05*	8,335,491.97**
	(3,390,568)	(0)	(3,204,003)	(3,513,177)
Reminder	22,467,762.63***	0.04	32,299,017.73**	29,779,091.59**
	(5,880,489)	(0)	(13,150,904)	(10,714,815)
SMS	4,931,374.78**	0.02	6,685,501.76***	10,952,616.67***
	(2,026,196)	(0)	(1,887,814)	(3,314,231)
Letter	16,671,489.49***	0.00	22,675,553.20**	23,871,587.45***
	(4,651,086)	(0)	(9,093,498)	(8,201,630)
Email	20,157,511.72**	-0.05	24,440,671.27	19,036,376.94
	(8,677,631)	(0)	(14,088,796)	(13,136,948)
Observations	1446	1446	913	949

Notes: Cluster robust standard errors are in parentheses (a jackknife variance estimation is used in the LATE regressions). *** $p<0.01$, ** $p<0.05$, * $p<0.1$, using a T distribution with $G–L$ degrees of freedom. All LATE rows report estimates from instrumental variable (IV) models. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. Marginal effects are reported for ITT in columns 1 and 2 but not for LATE. All regressions include a constant, controls for size, lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit. Reported coefficients for pooled treatments by methods (email, SMS, letter) and contents (public service, deterrence, reminder) result from two separate regressions where the disaggregated treatments are not included.

Table 10 Results on pooled treatments for small and medium taxpayers (bottom 90%)

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
		ITT		
Public service	71,379.87	-0.02	203,365.12	172,870.61
	(59,433)	(0)	(177,688)	(120,017)
Deterrence	83,668.79***	-0.02	407,569.38**	357,160.02**
	(20,040)	(0)	(186,688)	(157,132)
Reminder	43,290.04**	-0.03***	52,280.38	131,003.18
	(18,589)	(0)	(97,069)	(97,238)
SMS	54,887.48***	-0.03**	118,409.71	111,039.64*
	(8,301)	(0)	(177,726)	(56,947)
Letter	46,600.14	-0.02	114,368.56*	183,386.51*
	(66,606)	(0)	(56,997)	(99,649)
Email	95,928.08***	-0.02	419,158.44***	356,433.71***
	(30,942)	(0)	(39,111)	(61,113)
		LATE		
Public service	355,797.54	-0.08	239,755.00	207,804.37
	(351,975)	(0)	(193,401)	(131,457)
Deterrence	432,246.03***	-0.06	493,886.59**	429,379.11***
	(142,495)	(0)	(226,570)	(60,142)
Reminder	220,853.18*	-0.09**	62,457.33***	159,859.57
	(116,540)	(0)	(10,577)	(123,960)
SMS	227,928.72***	-0.07**	119,255.32	111,100.10*
	(26,502)	(0)	(79,566)	(54,485)
Letter	327,406.42	-0.09	192,946.12***	319,229.06*
	(596,536)	(0)	(42,099)	(171,969)
Email	450,071.88**	-0.06	466,759.97***	392,770.36***
	(174,580)	(0)	(29,252)	(6,207)
Observations	7235	7235	2868	2676

Notes: Cluster robust standard errors are in parentheses (a jackknife variance estimation is used in the LATE regressions). *** $p<0.01$, ** $p<0.05$, * $p<0.1$, using a T distribution with $G–L$ degrees of freedom. All LATE rows report estimates from instrumental variable (IV) models. Marginal effects are reported for ITT in columns 1 and 2 but not for LATE. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for size, lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit. Reported coefficients for pooled treatments by methods (email, SMS, letter) and contents (public service, deterrence, reminder) result from two separate regressions where the disaggregated treatments are not included.

