



Working Paper 46

# From the Lab to the Field: a Review of Tax Experiments

Giulia Mascagni

February 2016



ICTD Working Paper 46

# **From the Lab to the Field: a Review of Tax Experiments**

Giulia Mascagni

February 2016

From the Lab to the Field: a Review of Tax Experiments  
Giulia Mascagni  
ICTD Working Paper 46  
First published by the Institute of Development Studies in February 2016  
© Institute of Development Studies 2016  
ISBN: 978-1-78118-283-3

A catalogue record for this publication is available from the British Library.  
All rights reserved. Reproduction, copy, transmission, or translation of any part of this publication may be made only under the following conditions:  
- with the prior permission of the publisher; or  
- with a licence from the Copyright Licensing Agency Ltd., 90 Tottenham Court Road, London W1P 9HE, UK, or from another national licensing agency; or  
- under the terms set out below.

This publication is copyright, but may be reproduced by any method without fee for teaching or nonprofit purposes, but not for resale. Formal permission is required for all such uses, but normally will be granted immediately. For copying in any other circumstances, or for reuse in other publications, or for translation or adaptation, prior written permission must be obtained from the publisher and a fee may be payable.

Available from:  
The International Centre for Tax and Development  
at the Institute of Development Studies,  
Brighton BN1 9RE, UK  
Tel: +44 (0) 1273 606261 Fax: +44 (0) 1273 621202  
E-mail: [info@ictd.ac.uk](mailto:info@ictd.ac.uk)  
Web: [www.ictd/en/publications](http://www.ictd/en/publications)

IDS is a charitable company limited by guarantee and registered in England (No. 877338)

# **From the Lab to the Field: a Review of Tax Experiments**

Giulia Mascagni

## **Summary**

Tax experiments have been gaining momentum in recent years, although this literature dates back several decades. With new developments in methods and data availability, tax experiments have gradually moved away from lab settings and towards the field. This movement from the lab to the field has happened against the background of the 'credibility revolution' in applied economics, which has seen more rigorous methods applied to policy relevant questions, and of the availability to researchers of administrative data from tax returns. These developments have allowed significant advances in the experimental literature on tax compliance. This paper reviews this literature, giving particular attention to field experiments using administrative data, but putting them in the broader context of the compliance literature. A particular effort is made to take a global perspective, in a literature that is only recently seeing the emergence of evidence from Africa, Latin America and Asia.

**Keywords:** tax experiments; administrative data; tax compliance; randomisation.

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

# Contents

	Summary	3
	Acknowledgements	5
	Acronyms	5
<b>1</b>	<b>Introduction</b>	<b>6</b>
1.1	Speaking to theorists, searching for facts, and whispering in the ears of princes	7
1.2	Aim and relevance of this review	9
<b>2</b>	<b>Review of tax experiments</b>	<b>10</b>
2.1	Traditional economic factors of deterrence	10
2.2	Social norms, equity, and moral appeals	13
2.3	Positive incentives	16
2.4	Fiscal exchange and accountability	18
2.5	Information and third-party reports	19
2.6	Tax collectors and their incentives	21
2.7	Complexity and compliance costs	22
2.8	Learning and long-term perspective	23
<b>3</b>	<b>Conclusions</b>	<b>24</b>
	<b>References</b>	<b>27</b>
<b>Tables</b>		
Table 1	Types of tax experiments	9

# Acknowledgements

This paper was produced with the support of the International Centre for Tax and Development (ICTD), which is financed by DFID and NORAD. An earlier version of this review was published as IDS Evidence Report No. 97. I am particularly grateful for the comments provided by Mick Moore and two anonymous reviewers. All remaining errors are my own.

# Acronyms

EBM	Electronic Billing Machines
EITC	Earned Income Tax Credit
IRS	Internal Revenue Service
TE	Tax Experiments
VAT	Value Added Tax

# 1 Introduction

Tax compliance is certainly amongst the greatest obstacles that tax administrations face in mobilising revenues. It has to do with the extent to which citizens comply with tax laws or, conversely, evade and avoid taxes. Tax compliance has obvious effects on the level of tax collections: the lower the level of compliance, the less the revenues readily collected by the government. However, it also has important implications in terms of tax administration. When compliance is low, the government has to resort to coercion and aggressive enforcement to raise revenues. Not only are these methods of collection costly for tax administrations, particularly in low-income countries; but also they denote a relation between citizens and the state based on force and fear rather than a healthier understanding of the fiscal contract underlying taxpaying.

Clearly in any country taxpaying is the result of both quasi-voluntary compliance and strong institutions. It would be simplistic to think that all taxpayers comply with tax laws just because they think this is a good thing to do. An element of coercion is present in most cases. For example, withholding procedures make it very difficult for employees to evade taxes on their wages. Therefore most of them do not have the choice of whether to evade or not. At the same time, it is also clear that the observed level of tax compliance is too high to be only the result of coercion (Andreoni et al., 1998; Alm et al., 1992). Most countries simply do not have the capacity to control and audit a large amount of taxpayers. Therefore the right question to ask may be ‘Why do people pay tax?’ rather than ‘Why do they evade?’ (Alm et al., 1992).

The answer is likely to be a mix of deterrence, related to the probability of audit and sanctions, and a range of factors that are often referred to as tax morale. Although there is no uniform definition for tax morale, a recent review (Luttmer and Singhal, 2014) unpacked five main mechanisms through which it can operate: 1) intrinsic motivation, 2) reciprocity, 3) peer effects and social influences, 4) culture, and 5) information imperfections and deviations from expected utility.

The literature on tax compliance is vast and diverse, spanning from quantitative to qualitative analyses and from surveys to lab games. Despite several differences in methods and approaches, the kind of questions asked, the underlying hypotheses and the theoretical foundations are very similar. In most cases the theoretical framework can be traced back to the model of taxpayer behaviour of Allingham and Sandmo (1972). These authors model tax evasion based on the theoretical frameworks used in the economics of crime (such as Becker (1968)) and the economics of uncertainty (such as Mossin (1968)). They hold that a taxpayer’s decision on whether to evade, and to what extent, is influenced by the expected economic returns to evasion. These returns are determined by the benefits of evasion, affected by the tax rate and taxable income; and by its costs, influenced by the probability of being detected and the size of sanctions if found guilty. However researchers soon realised that if taxpayers acted purely according to this model, we would observe much higher levels of tax evasion. This observation has spurred researchers to include other factors in the model, such as social and moral considerations, which are likely to affect the decision on evasion (Myles and Naylor, 1996; Traxler, 2010).

Based on these theoretical frameworks, the empirical literature has tested the validity of such determinants of compliance in practice. The obvious problem in doing so is that it is difficult to get honest answers about dishonest behaviour. For this reason, survey-based studies have focussed on perceptions of tax compliance rather than actual behaviour. Typically, they ask indirect questions such as ‘Do you think most people believe they should honestly declare income on their tax return?’. There are already a number of good reviews of this literature, such as Slemrod (2007) and, more specifically on African countries, Fjeldstad et al. (2012). The survey-based literature offers very useful insights into the question of why people pay tax. However, it suffers from a major drawback: the attitudes that people reveal in surveys may not accurately reflect their actual behaviour when faced with the real consequences of tax evasion and avoidance (Ajzen and Fishbein, 1977).

Tax experiments (TE) address this drawback by looking at real taxpayer behaviour rather

than perceptions – although section 1.1 makes important distinctions on the level of realism of different types of TE. This paper reviews the evidence emerging from this literature, paying particular attention to the growing number of large-scale field experiments. It proceeds as follows. Section 1.1 provides an overview of different types of tax experiments and their evolution in time. Against this background, section 1.2 states the specific aim and relevance of this review. Section 2 represents the core of this paper and summarises the main findings in the literature, classified by determinant of compliance. Finally, section 3 critically takes stock of this literature and looks forward to its future.

## 1.1 Speaking to theorists, searching for facts, and whispering in the ears of princes

Tax experiments use experimental methods to identify causal effects of the determinants of compliance. The authors in the early TE literature (Spicer and Thomas, 1982; Friedland et al., 1978) argue that such causal links could be identified in a lab setting, where it is possible to control the environment and independent variables better than in real-life or survey-based research. They also argue that the artificiality of the lab allowed them to circumvent the problem of getting honest answers about dishonest behaviour by working within hypothetical settings. The more recent TEs instead turned to administrative data from tax returns, and the causal links are established thanks to randomisation. These are two separate types of TE, as discussed in more detail in this section. While the latter (field TE) is more relevant to understand real taxpaying behaviour and inform policy, the former (lab TE) helps to understand the origins and theoretical foundations of this literature.

In his 2002 review of tax experiments, Benno Torgler classified the first type of TE in two categories labelled ‘speaking to theorists’ and ‘searching for facts’, based on the definition provided by Roth (1995) (Torgler, 2002). In both cases, the primary issue of interest is tax compliance, with a common focus on income taxation – although similar arguments can be made for other taxes. The first category includes lab experiments aimed primarily to test theories, and whose results feed back into the theoretical debate and literature. In particular, several early studies tested the empirical validity of the economic model of taxpayer behaviour of Allingham and Sandmo (1972). These *speaking to theorists* studies therefore focused primarily on the role of deterrence, sanctions and expected economic returns to explain compliance (Spicer and Thomas, 1982; Friedland et al., 1978; Beck et al., 1991; Spicer and Hero, 1985). In this case, experimenters speak to theorists. The typical lab experiment would involve a group of students who are faced with the decision on how much income to report, based on a set of given variables including, for example, the tax rate, the probability of audit, and sanctions. The artificial lab setting then allows the researcher to change these variables and to observe how the participants’ decisions change as a response.

The second category, *searching for facts*, includes experiments that analyse variables that are not explicitly included in the traditional economic models, but that are motivated instead by the results of other experiments and by observations of what actually happens in the real world. One of these observations is that the levels of compliance in many countries are much higher than those predicted by purely economic models. This has prompted researchers to explore the role of public goods, and moral and social factors in tax compliance, in addition to the traditional economic considerations. In this case, experimenters are talking to each other (Torgler, 2002). Although experiments looking at these issues are searching for the facts, they speak to theorists too, since the new observed patterns can feed back into the theory to develop existing models further.

*Speaking to theorists* and *searching for facts* studies have some common characteristics. First, they typically use students as experimental subjects. While students respond to the incentives set out in the experiment, they are not real taxpayers and they may not have any experience at all with taxpaying in real life. Second, these studies usually rely on small samples, often

involving only a small number of students in repeated rounds. This presents problems related to statistical power, and the resulting estimates may not be stable and reliable. Finally, they use largely artificial experimental settings, despite efforts to move to more realistic environments. In a lab setting, the individuals involved obviously know that they are part of a study and they may behave differently than they would under the pressures of real life. Behaviour in the artificial environment of the lab is likely to differ from actual decisions when faced with a real social context and real consequences of wrongdoing. This problem has largely prevented lab TE from having much policy impact. Therefore lab TE have been well suited to understand the theoretical mechanisms that drive compliance, but the relevance of their findings in real taxpaying situations has to be tested in the field.

In an attempt to achieve greater realism in tax experiments, some researchers have moved the lab to the field (lab-in-the-field TE). Although this helped in getting insights into cross-cultural differences (Cummings et al., 2009), they still suffer from very similar methodological shortcomings.

Many of the drawbacks of lab TE were overcome with the emergence of large-scale field TE. Following the categorisation proposed by Roth (1995), these TE can be labelled as ‘whispering in the ears of princes’, as experimenters aim to speak to policymakers. To do so, the lab setting is substituted with real life situations: the participants in the study are real taxpayers, the data comes from real returns, and the experiment occurs in the normal taxpaying environment where the main counterpart is the tax administration. The typical field tax experiment is based on a comparison of randomly selected treatment and control groups. The treatment consists of the provision of information that is meant to change key determinants of tax compliance, like perceptions on the probability of audit, fairness and tax morale (with all the aspects it encompasses). This information is typically delivered through a letter that highlights, for example, the fact that the tax administration routinely carries out audits and that those who are caught evading face severe fines; or the importance of paying taxes to finance important public services such as schools and hospitals. Because the participants in the treatment and control groups are chosen based on randomisation, they are very similar with respect to all other aspects such as age, income, and location. The only difference between groups is therefore whether or not they received the informational treatment in the form of a letter. By observing the differences in taxpayer behaviour between groups, and before and after receiving the letters, researchers can establish the causal effect of different determinants on compliance.

By doing this, they can show, for example, whether taxpayers are more affected by the probability of being caught or by an understanding of the social value of taxpaying, and by how much. Field experiments, by investigating behavioural factors of taxpaying, can potentially provide guidance on how to raise tax revenues in a cost-effective way by changing perceptions and the information available to taxpayers. This is particularly important given that credibly increasing enforcement is very costly and not always desirable. With this information, policymakers can weigh their policy options and target their compliance strategies more specifically to the factors that generate the larger responses at the most reasonable cost. In some cases, researchers have studied existing initiatives of the tax administration, or variations in the law or other relevant variables, to evaluate their effect on compliance. These studies are natural experiments or quasi-experiments, which are also included in this review. Although they do not use randomisation through a formal control and treatment group, they are useful to evaluate policy and are often able to establish causal effects.

The types of experiment described so far also represent the evolution in time of this field of research, which has seen tax experiments progressively moving away from theory and closer to real life experiences. This movement from the lab to the field has involved both the introduction of new questions and changes in the design of experiments. In addition to deterrence and tax morale, new questions have focussed around more practical issues, such as the complexity of tax systems, the incentives of tax collectors, and compliance costs for taxpayers. As far

as research design is concerned, lab experimenters have made greater efforts to approximate real life conditions by using tax specific language, simulations of exchanges of resources, and participants with experience of taxpaying rather than students. However the lab is still an artificial setting, so the movement to reality of TE can be considered completed only with large-scale field experiments. In addition to being more realistic, field experiments involve much larger samples than lab TE, therefore also benefiting from higher statistical power.

The types of TE described in this section are summarised in table 1, which necessarily makes some over-simplifications – such as dividing the determinants of compliance into three general categories of deterrence, tax morale and practical issues.

**Table 1:** Types of tax experiments

	Lab	Field	Deterrence	Tax morale	Practical issues
<i>Speaking to theorists</i>	✓		✓		
<i>Searching for facts</i>	✓		✓	✓	
<i>Whispering in the ears of princes</i>		✓	✓	✓	✓

## 1.2 Aim and relevance of this review

Against the background of section 1.1, this paper aims to update Torgler’s review of 2002 by adding tax experiments that are *whispering in the ears of princes*. At the time Torgler (2002) was published there was not an extensive literature of this sort - with the exception of Slemrod et al. (2001) - which justified its exclusion from his review. However in the last decade field experiments have been thriving, forming a growing strand of literature that can no longer be ignored. This review aims to fill this gap, and is particularly relevant for three reasons.

The first one has to do with methods and to the ‘credibility revolution’ that applied economics has been experiencing in recent years (Angrist and Pischke, 2010). Field experiments combine rigorous methods with high policy relevance better than other methods in the tax compliance literature. For this reason, Slemrod and Weber (2012) argue that they are the most credible method to evaluate policy interventions to address tax compliance and, as such, they should be used as a benchmark. Despite being at the basis of this ‘credibility revolution’, randomisation has attracted both supporters and critics. The critics are particularly concerned with external validity and with the possibility of scaling up interventions (Deaton, 2009; Basu, 2013). It is certainly true that an experiment’s results are strictly valid only in the specific time and context in which the experiment takes place. However, field experiments have an important advantage over the typical randomised control trial. Because they can use population-wide data that is nationally representative, they can directly evaluate options for real policy interventions that, if proven effective, are more easily scaled up to the national level. While the external validity of a specific study to other countries remains a concern, it is possible to draw more general conclusions as evidence keeps building up from different countries.

Secondly, and related to the previous point, this review makes a particular effort to include studies that focus on middle- and low-income countries. Most of the TE literature, as well as the compliance literature more generally, has been largely concerned with Europe and the U.S. However, more and more studies are being published with evidence from Latin America and Asia, while a few are ongoing in Africa. Tax compliance can be seen as a global issue, which all revenue administrations face to different extents, and whose possible determinants are likely to be very similar across countries. However there are at least three reasons that so far prevented field experiments from spreading more widely outside Europe and the U.S. The

first one is related to data availability. In the past two decades, most low- and middle-income countries have undergone great transformation in their tax administration by establishing semi-autonomous revenue authorities, and more generally by modernising the way tax collection is organised. This has entailed the digitalisation of tax records, which have become available to researchers only recently. There are still data quality concerns that are likely to be more severe than for more established administrations. Secondly, field tax experiments require a high degree of commitment by the local government. Since most tax administrations in low-income countries operate under great capacity constraints, it can be particularly challenging for them to find the necessary resources to dedicate to tax research. Thirdly, because revenue administrations in these countries are relatively new, they are developing only recently the practice of working with researchers – this includes practical issues such as the appropriate anonymisation of data, but also more generally trust and appreciation of the value of policy-relevant research.

Finally, field tax experiments use anonymised administrative data from tax returns. The use of this data by researchers is a relatively new development in many countries, which partly explains why this literature has expanded only recently. This type of data has at least two important advantages. First, it allows the observation of actual taxpaying behaviour in real life. However, since tax returns only capture what is known by the tax administration, they do not allow the direct observation of tax evasion and avoidance. Experimental methods allow researchers to address this problem, but only partially. Those taxpayers who are completely informal - not registered for tax purposes - are still missed by field TE. Secondly, the effective use of administrative data requires a high degree of collaboration with the tax administration that originally compiled and maintains the data. Therefore, these studies have a built-in connection with policymakers, which is beneficial both to design relevant research projects and to make results useful for supporting policy.