Appendix 6 Robustness

Table 11 CIT ITT regressions with top 1% tax payments capped at 99th percentile

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
SMS public service	160,479.06**	-0.04***	209,011.08	593,983.27***
	(65,254)	(0)	(192,439)	(130,887)
SMS deterrence	102,415.68	-0.03	162,312.80	530,324.54***
	(69,027)	(0)	(313,835)	(112,564)
SMS reminder	-4,926.06	0.00	50,455.51	592,539.95***
	(70,086)	(0)	(69,997)	(86,994)
Letter public service	583.10	0.00	-7,066.72	185,442.62
	(91,191)	(0)	(252,915)	(230,935)
Letter deterrence	100,296.21	-0.03*	-23,426.30	122,113.67
	(84,545)	(0)	(231,628)	(165,707)
Letter reminder	29,814.40	-0.02	-15,305.87	311,492.77
	(42,690)	(0)	(250,426)	(351,267)
Email public service	9,489.76	-0.01	-188,915.61	-216,599.89
	(42,508)	(0)	(109,746)	(190,034)
Email deterrence	73,302.04	-0.00	307,808.37	632,703.06**
	(108,622)	(0)	(338,443)	(272,746)
Email reminder	211,903.87***	-0.04***	595,784.60***	638,150.67***
	(71,446)	(0)	(146,542)	(175,236)
Constant			1,491,158.30***	831,100.05***
			(50,110)	(81,422)
Observations	9096	9096	4053	4002

Notes: Cluster robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$, using a *T* distribution with *G–L* degrees of freedom. Columns 1 and 2 report marginal effects evaluated at the mean. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for size, lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit. All controls are highly significant but are omitted for the sake of clarity.

Table 12 CIT LATE regressions with top 1% tax payments capped at 99th percentile

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
SMS public service	531,196.24**	-0.10***	211,007.04*	603,306.67***
	(214,149)	(0)	(118,186)	(111,574)
SMS deterrence	341,448.85*	-0.07	165,121.22	539,159.27***
	(179,513)	(0)	(276,189)	(134,704)
SMS reminder	-16,825.70	0.00	51,173.95	596,261.06***
	(52,586)	(0)	(56,812)	(13,505)
Letter public service	-67,241.71	0.01	-10,446.87	290,869.30
	(675,073)	(0)	(179,456)	(215,283)
Letter deterrence	613,713.68	-0.16	-43,779.57	227,989.78
	(639,930)	(0)	(475,337)	(359,836)
Letter reminder	132,828.99	-0.06	-24,632.09	518,581.61
	(78,694)	(0)	(335,584)	(590,814)
Email public service	31,767.20	-0.04	-216,037.86	-251,972.18***
	(142,427)	(0)	(157,840)	(71,505)
Email deterrence	244,140.82	-0.00	311,915.30	638,563.47**
	(412,287)	(0)	(400,299)	(280,140)
Email reminder	829,094.03***	-0.13***	711,548.99***	760,540.60***
	(195,336)	(0)	(122,066)	(206,085)
Constant	363,807.71	-0.80***	1,489,929.63***	826,370.86***
	(356,201)	(0)	(84,468)	(133,005)
Observations	9096	9096	4053	4002

Notes: Cluster robust standard errors are in parentheses, using a jackknife variance. *** $p<0.01$, ** $p<0.05$, * $p<0.1$, using a T distribution with $G–L$ degrees of freedom. Columns 1 and 2 do not report marginal effects evaluated at the mean. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for size, lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit. All controls are highly significant but are omitted for the sake of clarity.

Table 3 PIT regressions with top 1% tax payments capped at 99th percentile

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
		ITT		
SMS public service	225,515.61**	-0.03	662,463.28	551,194.07
	(90,920)	(0)	(437,401)	(305,992)
SMS deterrence	38,597.35	0.02	331,718.04	211,661.40
	(83,569)	(0)	(307,716)	(132,909)
SMS reminder	159,260.50**	-0.03**	384,549.19	447,455.89***
	(76,602)	(0)	(214,722)	(128,426)
		LATE		
SMS public service	580,724.12***	-0.07	697,240.10	576,692.76**
	(170,257)	(0)	(396,331)	(230,611)
SMS deterrence	102,186.79**	0.05	347,747.86*	222,754.36***
	(45,149)	(0)	(166,382)	(35,578)
SMS reminder	405,547.02***	-0.07***	392,817.85***	461,649.28***
	(45,557)	(0)	(18,922)	(16,583)
Observations	1919	1919	1097	1268

Notes: Standard errors are in parentheses (robust for columns 3 and 4). *** $p<0.01$, ** $p<0.05$, * $p<0.1$, using a T distribution with G–L degrees of freedom. Columns 1 and 2 report marginal effects evaluated at the mean. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit.