## 2 Review of tax experiments

This section reviews the tax experiments literature and is organised around eight issues that matter for tax compliance. They are often interrelated, sometimes overlapping, and almost always co-exist in various combinations. The main focus is on field TE, but evidence from lab TE is also included. Although most lab TE belong to the early literature, they have recently made advances in integrating methods from psychology to get additional insights into taxpayer behaviour. Many of the hypotheses tested in field TE come from the lab literature, and it therefore seems relevant to draw a clear connection between these distinct but highly interrelated strands of literature. Furthermore, the main results reviewed here are sometimes cross-validated against non-experimental studies that use data from tax returns. Since tax experiments are still a relatively narrow strand of literature, linking them to the broader context helps to support the empirical validity of their results.

### 2.1 Traditional economic factors of deterrence

The traditional economic factors set out in Allingham and Sandmo (1972) include variables that influence the costs and benefits of evasion, such as: the tax rate, the probability of audit, sanctions, and enforcement. Most of the early *speaking to theorists* lab TE test the effects of these variables on the probability of evading (the extensive margin), and the amount of taxes evaded through income under-reporting (the intensive margin). The pioneering simulation study by Friedland et al. (1978) analyses changes in the tax-paying behaviour of fifteen Israeli psychology students when these economic factors vary. When the tax rate rose above a certain level, the participants responded by sharply decreasing their reported income. Moreover, large sanctions appeared to be more effective than frequent audits in improving compliance. Finally, decisions on the extensive and intensive margins are driven by different factors. While the tax rate seems

to have the largest impact on the former, socioeconomic factors are more important to determine the latter. However, conclusions from this experiment need to be taken with caution, because of the small sample size and related power problems. An additional drawback is that in the lab environment all the key variables are fully known to participants, while in real life probability of audit is not known precisely and even the exact amount of sanctions can be uncertain.

Confirming the importance of uncertainty, Spicer and Thomas (1982) found that the predicted deterrence effect of the probability of audit is only empirically valid when taxpayers have full information about it. However, even when information is available, an increase in the probability of audit only affects the decision to evade and not its extent. Beck et al. (1991) reinforce these results by showing that, although evasion is lower when audit probabilities and sanctions are higher, uncertainty weakens these effects. Moreover, when actual variables are not known due to uncertainty, perceptions play an important role. Spicer and Hero (1985) test two hypotheses with a group of thirty-six psychology students from the University of Colorado: first, taxpayers evade more when they perceive that the level of evasion in general is higher; second, those who have been audited in the past perceive a higher probability of audit and therefore tend to evade less. The authors found no support for the former, but did find evidence for the latter. Although these results still stem from highly artificial settings (including other studies such as Mittone (1997)), with small samples of students, they help to understand the possible reasons why compliance in real life is higher than in the theoretical model of Allingham and Sandmo (1972). In the presence of uncertainty, taxpayers tend to evade less even if audit probabilities and sanctions are low. This may be partly due to upwardly-biased perceptions of the risk of being caught.

Turning to field experiments, Slemrod et al. (2001) carry out a pioneering study of this kind based on a sample of 22,368 tax returns from taxpayers in Minnesota, USA. The authors, in close collaboration with the Department of Revenue, collected income tax data for two subsequent years: 1993 and 1994. The treatment consisted of a letter stating that the taxpayer had been selected to be part of a study and that their state and federal tax returns would be closely examined. The letter also stated that taxpayers would be contacted regarding any discrepancy and that if any irregularities were found, their previous tax returns would be scrutinised as well. The letters were sent by the Commissioner of Revenue in January 1995, at the beginning of the filing season for 1994, to a treatment group of about 1,500 taxpayers. The letter was meant to increase the perceived probability of audit and therefore to affect the accuracy of income tax reporting. A difference-in-difference methodology was used to analyse the changes in reported taxable income between 1993 and 1994 in the treatment and control groups. Changes in reported income were then used as an indication of changes in compliance, as data on the audit results (the close examination) were not available to the authors.

Before discussing the results, it is worth noting that Slemrod et al. (2001) introduced two methodological elements that are common to the rest of this literature. Firstly, the letters are sent directly from the revenue administration. This helps to make the experiment real and therefore to capture real behaviour. Secondly, the treatment does not change actual tax parameters, such as the rate or the number of audits. Instead it only varies the taxpayers' perceptions and the information available to them. On the other hand, Slemrod et al. (2001) differ from other studies in the literature on at least two methodological choices. Firstly, in this case the treatment is relatively strong, as it hints with some degree of certainty to an upcoming scrutiny of the taxpayer's return. Other experiments have adopted softer approaches that only increase the salience of audits without hinting to an upcoming audit for any specific taxpayer (Castro and Scartascini, 2013). Secondly, Slemrod et al. (2001) decide to let taxpayers know that they were part of a study. Other researchers, such as Kleven et al. (2011), instead prefer to conduct the study as if the treatment was a normal letter from the tax administration. By doing this, they believe that the message may be taken more seriously. Clearly, taxpayers' reactions to the treatment are a mix of both things: they may know it is part of a study, but if the treatment

is strong enough, as in the case of Slemrod et al. (2001), they are still likely to respond.

Indeed Slemrod et al. (2001) find that the letter had a significant impact on reported income, in some cases increasing compliance by 12 per cent. However this effect varies across categories of taxpayers, who were stratified along two dimensions: income and opportunity to evade. The latter is defined according to the type of income being taxed: high-opportunity occurs in presence of self-reported income, or income from rents and royalties. In a situation of high opportunity to evade, tax payments increased by 12 per cent for low- and middle-income taxpayers. Among the high-opportunity taxpayers, those with self-reported income represented a major component of the treatment effect. For low-opportunity taxpayers in the low- and middle-income groups, the letter had a much smaller effect, although still positive. As far as high-income taxpayers are concerned, the results instead are somewhat perverse: reported income decreased as a result of the ‘threat of audit’ letter, indicating lower compliance as a result of increased audit probability. One possible explanation is that high-income taxpayers sought professional advice in response to the letter, which may have helped them to find legal ways to reduce their taxable income.

The authors underline a couple of caveats of the experiment design, which again are relevant for other field TE as well. Firstly, some taxpayers may already have thought, before receiving the letter, that the authority closely examines most returns, including theirs. In this case, the taxpayers would not respond to the letter, because this treatment would not change their perceptions on the probability of detection. Similarly, they may not believe that such close inspection is actually capable of uncovering any wrongdoing, and therefore they may still not modify their behaviour in response to the message. As highlighted in section 1 and further discussed in section 2.5, the credibility of enforcement is a key issue. In the long term, a high perceived probability of audit can only be sustained by a relatively high number of actual audits and detections. Second, the indication in the letter that previous returns could also be inspected may have induced taxpayers to maintain their reporting pattern to avoid uncovering a history of non-compliance.

Kleven et al. (2011) largely confirmed the importance of deterrence by using administrative data from the Danish tax collection agency, involving 42,800 randomly selected taxpayers in 2007 and 2008. In 2007, a group of these taxpayers was randomly selected for unannounced tax audits. Irregularities were corrected and penalised according to the law. In 2008, both audited and non-audited taxpayers were selected for pre-announced audits. Two treatments were tested against a control group: one letter informed them that their tax return would certainly be audited; and the other that half the people in the group would be audited. By doing this, the treatment exogenously changed audit probabilities. In this case, participants were not aware of being part of a study. The threats of audit had a large effect on compliance, which was entirely driven by self-reported income. An important advantage of Kleven et al. (2011) in comparison to other field TE is that they had access to audit data and could therefore directly measure variations in compliance. Importantly, audited taxpayers were selected randomly and not based on risk, which would be endogenous to their evasion behaviour. Using this data, the authors show that compliance rates in Denmark are high despite relatively high tax rates. Consistent with the importance of having the opportunity to evade, third-party reporting environments display close to full compliance, while substantial evasion occurs when income is self-reported. In addition, the authors used a quasi-experimental setting to assess the effect of changes in the tax rate on compliance. In this case the research design exploits salient kinks in the income tax schedule, and compares taxpayers on each side. Marginal tax rates do not seem to have a large effect on compliance.

Castro and Scartascini (2013) evaluated the effectiveness of similar threatening letters in Latin America. They analysed property taxes in a municipality in Argentina to test the effect of different factors on taxpayer compliance, including deterrence but also equity and fairness. This property tax is based on the linear front size of the property, making the tax base observable by the authorities. Therefore in this case, as in Kleven et al. (2011) though with a different

method, the authors can directly measure compliance. Importantly, this tax is linked to the provision of public goods by the same level of government that levies the tax. In total 23,000 individual taxpayers were randomly divided into four groups, three of which received a treatment letter influencing their perceived probability of detection or their beliefs. All letters provided information: the first one about enforcement and possible fines; the second one about other taxpayers' behaviour, thus affecting beliefs on equity; and the third about how government resources are used, thus underlining fiscal exchange. Compared with Slemrod et al. (2001) and Kleven et al. (2011), the deterrence message here is much softer, along the following lines: "did you know that if you do not pay your taxes on time you will have to pay sanctions and you may be prosecuted?". To make the message clearer and more salient, the authors included a monetary example of underpayment with the relative amount of fines potentially due. This message increased compliance by almost five percentage points, while the other treatments did not have any effects compared with the control group.

Other studies in the literature typically include a deterrence treatment along with other factors. They generally find deterrence to be effective (Fellner et al., 2013; Ariel, 2012; Bott et al., 2014; Dwenger et al., 2014; Cummings et al., 2009). Since the main results are related to the other factors explored, they are reviewed in other sections.

## 2.2 Social norms, equity, and moral appeals

As noted in section 1, tax morale is a broad concept that encompasses various factors. Based on the classification of Luttmer and Singhal (2014), this section focuses particularly on intrinsic motivation, culture, and peer effects and social influences.

Some of the studies reviewed in the previous section also test the importance of social and moral factors, in addition to traditional economic ones. Castro and Scartascini (2013) do not find any significant effect from the messages on equity and public services. In addition, a study by Blumenthal et al. (2001) uses a very similar design and method as Slemrod et al. (2001) and was also conducted in Minnesota. This time however the authors investigated the effect of moral appeals, using two treatments. The first one is a 'support valuable services' message, communicating how taxes are spent and encouraging taxpayers to comply voluntarily to support them. The second one is a 'join the compliant majority' message, indicating that the majority of citizens comply with tax laws and that one should be compliant in order to be part of that majority. Contrary to Castro and Scartascini (2013), who included their messages directly in the tax bill, Blumenthal et al. (2001) sent dedicated letters. As in the rest of this literature, the letters would start and end with the same text, a reminder of deadlines and indications on where to get additional information, while the middle paragraph reported one of the two messages above. The authors found little or no evidence that the letters resulted in any changes in compliance behaviour, with high-opportunity taxpayers (defined as in Slemrod et al. (2001)) being even less likely to respond to normative appeals.

The lack of empirical support for the importance of moral and social factors may be due to various reasons. Moral suasion may not have a large effect if the elasticity of compliance to perceptions is low. Moreover, evaders may have generally lower trust in government, therefore making the moral appeals less effective for them. However variations in compliance come from precisely this group, as they have some margin to increase reported income. From a methods perspective, Castro and Scartascini (2013) suggest that their 'equity' letter may have actually changed beliefs in unexpected ways. Taxpayers may have discovered that evasion is more widespread than they thought, therefore being encouraged to evade more. This is probably related to the specific message, illustrating the importance of the context and design of field TEs. While for the U.S. Blumenthal et al. (2001) could state that 'people who file tax returns report correctly and pay voluntarily 93 per cent of the income taxes they owe'; the message in Castro and Scartascini (2013) reads 'did you know that only 30 per cent of taxpayers do not pay the [property tax]?'. Since compliance levels are much lower for the Argentinian property

tax than for the U.S. income tax, similar information messages may have very different effects: a feeling of unfairness towards the 30 per cent that do not pay in the former case and a feeling of social shame for being part of a small minority (7 per cent) in the latter.

Moral appeals are also explored by Torgler (2004) in a field experiment conducted in collaboration with the local tax administration in Trimbach, Switzerland. Out of 580 randomly selected taxpayers, in 2002 a sub-group of them received a letter signed by the commune's fiscal commissioner. The letter was sent in a separate envelope and printed on a pink sheet to increase taxpayers' awareness of it. Participants were not informed they were part of a study, to elicit true behaviour. The relatively complex text of the letter alluded to factors such as: active citizen participation in the life of the commune, citizens' sense of responsibility, and the importance of tax compliance in 'preserving this democratic and liberal structure' and to 'keep Trimbach attractive for its inhabitants'. The effect of the letter was assessed against two compliance indicators: timely filing and timely paying. From a methodological perspective, these compliance indicators have the advantage of being observable – unlike the level of under-reported income used in other studies. The results largely confirm the previous findings that moral suasion has virtually no effect on compliance. The treatment group did not significantly change the timeliness of their returns or payments in response to the moral appeal. The relatively complex treatment message may also be a reason for this lack of results.

Evidence on TV licences from Austria further supports the previous findings (Fellner et al., 2013). This experiment involved a sample of 50,498 individuals that, contrary to the other studies reviewed so far, is composed exclusively of potential evaders. This design feature has implications for the interpretation of results since, as discussed above, evaders may be less motivated by moral appeals. All taxpayers, except a control group of about 5 percent of the total sample, received a letter from the Fee Info Service, the managing agency for the TV licence. One group received a control letter containing a standard, neutral text. This standard text was then complemented with three treatment messages: a threat, social information emphasising the high level of compliance, and a moral appeal stressing fairness. Moreover, for some participants, the threat treatment was combined with either social information or a moral appeal, making a total of six different letters. The messages clearly indicated that participants were selected because they were suspected of evasion, making treatments stronger. Compliance here is measured in terms of new registrations, as a binary variable indicating whether an evader started to comply or not. This feature makes this study different than the typical field TE, as tax administrations usually do not know who the non-registered evaders are. Generally, the letters were effective in increasing new registrations, compared to the 'no letter' control group. However the bulk of the effect came from the neutral letter, indicating that the pure fact of receiving any letter may be more effective than the specific contents. Nonetheless, amongst the different treatments, the threat message was the only one to have any effect in addition to the neutral mail. A survey designed to investigate the mechanisms of increased compliance, shows that it is mainly due to a signal of surveillance and to increased sanction risk. Although the social information did not increase compliance overall, there is substantial heterogeneity across regions. Where compliance is relatively high, the message had a weakly negative effect, while where it is relatively low, it had a weakly positive one.

Further evidence on VAT payments by corporations in Israel supports the apparent ineffectiveness of moral suasion (Ariel, 2012). Out of 4,395 Israeli corporations, two groups of about 730 taxpayers each received either a deterrence message or a moral appeal. The latter stressed the social implications of non-compliance ('we all suffer from noncompliance') and informed companies how public money is spent. The study used gross sales values reported to the authority, tax payments, and tax deductions as measures of compliance. None of the messages significantly increased compliance. However, corporations receiving the moral appeal claimed higher deductions, implying a decrease in government revenues. This backfiring effect recalls the perverse finding of Slemrod et al. (2001) where high income taxpayers decreased compliance

as a result of the deterrence letter, and it may be motivated by similar reasons. Corporations, much like high-income individuals, have large resources to dedicate to aggressive tax planning.

While these results may seem discouraging, lab experiments and recent field studies show that moral and social factors can have an important role to play in increasing compliance.

Building on the methods of the early ‘searching for facts’ experiments (Bosco and Mittone, 1997; Mittone, 1997), recent lab TEs linked the traditional economic factors even more closely with methods and theories from psychology. For example, Coricelli et al. (2010) analyse psychological costs and benefits of evasion using data from self-reports and skin conductance responses.<sup>1</sup> The experiment, involving eight subjects in each repeated round, showed that taxpayers’ decisions to evade are dictated by the moral and emotional implications of cheating, and not by monetary calculations alone. In addition, they show that public disclosure of deceptive behaviour, achieved by displaying the picture of cheaters, deters evasion. Looking at repeated rounds, they can show that audits decrease the extent of cheating in the next period, but larger fines increase the proportion of evaded income. Still in a lab setting, Coricelli et al. (2014) investigate two types of shaming: one that results in stigmatisation and exclusion from the community, and another that leads to reintegration of the evader. In both cases the shaming ritual involved displaying in public a picture of the evader, but in the reintegration treatment the picture immediately disappeared. Stigmatisation and exclusion from the community significantly increased evasion compared to reintegration. Calvet Christian and Alm (2014) further show that the promotion of empathy and sympathy can help to increase compliance.

Clearly, these studies suffer from statistical power and from the artificiality of the lab setting. However, they offer interesting insights into the emotional dimension of tax evasion that may guide empirical investigations in the field. For example, some countries adopt policies of naming and shaming that are the real-world equivalent of displaying images of cheaters in the lab. The lab evidence shows that although they can be a powerful tool, the intense negative emotions associated with stigmatisation should also be taken into account.

As far as large-scale field experiments are concerned, a recent study conducted in the United Kingdom (Hallsworth et al., 2014) unpacks moral factors into social norms and concerns over public goods. The basic message for the former was that the majority of people pay tax, while the latter underlined the importance of taxes to finance public goods for all. Respectively three and two variations of these short messages were used as alternative treatments, in addition to a neutral control letter. For example the public goods message was phrased in a positive (‘Paying tax means we all gain from vital public services’) and negative way (‘Not paying taxes means ...’); while the three versions of the ‘minority norm’ message presented the same message in various degrees of specificity to the individual. The most specific and strongest one stated that ‘nine out of ten people in the UK pay their tax on time. You are currently in the very small minority of people who have not paid us yet.’ Similarly to Fellner et al. (2013), the letter is directed to those who failed to pay on time (a sample of 100,000 individual taxpayers). Thus, compliance can be fully measured, since the tax administration can observe whether and to what extent taxpayers have paid due taxes. All letters resulted in an increase in the probability to clear outstanding taxes, with the greatest effect from the strongest version of the ‘minority norm’ treatment. Although this treatment is designed around social norms, its effect may also be related to deterrence since it makes explicit that the revenue administration knows that the recipient is not compliant. Following these results, the authors carried out a second experiment where they compared the effect of descriptive norms (i.e. how other taxpayers behave) and injunctive norms (what other people think is the right thing to do). The former is found to have a significantly larger effect than the latter. These experiments show that large revenue gains can be made by increasing moral costs to non-compliance, with negligible additional costs to

---

<sup>1</sup>Coricelli et al. (2010) report that “The SCR is a phasic component of the electrodermal activity primarily controlled by the sympathetic division of the autonomic nervous system. It is associated to emotional arousal in a wide range of psychological states and processes.”

the revenue administration.