Table 14 CIT ITT regressions controlling for lagged tax due

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
SMS public service	1,109,972.55***	-0.04***	1,883,454.88**	3,154,896.62***
	(308,886)	(0)	(688,894)	(611,039)
SMS deterrence	364,423.39	-0.03	-310,853.10	653,559.15
	(493,614)	(0)	(1,354,922)	(1,259,831)
SMS reminder	102,482.05	0.00	1,612,168.88	2,935,756.46
	(352,142)	(0)	(2,006,015)	(2,359,002)
Letter public service	147,578.28	0.00	182,937.19	1,244,638.44
	(1,201,282)	(0)	(3,007,842)	(2,616,616)
Letter deterrence	349,388.06	-0.03*	-718,541.98	-447,051.03
	(634,794)	(0)	(1,640,812)	(1,616,064)
Letter reminder	891,986.53	-0.02	4,118,949.74	4,433,945.93
	(620,469)	(0)	(3,949,346)	(3,934,022)
Email public service	-20,147.50	-0.01	569,508.49	-2,646,134.02
	(599,123)	(0)	(1,428,232)	(2,825,074)
Email deterrence	180,842.03	-0.00	1,231,385.78	2,226,024.37
	(535,205)	(0)	(1,560,506)	(1,572,913)
Email reminder	2,279,725.93***	-0.04***	8,488,119.65*	7,442,981.62
	(861,907)	(0)	(4,112,397)	(4,525,796)
Observations	9096	9096	4053	4002

Notes: Cluster robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$, using a T distribution with $G–L$ degrees of freedom. Columns 1 and 2 report marginal effects evaluated at the mean. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for size, lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit. All controls are highly significant but are omitted for the sake of clarity.

Table 15 CIT LATE regressions controlling for lagged tax due

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
SMS public service	4,175,055.77***	-0.10***	1,897,353.24***	3,206,906.05***
	(373,632)	(0)	(549,615)	(631,211)
SMS deterrence	1,391,729.84	-0.07	-313,482.90	659,534.78
	(1,773,232)	(0)	(1,734,292)	(1,206,208)
SMS reminder	375,119.78	0.00	1,620,943.39	2,953,536.76
	(1,947,214)	(0)	(2,899,416)	(3,597,364)
Letter public service	211,997.35	0.01	277,130.19	1,953,757.55
	(8,434,277)	(0)	(3,289,172)	(3,235,483)
Letter deterrence	2,615,636.30	-0.16	-1,339,914.37	-852,600.24
	(5,436,105)	(0)	(3,795,311)	(3,525,648)
Letter reminder	5,873,039.86***	-0.06	6,652,254.61***	7,384,025.27***
	(738,890)	(0)	(1,499,248)	(1,415,282)
Email public service	-131,782.05	-0.04	645,070.76	-3,060,851.68
	(2,904,935)	(0)	(2,295,666)	(3,473,355)
Email deterrence	685,461.26	-0.00	1,247,715.93	2,247,480.56
	(2,319,907)	(0)	(1,951,461)	(1,642,502)
Email reminder	10,328,426.59**	-0.13***	10,137,326.42	8,877,961.70
	(4,157,165)	(0)	(6,969,936)	(6,417,120)
Observations	9096	9096	4053	4002

Notes: Cluster robust standard errors are in parentheses, with a jackknife correction. *** $p<0.01$, ** $p<0.05$, * $p<0.1$, using a T distribution with $G–L$ degrees of freedom. Columns 1 and 2 do not report marginal effects evaluated at the mean. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for size, lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit. All controls are highly significant but are omitted for the sake of clarity.

Table 16 PIT regressions controlling for lagged tax due

	(1) Tobit	(2) Part 1: probit	(3) Part 2: OLS	(4) Baseline tax>0
		ITT		
SMS public service	188,562.73***	-0.03	567,647.02	374,709.53*
	(61,411)	(0)	(371,586)	(194,272)
SMS deterrence	-53,757.08	0.02	79,725.36	-28,907.12
	(66,211)	(0)	(178,763)	(57,826)
SMS reminder	43,964.76	-0.03***	-11,259.94	113,552.98
	(142,426)	(0)	(423,202)	(304,619)
		Late		
SMS public service	485,506.90**	-0.07***	597,476.94	392,137.32
	(187,437)	(0)	(429,041)	(253,899)
SMS deterrence	-140,220.23***	0.06***	83,695.09	-30,472.18
	(33,217)	(0)	(112,958)	(35,951)
SMS reminder	108,167.61	-0.08***	-11,431.35	117,158.04
	(140,225)	(0)	(167,864)	(138,467)
Observations	1919	1919	1097	1268

Notes: Standard errors are in parentheses (robust for columns 3 and 4). *** $p<0.01$, ** $p<0.05$, * $p<0.1$, using a T distribution with $G–L$ degrees of freedom. Columns 1 and 2 report marginal effects evaluated at the mean. The dependent variable is tax due for fiscal year 2015 in columns 1, 3, and 4; it is a binary variable taking the value of 1 if the taxpayer is 'zero-tax' in column 2. All regressions include a constant, controls for lagged gross profit, dummies for the districts Nyarugenge and Kicukiro, as well as a dummy variable indicating lagged zero tax due and an interaction variable between lagged zero tax due and lagged gross profit.