The importance of moral appeals is confirmed in a recent experiment carried out in Norway among taxpayers who earned income in a foreign country (Bott et al., 2014). This type of income is not pre-populated in the tax return, so compliance fully relies on self-declaration. However the Norwegian Tax Administration recently gained access to third-party reports on this type of income. These reports allowed the identification of a sample of taxpayers whose self-reported amount was substantially different than the third-party reported one. In addition to a neutral letter, three treatments were used: 1) an equity message reminding taxpayers that most people pay their taxes fully and on time; 2) a message underlining the importance of taxes to finance public goods and services; 3) a deterrence letter. The main result is a large and significant effect of the moral appeal, particularly the equity treatment. The amount reported by those who received the moral treatment was almost double that of the base letter. Increasing the probability of detection also had a large effect on reported income, but in a different way. While moral appeals resulted in increasing the amount of income reported – the intensive margin, the detection message increased the number of people declaring foreign income – the extensive margin. A complementary survey explored how the letters were perceived by recipients and whether there were cross-contaminations whereby, for example, the moral treatment also affected the perceived probability of detection. The survey shows that moral appeals actually affected moral motivation rather than working through the probability of detection.

Del Carpio (2013) uses a mix of surveys and experimental methods to look at perceptions of compliance and enforcement, where the latter is very similar to the social or minority norm treatments used in the rest of the literature. The baseline survey reveals that in Peru the true levels of both compliance and enforcement are generally underestimated. Exploiting this feature, the experiment focuses on providing the true information on compliance and enforcement, both in separate treatments and jointly in the same letter, as well as a neutral message to remind of deadlines. The treatments' effect is evaluated on property tax payments in Peruvian municipalities. As in the case of Castro and Scartascini (2013), compliance is fully observable thanks to cadastral registries that allow the distinction between taxpayers who comply and those who don't. The strongest effect is generated by the information letter on true levels of compliance, which increases payments by 20 per cent with respect to the no-letter control group. Comparing this effect with that of the neutral message reveals that about half of it is due to the pure reminder effect of the letter. Interestingly, the joint treatment, where both true compliance and enforcement levels are revealed, has a lower effect than the social norms treatment alone. This can be interpreted as a crowding out of intrinsic motivations to comply. Moreover a post-experiment survey shows that all treatments raise beliefs about compliance and enforcement (except for the pure reminder that only affects the former).

Other studies, which are reviewed in other sections, looked at moral suasion amongst other factors and found both positive effects (lab-in-the-field study by Torgler (2003)) and no effect (field TE by Dwenger et al. (2014)). Generally, the literature is not fully conclusive on the effectiveness moral and social factors. Within countries, they are likely to play an important role in explaining the observed *level* of compliance. However, *variations* in compliance, which is what field experiments capture, rely on evaders starting to comply. The fact that evaders may be less affected by moral and social appeals could therefore partly explain the lack of significant results in many studies. Across countries, the same appeal is likely to have different effects in different social, institutional and cultural settings. However differences in design prevent direct cross-country comparisons, so we cannot identify how country-specific conditions influence results.

### 2.3 Positive incentives

Positive incentives are closely related to moral and social factors, as they are meant to strike a chord related to feelings of social acceptance and recognition. However, while moral messages

appeal to intrinsic motivations, positive rewards have an element of extrinsic motivation, as they can still be driven partly by self-interest. Contrary to deterrence, however, these incentives provide positive rewards to taxpaying, rather than negative sanctions for evaders.

In an early lab study, Alm et al. (1992) tested four different options of positive rewards as incentives for compliance: 1) a lottery that could be accessed by those who were tax compliant, where the chance of winning was one in twenty-five; 2) a fixed reward for compliant subjects; 3) an audit reduction; and 4) a public good. The lottery had the largest effect on compliance, but it had the same expected value of the fixed reward. Audit reductions and public goods improved compliance but less so than other treatments. The authors therefore concluded that positive rewards need to be immediate and salient to have a significant effect. In an attempt to improve the realism of lab evidence, Torgler (2003) conducted a lab-in-the-field experiment in Costa Rica using thirty-seven real taxpayers and real money, but neutral rather than tax-specific language. The treatments focussed on fiscal exchange, moral suasion and positive rewards. The results showed that positive rewards were most effective in increasing compliance, followed by the other two treatments. The number of subjects per treatment is very small (between eight and thirteen) and hence, the statistical analysis may have suffered from power problems. Moreover the experiment only included one round, so questions on the sustainability of moral suasion over time and on learning effects remain unanswered. While these two studies start shedding light on the role of positive rewards, they suffer from the typical limitations of the lab setting, including artificiality and small sample sizes.

A field experiment from Germany (Dwenger et al., 2014) points to the importance of distinguishing between intrinsic and extrinsic motivations of taxpaying. Intrinsic motivations are related to social norms and feelings of guilt and shame, while extrinsic motivations have more to do with enforcement and detection. The authors study a legally binding local church tax that encourages overpayments, so individuals can be evaders, compliers, or donors. These groups can be clearly distinguished because the relevant tax base used by the Church, reported taxable income, is fully observable from tax declarations. While the Church in principle could cross-check this data from the tax administration, it has never done so. This feature allows the identification of taxpayers who comply entirely due to intrinsic motivations (20 per cent) and those who evade in part or in full in absence of enforcement (80 per cent). Both groups received three sets of treatments, in addition to a neutral letter, which aim to: 1) increase the salience of the tax and inform taxpayers of the zero audit probability, to correct possible misperceptions about audits; 2) underline deterrence and communicate positive audit probabilities; and 3) appeal to moral values and provide positive rewards. The first letter just confirms that there are no misperceptions about the absence of deterrence, confirming that the group of compliers is intrinsically motivated. Increasing the probability of deterrence increases compliance substantially only for those taxpayers that are extrinsically motivated. Amongst the third set of treatments, a first sub-set included social norms and moral appeals similar to other studies reviewed in section 2.2. These treatments do not produce any significant effect, in line with most of this literature. The second sub-set included public recognition and positive rewards. The former involved the publication of compliant contributors' names in the local newspaper, while the latter allowed compliant contributors to take part in a lottery and possibly gain a monetary reward. An additional treatment combined both the recognition and monetary reward in the same letter. These treatments had a positive effect on compliance for the intrinsically motivated, with larger effects when social recognition was part of the treatment. However rewards further reduced compliance for the extrinsically motivated – those who evade in absence of enforcement. A possible explanation is that such rewards reinforced a view that tax payments are voluntary rather than mandatory, and that enforcement is weak. These results show, importantly, that taxpayers' responses to various treatments depend on what motivates them to pay taxes to start with.

An ongoing field experiment in Uruguay also focuses specifically on the relative effectiveness of positive and negative incentives on compliance. Although the study is currently ongoing, a

pre-analysis plan includes the details of the research design (Dunning et al., 2014). The study evaluates a policy initiative adopted by the municipality of Montevideo where a year-long tax holiday is randomly allocated to compliant taxpayers as a reward for their honesty. Taxpayers can enter a lottery if they have paid their taxes fully and in time, and thus have the chance to win the prize. The study aims to evaluate the effect on compliance for winning taxpayers, at different time periods after they won the lottery. A complementary survey is planned to capture the perceptions and other characteristics of winners and eligible non-winners. In addition a letter experiment, focusing particularly on property taxes, will involve the following treatments: 1) a reminder of deadlines as a placebo, 2) a reminder of the tax lottery and information about it, 3) a reminder about the sanctions attached to non-compliance. Furthermore, additional treatments add social motivations both to the positive and negative incentive messages. While the results of this study are not yet available, they will certainly represent an important contribution to the rather thin literature on the relative effects of positive and negative incentives.

Taken together, the field studies in this section and section 2.2 show that the response to social and moral factors is more complex and nuanced than previously documented in the lab. This response is likely to be influenced by existing beliefs on compliance and enforcement (Del Carpio, 2013) and by taxpayers' initial motivations to comply (Dwenger et al., 2014). Given this heterogeneity in taxpayers' responses, lab TE's small samples seem particularly unsuited to analyse the complexity of real life behaviour.

While the existing evidence on rewards is still relatively thin, positive incentives are a concrete and potentially important policy option. Several countries in Africa have already recognised this and have adopted various types of initiatives going in that direction. For example, countries like Rwanda, Zimbabwe and Gambia regularly hold events called 'Taxpayer appreciation day' where compliant taxpayers are publicly recognised by the authorities and rewarded for their contribution to the country's development efforts. In some cases they are also provided with an official certificate of compliance.