References

Ali, M., Fjeldstad, O-H. and Sjursen, I. H. (2015) 'To Pay or Not to Pay? Citizens' Attitudes towards Taxation in Kenya, Tanzania, Uganda and South Africa', *World Development* 64: 828–42

Allingham, M. G. and Sandmo, A. (1972) 'Income Tax Evasion: A Theoretical Analysis', *Journal of Public Economics* 1 (3–4): 323–38

Alm, J., Cherry, T., Jones, M. and McKee, M. (2010) 'Taxpayer Information Assistance Services and Tax Compliance Behavior', *Journal of Economic Psychology* 31 (4): 577–86

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

Angrist, J. D., and Pischke, J-S. (2009) *Mostly Harmless Econometrics: An Empiricist's Companion,* Princeton, NJ: 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 (2016) *African Tax Outlook*, www.ataftax.org/en/Documents/ATO Africa Tax Outlook_Book_Eng Version_Final_Lowres.pdf (accessed 11 January 2017)

Becker, G. S. (1968) 'Crime and Punishment: An Economic Approach', *Journal of Political Economy* 76: 169

Bell, R. M. and McCaffrey, D. F. (2002) 'Bias Reduction in Standard Errors for Linear Regression with Multi-Stage Samples', *Survey Methodology* 28 (12): 169–81

Besley, T. and Perrson, T. (2013) 'Taxation and Development', in A. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.), Handbook of Public Economics, Elsevier

BIT (2012) *Applying Behavioural Insights to Reduce Fraud, Error and Debt*, London: Cabinet Office Behavioural Insights Team

Bloom, H. S. (2008) 'The Core Analytics of Randomized Experiments for Social Research', in Alasuutari, P., Bickman, L. and Brannen, J. (eds) *The SAGE Handbook of Social Research Methods*, Thousand Oaks, CA: SAGE

Blumenthal, M., Christian, C. and Slemrod, J. (2001) 'Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota,' *National Tax Journal* 54 (1): 125–38

Bott, K., Cappelen, A. W. and Sørensen, E. Ø. (2014) *You've Got Mail: A Randomised Field Experiment on Tax Evasion,* Norwegian School of Economics

Cameron, A. C. and Trivedi, P. K. (2010) *Microeconometrics Using Stata,* revised ed., Stata Press

Cameron, A. C. and Miller, D. L. (2015) 'A Practitioner's Guide to Cluster-Robust Inference', *Journal of Human Resouces* 50 (2): 317–72

Carrillo, P., Castro, E. and Scartascini, C. (2016) *Do Rewards Work to Maintain and Increase Tax Compliance? Evidence from the Randomization of Public Goods*

Carrillo, P., Pomeranz, D. and Singhal, M. (2016) 'Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement', *American Economic Journal: Applied Economics*

Castro, L., and Scartascini, C. (2015) 'Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment', *Journal of Economic Behavior and Organization* 116: 65–82

——— (2014) *The Devil Is in the Details: Policy Design Lessons from Field Experiments in the Pampas*, Policy Brief, IDB-PB-232, Inter-American Development Bank

Coolidge, J. (2012) 'Findings of Tax Compliance Cost Surveys in Developing Countries', *eJournal of Tax Research* 10 (2): 250–87

Coricelli, G., Joffily, M., Montmarquette, C. and Villeval, M-C. (2007) *Tax Evasion : Cheating Rationally or Deciding Emotionally?* IZA Discussion Paper 3103, Bonn: IZA

Curtis, D. E. A. (2015) 'Development Assistance and the Lasting Legacies of Rebellion in Burundi and Rwanda', *Third World Quarterly* 36 (7): 1365–81

Del Carpio, L. (2014) *Are The Neighbors Cheating? Evidence From A Social Norm Experiment On Property Taxes In Peru*