## 2.4 Fiscal exchange and accountability

Fiscal exchange is another one of the mechanisms of tax morale identified by Luttmer and Singhal (2014). It refers to the hypothesis that taxpayers contribute to the public purse because they appreciate the goods, services and guarantees that the government gives them in return. This aspect is sometimes analysed by field TE in the context of moral appeals, when the message underlines the importance of taxes to finance public goods (Blumenthal et al., 2001; Castro and Scartascini, 2013). Some of the experiments discussed in section 2.2 include treatments related to fiscal exchange. They generally find little support for its effectiveness in increasing compliance, revealing either no effect (Blumenthal et al., 2001; Ariel, 2012), or smaller effects than other moral appeals (Hallsworth et al., 2014; Bott et al., 2014; Torgler, 2003). However others find some evidence that public service delivery may be an important component. For example, Castro and Scartascini (2013) show that their treatments have heterogeneous effects based, amongst others, on the level of provision of public goods. While their evidence cannot be considered as causal, their findings suggest that a higher level of public goods provision is associated with higher compliance.

Building on previous lab studies showing that taxpayers evade less under conditions of uncertainty (see section 2.1), Alm et al. (1992) show that this relation is reversed when public goods are introduced in the experiment. When participants know that tax revenues are used to benefit everyone, to simulate a public good, then increased uncertainty leads to less compliance. This finding highlights the importance of the institutional environment in which the compliance decision is made. Other lab studies argue that rational utility-maximising taxpaying behaviour must take into account both tax payments and the transfers that result from the redistribution of the tax yield. Two lab studies (Becker et al., 1987; Mittone, 1997) confirm this hypothesis by showing that when taxpayers benefit from public spending, they are less likely to evade taxes.

In addition Becker et al. (1987) find that the perceived equity in government transfers helps to fight evasion. Still in the lab setting, Alm and Jackson (1993) analyse fiscal exchange, the use of tax revenue, and the decision processes for deciding such uses. Individuals tend to comply more when they know that the tax yield is spent on programmes they approve of and when they take an active part in the decision on spending. Compliance is higher when the outcome of the vote reflects widespread support rather than a close win. On the contrary, compliance is reduced when unpopular expenditure programmes are imposed top-down on participants. Similarly, the lab study by Lamberton et al. (2014) shows that allowing taxpayers to indicate preferences on government spending increases tax compliance.

Cummings et al. (2009) use a lab-in-the-field experiment to analyse the role of governance quality on compliance. Recognising the limitations of the lab, they also use survey data from Afrobarometer to corroborate the results. To test the hypothesis that perceptions on the quality of governance increase compliance, two parallel experiments are carried out in Botswana and South Africa – two countries with similar tax systems but different political histories and institutions. This hypothesis is clearly related to tax morale more broadly, whereby the motivation to pay tax derives from a social and moral obligation to contribute to society. However the link between taxpaying and public goods, through service delivery, depends on the quality of governance. The authors show that a lower level of governance is associated with lower compliance. Moreover, while increasing enforcement always increases compliance, this effect is larger in presence of better governance. This finding is echoed in Torgler et al. (2007) who combine findings obtained using survey data, macroeconomic indicators and experimental design. Not only do they argue that governance quality matters for compliance; but also they identify the quality of political institutions as one of the most important determinants. This broad indicator includes aspects such as accountability, rule of law, and control of corruption.

Other studies have focussed on the relation between taxpaying and public accountability using experimental or quasi-experimental methods, such as Paler (2013), Martin (2014) and Gadenne (2013). However they have been more concerned with testing the hypothesis that reliance on taxes has beneficial effects on public accountability, than with the importance of the latter as a motive for tax compliance. The first two are lab-in-the-field experiments, respectively in Indonesia and Uganda, while Gadenne (2013) uses municipality-level data from Brazil in a non-experimental setting. While they are beyond the scope of this review, it is worth mentioning that they also provide evidence of a positive relation of tax revenues with public accountability and the quality of public spending. In addition survey-based studies also support the fiscal exchange hypothesis (Ali et al., 2015; Fjeldstad and Semboja, 2001). Therefore they confirm the potential importance that this aspect can play in increasing compliance. Clearly however, the fiscal exchange motive involves important considerations about the quality of the institutional environment. Although the evidence so far does not allow us to draw cross-country comparisons, it is likely that low-income countries would find it particularly challenging to leverage this determinant of compliance, given the lower quantity and quality of public services available. Moreover, different types of public goods may generate different taxpayers' responses. For example, countries where violent crime is more widespread may respond more to appeals on the importance of taxes to finance policing than to support other public services. A recent study based on Afrobarometer data compares four African countries and explores fiscal exchange, amongst others (Ali et al., 2015). It shows that the link between compliance and public service delivery is based on specific services that are different in each country.

## **2.5 Information and third-party reports**

One way in which revenue authorities can enforce tax laws is by making full use of information, and particularly of third-party reports. This aspect is closely linked with the issue of deterrence and probability of detection, since the effective use of appropriate information increases the probability of detection. Some of the studies reviewed in previous sections specifically show that

the treatments are only or mostly effective on those who self-report income (Kleven et al., 2011; Slemrod et al., 2001; Bott et al., 2014). Kleven et al. (2011) argue that compliers are unable, rather than unwilling, to cheat. In particular, taxpayers who are only subject to third-party reporting (e.g. those only receiving wage income) may not be able to cheat even if they wanted to. Therefore for this group the probability of audit is a poor proxy for the probability of detection, as the authorities could easily uncover any irregularity simply by matching information reports to tax returns. Consistently, Kleven et al. (2011) find the audit letters to be largely ineffective amongst this group. Using audit data, they also show that while evasion in the third-party reporting environment is close to zero, it is instead substantial in the case of self-reporting. Interestingly this finding holds even considering individuals who earn both types of income, as they fully comply for the portion of income that is third-party reported while they tend to evade on self-reported income. This finding suggests that it is information reporting, rather than an emotional or social aversion to cheat that drives compliance.

Pomeranz (2013) conducted a field experiment to look further into information as a key determinant of compliance in Chile, in close collaboration with the Chilean tax authority. The study focussed on VAT, which is often praised for its built-in incentive structure, triggered by the verifiable paper trail it generates. If the self-enforcement properties of the VAT are effective, taxpayers' response to increased probability of audit would be very low due to the preventive deterrence effect of the paper trail. To evaluate the validity of this hypothesis, transactions that are covered by the paper trail (those between firms) are compared to those that are not (those with final consumers). A first experiment involved sending letters to over 100,000 randomly selected firms, indicating increased audit probabilities. Consistent with the hypothesis, the letters generated a strong increase in tax payments that is largely driven by sales to final consumers, rather than by transactions between firms. Moreover small firms appear to respond more, perhaps because they tend to have higher shares of retail sales. A second 'spillover' experiment looked particularly into the mechanism of the paper trail. In this case half the sample received a pre-announcement of an upcoming audit. The full sample was then audited and the information was made available to the authors. In addition data was collected for the trading partners of the audited firms, before the audit. In presence of the paper trail, one should observe a spillover effect of the audit on the suppliers of the audited firms, which will increase compliance as a result. Finding empirical support for this mechanism, this study argues that the effectiveness of the paper trail gets multiplied thanks to spillover effects. The paper trail allows the tax authority to easily identify discrepancies in reporting between the supplier and the audited firm by cross-checking records, therefore encouraging suppliers to also comply more.

Carrillo et al. (2014) show a fundamental limit of the role of third-party reporting, particularly in countries where information and enforcement constraints are high. The study is based on a natural experiment in Ecuador and focuses on corporate income taxes. While third-party information has been available in Ecuador for some time, it is only recently that the tax authority has started using it to increase compliance. In 2011 and 2012, the tax administration notified about 8,000 firms of discrepancies between their self-reported revenue and third-party information. Because the notification concerned tax returns for previous years, the response is not due to changes in actual behaviour but only to changes in reporting. Therefore, even without a control group, the reported values before the notification provide a counterfactual for the true, post-notification reports, therefore allowing the identification of causal effects. Third-party reports show higher taxable income than the self-reported amount in 24 per cent of the cases. Interestingly, under-reporting also seems to occur on firms' expenditure, although third-party information on costs is much more limited than for revenues. This experiment presents two important results. Firstly, a substantial proportion of firms fail to file an amendment even after receiving the notification. Under weak credible enforcement, firms still under-report revenues because they do not believe the revenue authority has the ability to punish them even when

it has detected under-reporting. Increasing credibility would require the tax administration to follow up on the notifications about discrepancies with actual audits. Failing that, taxpayers do not seem to fully respond to the notifications. However audits are costly and many countries simply do not have the capacity to follow up every case where evasion is suspected. Therefore when traditional enforcement mechanisms are weak, third-party information may not be effective.

Secondly, amongst firms that do file an amendment, reported income increases by large amounts. When firms are notified the exact amount in the third-party reports they tend to amend for larger amounts (93 per cent of the discrepancy) rather than when they do not know the specific amount (36 per cent on average). However, for every dollar of revenue adjustment, costs tend to increase by 96 cents. The line item of increased costs most frequently adjusted is ‘other administrative costs’, which is particularly difficult to verify because of its miscellaneous nature. By amending both revenue and costs, post-amendment profits remain similar and therefore increases in tax payments are limited. In other words, firms seem to adjust less monitored margins (i.e. costs) as a response to increased monitoring on income. This result illustrates how, in presence of incomplete information on all relevant dimensions of compliance, third-party reporting may not have a large impact on revenue mobilisation. Importantly, the same would happen when the information is available but taxpayers know the tax administration cannot use it, typically because of administrative constraints.

While administrative constraints are certainly higher in low- and middle-income countries, Slemrod et al. (2015) show that similar behaviour as documented by Carrillo et al. (2014) occurs in the USA too. Since 2011, firms that process electronic payments (e.g. credit cards) are requested to report this information to the Internal Revenue Service (IRS). These reports are then used by IRS to cross-verify the income declared by small businesses. Although a clear effect of these information reports is difficult to identify, they seem to result in higher declared income particularly for those firms that did not file in the previous year. Firms bunch at the exact amount of the information reports, either because they fear that reporting less would trigger an audit or because that amount accurately reflects true income. However many businesses report equal amounts for income and costs, resulting in zero profits. This study reports some evidence that the new information reports made firms more likely to do so. Therefore, as in the case of Ecuador, the potential benefits from increased declared revenues were partly offset by increased costs. Also in this case the adjustment occurred mostly through items that are difficult to verify, such as ‘other expenses’. This study shows that using third-party reporting to improve compliance of businesses may be challenging in all contexts. Contrary to wage income, business income involves complex reporting requirements that offer more opportunities to minimise tax payments through less verifiable components.

Despite these limitations, there is increasing evidence that a more effective use of information and IT can be beneficial for revenue administrations in developing countries. For example, two recent non-experimental studies looked at the effectiveness of electronic billing machines (EBM) in Africa using data from VAT tax returns (Ali et al., 2015; Eissa and Zeitlin, 2014). They found that both in Rwanda and in Ethiopia, firms substantially increased VAT payments after EBM adoption, although the size of the increase varies across firms’ type, size, and sectors. However it is still unclear whether these increased payments have been offset by other adjustments.

## **2.6 Tax collectors and their incentives**

In addition to problems of capacity and enforcement, tax collections are likely to be constrained by the presence of widespread corruption in tax administrations, particularly in low and middle income countries. Particularly in countries where the use of IT systems and third-party reporting is still limited, contact with tax collectors can be an important factor in explaining compliance. Tax administrations may offer incentives to collectors to minimise the leakage of revenue into corruption. The incentive may be, for example, to allow collectors to cash a share of tax revenues,

thus increasing their overall salary. Particularly in low- and middle-income countries where wages are relatively low, this monetary incentive is particularly salient. However, an adverse consequence of such incentives may be to encourage collectors to use coercion and aggressive methods to raise tax. There is evidence that shows that this may have the counterproductive effect of generating tax resistance and confrontation between taxpayers and the authorities, thus eventually decreasing compliance (Fjeldstad and Semboja, 2001). Although ‘tax farming’ may be effective in increasing revenue in the short run, it also bears some important political costs that may be detrimental in the longer run.

A recent field tax experiment investigates the effect of incentives for tax collectors both on revenues and on taxpayers’ satisfaction with the practices of the tax department. Khan et al. (2014) focus on property taxes in Punjab, Pakistan, and use both tax returns data and a satisfaction survey. The treatments consist of three incentive schemes for collectors. The first one is a pure revenue incentive, where they receive a percentage of any increase beyond a benchmark level of revenue. This incentive is relatively large, as it can almost double their base salary. The second incentive adds to the first one non-revenue outcomes, by adjusting the reward according to measures of taxpayer satisfaction and the accuracy of tax assessments. The third scheme aims to take into account the complexities of tax collection, which may not be easily captured by objective measures. To do so, it is based on subjective assessments of managers who distribute a fixed amount of bonus to a pool of good collectors.

The incentives are expected to interact with existing bribery in two possible ways: they might result either 1) in higher tax payments with a lower level of bribes, or 2) in a constant level of tax payments but higher bribes, as collectors now want to be compensated for a higher foregone gain (the incentive). The results show that on average the treatments are effective in increasing revenue, while taxpayer satisfaction does not change significantly. The most effective treatment is the pure revenue incentive, while the second one has a smaller effect and the third no effect. The additional revenue seem to come primarily from an increase in the tax base, through new assessments and re-assessments, and from the recovery of arrears. Most of the increases come from high-value properties, suggesting that collectors may be maximising the amount of revenue from a few large taxpayers. This would also explain why the perception of corruption and satisfaction does not decrease significantly on average. Overall, this study suggests that incentives can provide an effective way to increase revenue. Although they seem to do so without much political cost, it is still unclear whether such costs may occur in a longer time frame than the one in the study.

## **2.7 Complexity and compliance costs**

Tax systems are typically complex and it is often difficult for taxpayers to fully understand them in detail. This can result both in misperceptions about the system and in compliance costs, due to the time and resources dedicated to tax filing. Therefore, complexity and compliance costs may be an important element in explaining mis-reporting and non-compliance, in addition to plain tax evasion. Compliance costs are particularly relevant for small taxpayers and businesses, as well as first-time filers. There is evidence that small businesses face a cost of up to 15 per cent of their turnover, while this percentage is lower for larger businesses (Coolidge, 2012). Moreover, taxpayers in low- and middle-income countries are particularly affected, due to lower literacy and less availability of trained accountants. As a result, they may file only partially, not file at all, or not do so in time, therefore being partially or fully non-compliant. Modern revenue authorities are making greater efforts to simplify their procedures and to act as tax educators. However, complexity and compliance costs remain real, practical constraints in many countries.

A recent natural experiment used administrative data to show that taxpayers are often confused by complex tax systems (Feldman et al., forthcoming). This study, conducted in the United States, analyses labour responses to the loss of Child Tax Credit when a household’s child turns seventeen. The loss of Child Tax Credit results in a lump sum increase in tax

liabilities but not in a change of the marginal tax rate. However the authors show that confused taxpayers respond to this change by interpreting it as an increase in their marginal tax rate, thus documenting a substantial degree of misperception of the tax system.

In addition to misperceptions, taxpayers may actually lose out because of the complexity of the system and the compliance costs associated with it. Benzarti (2014) uses administrative data from the IRS to show that taxpayers forego large amounts of money because of the compliance cost associated with reporting deductions. Taxpayers in the U.S. have two choices: they can itemise all deductions or claim a standard amount. Itemising results in a monetary benefit when the sum of all itemised deductions is larger than the standard one. If tax filing is costly, rational taxpayers should only itemise when the benefits are higher than the costs of doing so. This is mostly a time cost, due to two elements: record keeping and actual filing. The study reveals that the average taxpayer foregoes \$617 that could have been saved by itemising deductions. However, there is large heterogeneity amongst taxpayers. Richer households lose more than poorer ones through this mechanism, perhaps because they value their time more (as a result of wage differences). These compliance costs are higher than previously documented, either because of actual compliance costs or a psychological bias against tax filing. Similarly, Gillitzer and Skov (2013) report a large amount of under-reporting of charitable contributions, which could be deducted, with a resulting loss for taxpayers. The natural experiment in this case comes from the introduction of third-party reports for donations by the Danish revenue administration. Since 2008, these reports were used to pre-populate tax returns. This measure resulted in doubling the number of tax deductions claimed, and a 15 per cent increase in the value of these claims. This shows that taxpayers were foregoing tax benefits, as there is no evidence of an actual change in charitable giving around the time of the reform.

## 2.8 Learning and long-term perspective

Some of the letter treatments reviewed so far, and particularly those related to deterrence (section 2.1), have large effects on compliance. However, is that effect sustainable or is it a simple nudge? Do taxpayers learn from these treatments? In a way, these questions are related to the issue of credibility of enforcement discussed by Carrillo et al. (2014). Taxpayers may react to a threat of audit, but they may go back to their previous behaviour if no action follows that threat. Field tax experiments typically evaluate only the immediate effects of treatments, since obviously it takes time to observe longer-run effects. However, for policy purposes, it is important to know whether the same measure can be repeated every year with similar results; whether taxpayers will learn from such interventions and therefore do not need regular communications; or perhaps if such interventions should be used only for new taxpayers. All these questions bring into the discussion a longer-term perspective and considerations on sustainability.