Doerrenberg, P. and Schmitz, J. (2015) *Tax Compliance and Information Provision: A Field Experiment with Small Firms,* IZA Discussion Paper 9013, Bonn: IZA

Drago, F., Mengel, F. and Traxler, C. (2015) *Compliance Behavior in Networks: Evidence from a Field Experiment,* IZA Discussion Paper 9443, Bonn: IZA

Dwenger, N., Kleven, H., Rasul, I. and Rincke, J. (2015) 'Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany', *American Economic Journal: Economic Policy*

Erard, B., and Feinstein, J. S. (1994) 'Honesty and Evasion in the Tax Compliance Game', *RAND Journal of Economics* 25 (1): 1–19

Fellner, G., Sausgruber, R. and Traxler, C. (2013) 'Testing Enforcement Strategies in the Field: Legal Threat, Moral Appeal and Social Information', *Journal of the European Economic Association* 11 (3): 634–660

Fjeldstad, O-H. and Semboja, J. (2001) 'Why People Pay Taxes: The Case of the Development Levy in Tanzania', *World Development* 29 (2059–2074)

Glennerster, R. and Takavarasha, K. (2013) *Running Randomized Evaluations: A Practical Guide,* Princeton, NJ: Princeton University Press

Hallsworth, M. (2014) 'The Use of Field Experiments to Increase Tax Compliance', *Oxford Review of Economic Policy* 30 (4): 658–79

Hallsworth, M., List, J. A., Metcalfe, R. D. and Vlaev, I. (2014) *The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance*

Hashimzade, N., Myles, G. D. and Tran-Nam, B. (2013) 'Applications of Behavioural

Economics to Tax Evasion', *Journal of Economic Surveys* 27 (5): 941–77

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

Luttmer, E. and Singhal, M. (2014) 'Tax Morale', *Journal of Economic Perspectives* 28 (4): 149–68

Mascagni, G. (2016) *From the Lab to the Field : A Review of Tax Experiments,* ICTD Working Paper 46, Brighton: Institute of Development Studies

Mascagni, G. and Mengistu, A. (2016) *The Corporate Tax Burden in Ethiopia : Evidence from Anonymised Tax Returns,* ICTD Working Paper 48, Brighton: Institute of Development Studies

Mascagni, G., Monkam, N. and Nell, C. (2016a) *Unlocking the Potential of Administrative Data in Africa : Tax Compliance and Progressivity in Rwanda,* ICTD Working Paper 56, Brighton: Institute of Development Studies

Mascagni, G., Nell, C., Monkam, N. and Mukama, D. (2016b) *The Carrot and the Stick: Evidence on Voluntary Tax Compliance from a Pilot Field Experiment in Rwanda,* ICTD Working Paper 56, Brighton: Institute of Development Studies

Mukama, D., Karangwa, J. and Hakizimana, N. (2016) *Communicating to Improve Compliance: Taxpayers' Feedback on Message and Mode of Delivery in Rwanda*

Myles, G. D. and Naylor, R. A. (1996) 'A Model of Tax Evasion with Group Conformity and Social Customs', *European Journal of Political Economy* 12 (1): 49–66

Ortega, D. and Scartascini, C. (2016) *Don't Blame the Messenger: A Field Experiment on Delivery Methods for Increasing Tax Compliance,* IDB Working Paper No. IDB-WP-627, Inter-American Development Bank

Schneider, F. and Williams, C. C. (2013) *The Shadow Economy,* London: The Institute of Economic Affairs

Slemrod, J., Blumenthal, M. and Christian, C. (2001) 'Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota', *Journal of Public Economics* 79 (3): 455–83

Slemrod, J. and Weber, C. (2012) 'Evidence of the Invisible: Toward a Credibility Revolution in the Empirical Analysis of Tax Evasion and the Informal Economy', *International Tax and Public Finance* 19 (1): 25–53

The Economist (2016) 'Health Care in Rwanda. An African Trailblazer', 17 September, http://www.economist.com/news/middle-east-and-africa/21707226-how-poor-country-brought-health-insurance-91-population-african (accessed 11 January 2017)

Traxler, C. (2010) 'Social Norms and Conditional Cooperative Taxpayers', *European Journal of Political Economy* 26 (1): 89–103

Yesegat, W., Vorontsov, D. E., Coolidge, J. and Corthay, L. O. (2015) *Tax Compliance Cost Burden and Tax Perceptions Survey in Ethiopia,* Washington, DC: World Bank Group