A recent paper explored these issues by looking at take-up of the Earned Income Tax Credit (EITC) in the U.S. (Manoli and Turner, 2014). Naturally, the paper focuses on a country with a longer history of tax experiments, therefore building on previous work. In order to increase take-up of the EITC, the IRS sends out notices to taxpayers who seem eligible but who have not claimed the tax credit. The authors are particularly interested in understanding whether these notices have purely short-term nudge effects, or if they also produce longer term learning effects. Firstly, a descriptive analysis of tax returns data shows that information and knowledge of the benefit schedule is indeed linked to low take-up. The study is then based on three experimental components. The first one exploits a computer glitch that resulted in the failure to notify a number of taxpayers who should have been notified. By comparing the group who received the notice and those who did not (but should have), the authors show large nudge effect of the notices on take-up, with a substantial gain of about 4 per cent of labor income as a result of the credit. However such effect gradually disappears, decreasing drastically in the year following the notice and disappearing completely three years later. The second component largely confirms these results by using data from EITC-eligible taxpayers in California in 2009. It looks at

different informational treatments, rather than just the difference between receiving a notice or not. The sample is composed of taxpayers who had already received a notice previously but did not respond. The simplest and most salient notice has the largest nudge effect, as can be expected. However in the years after the treatment the differences between treatment notices disappear, as none of them induces meaningful learning effects. The third component looks at taxpayers who did not claim their tax credit and who were selected for a random audit. Following the audit, they were found to be eligible for the EITC. This can be seen as a more heavy-handed treatment than notices, and may be expected to have a larger learning effect for taxpayers. Although learning effects appear to be larger than those of the notices, they still seem to fade away two years after the audit. While the smaller sample size in this component calls for some caution, the results are in line with the previous ones. Generally, they suggest that regular notices may be needed to increase take-up, as it seems very difficult to generate learning effects about the tax code.

Importantly, learning does not happen only through the revenue administration or through official measures more generally. A recent natural experiment (Alstadsaeter et al., 2014) shows that family networks are an important channel for transmitting information about the tax system. In particular the authors look at a well-defined, legal tax shelter in Norway. Using tax returns data, they exploit a discontinuity in the availability of the tax shelter that allows them to compare similar individuals who have or do not have access to it. By doing this, they show that similar individuals, with similar social networks, have different taxpaying behaviour according to whether or not they are part of a network including an eligible individual. These results suggest that taxpayers affect each other along social networks, which are therefore an important element in explaining tax avoidance and evasion.

### 3 Conclusions

The TE literature has developed greatly in the last decade, from lab studies to large-scale field experiments. The latter have a few important advantages over the former: they are fully embedded in real taxpaying situations, thus being more suitable as a basis for policy advice; they benefit from more statistical power as samples are much larger than what the lab previously allowed; and they provide further insights into the heterogeneity of taxpayers' responses to various determinants of compliance. While it is clear that deterrence messages affect particularly taxpayers with self-reported income (Slemrod et al., 2001; Kleven et al., 2011), other factors such as the level of compliance in the community (Fellner et al., 2013), existing beliefs on enforcement and compliance (Del Carpio, 2013; Carrillo et al., 2014), and taxpayers' motivations to comply (Dwenger et al., 2014) all matter to explain the results. With field experiments, this literature is moving away from purely theoretical hypotheses to more practical policy questions, around enforcement constraints and third party reports. Therefore the label 'whispering in the ears of princes' seems particularly appropriate to underline this higher policy relevance.

Although field experiments are a relatively new and narrow strand of literature, they still present important differences in methods. The following list names the most notable ones:

- Natural experiments exploit existing circumstances, while the typical letter experiment is set up by researchers in collaboration with the tax administration.
- Sample sizes differ greatly, from a few hundred (Ariel, 2012) to several thousands (Pomeranz, 2013; Dwenger et al., 2014).
- Compliance is defined in various ways: in terms of outstanding payments (Hallsworth et al., 2014), as new registrations (Fellner et al., 2013), or as increases in reported income (Slemrod et al., 2001).

- Most recent field TE use a neutral ‘control’ letter to isolate the effect of simply receiving any letter from the effect resulting from the specific content of the letter. This has proven particularly important as indeed the control letter has been found to be responsible for a large portion of the results in some cases (Fellner et al., 2013).
- Although the treatment messages are inspired by similar theoretical hypotheses, they are translated in practice with various degrees of complexity. While Hallsworth et al. (2014) and Castro and Scartascini (2013) only make small variations to a standard message, Slemrod et al. (2001) or Torgler (2004) changed entire paragraphs.
- The treatment message can be embedded in the tax bill (Castro and Scartascini, 2013) or, as in most cases, be sent as an independent letter.
- The studies reviewed look at different tax types, including VAT (Pomeranz, 2013), personal income taxes (Bott et al., 2014), corporate taxes (Ariel, 2012), and property taxes (Castro and Scartascini, 2013). Although compliance *issues* are common across tax types, the compliance and enforcement *levels* differ, with possible implications on the results (see section 2.2).

Importantly, these differences are likely to matter for results – making rigorous and context-appropriate research designs crucial to their credibility. In addition, more nuanced results can be shown in studies that analyse sub-groups separately. The absence of any effect at the aggregate level may hide heterogeneous responses across groups (Fellner et al., 2013).

Despite these differences across individual studies, this literature points to a few common results on what matters for compliance. Deterrence factors seem to be the most effective method to increase compliance in the short run, but its effect may not be sustainable over time (Manoli and Turner, 2014) and in presence of weak follow-up enforcement (Carrillo et al., 2014). The evidence on moral appeals and social norms is less conclusive. However, these factors, as well as positive rewards, can be effective strategies to increase compliance in some cases (Hallsworth et al., 2014; Dwenger et al., 2014). Factors other than deterrence are particularly important for two reasons. The first one, mentioned before, has to do with the sustainability of threats in weak enforcement environments. Increasing actual enforcement is very costly and in most cases it is not optimal for tax administration to check all taxpayers who can be suspected of evasion. Therefore, it makes sense to explore compliance strategies that are more cost effective. Secondly, an approach based only on deterrence could fuel a confrontational relationship between citizens and the state, which in turn may increase taxpayers’ resistance and lower compliance. Last but not least, there are some important practical issues that act as obstacles to compliance, such as the complexity of tax systems and compliance costs.

This review has made a particular effort to include evidence from all regions of the world, in a literature that is largely dominated by Europe and the U.S. The increasing availability of tax returns data from low- and middle-income countries is allowing researchers to draw a more global picture than in the past. Generally, the types of issues that emerge as key determinants of compliance are very similar across countries. However, there are important context-specific factors that affect both results and policy implications, such as beliefs on enforcement and motivations to comply. A case in point are moral appeals including information on existing compliance levels: while they may be effective in countries with low evasion, they can backfire where evasion is widespread (see section 2.2). In addition, while in rich countries threats of audits are likely to be more credible, they may not be taken seriously in contexts where administrative constraints are well known. Again, heterogeneity matters as large taxpayers are still less likely to be affected by such threats in any country. In conclusion, it is still difficult to identify systematic differences across countries from this literature. Cross-country comparisons are particularly hard because any difference in results may be due to variations in context, design, or both.

Looking forward, this literature certainly promises to provide further useful insights on compliance both for the research community and for policymakers. There are at least three directions for further research emerging from this review. Firstly, there is still a lack of evidence from Asia, Latin America and, particularly, Africa. A few ongoing studies are starting to fill this gap, and more work needs to be done to build a truly global body of knowledge. Secondly, this literature would benefit from more standardised methods across studies. Although research design needs to take the context into account, researchers should make greater efforts to achieve comparability across studies. Finally, there are a number of practical issues that are known to matter for compliance but lack sufficient rigorous evidence. For example, we know that complexity and compliance costs are important constraints, particularly in low-income countries. However, there is still relatively little experimental evidence, mostly from natural experiments, on what tax administrations can do to tackle them. Moreover, the issue of informality remains largely untouched in this literature. Although this is largely due to the use of administrative data, which naturally excludes the informal sector, it may be possible in some cases to look at new registrations by matching tax data with other sources of information - for example, from banks. Finally, further studies should keep making efforts to disaggregate effects along dimensions of heterogeneity that can offer a more nuanced picture of compliance.

## References

- Ajzen, I. and M. Fishbein (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological bulletin* 84(5), 888–918.
- Ali, M., O.-H. Fjeldstad, and I. H. Sjursen (2015). To pay or not to pay? citizens' attitudes towards taxation in Kenya, Tanzania, Uganda and South Africa. *World Development*.
- Ali, M., A. Shifa, A. Shimeles, and F. Woldeyes (2015). Information technology and fiscal capacity in a developing country: Evidence from Ethiopia. Working paper 31, International Centre for Taxation and Development.
- Allingham, M. G. and A. Sandmo (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics* 1(3-4), 323–338.
- Alm, J. and B. Jackson (1993). Fiscal exchange, collective decision institutions and tax compliance. *Journal of economic Behavior and Organisation* 285(303).
- Alm, J., B. Jackson, and M. McKee (1992). Deterrence and beyond: towards a kinder, gentler IRS. In J. Slemrod (Ed.), *Why people pay taxes*. University of Michigan Press.
- Alm, J., G. H. McClelland, and W. D. Schulze (1992). Why do people pay taxes? *Journal of Public Economics* 21(38).
- Alstadsaeter, A., W. Kopczuk, and K. Telle (2014). Social networks and tax avoidance: Evidence from a well-defined norwegian tax shelter. mimeo.
- Andreoni, J., B. Erard, and J. Feinstein (1998, June). Tax Compliance. *Journal of Economic Literature* 36(2), 818–860.
- Angrist, J. D. and J.-S. Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives* 24(2), 3–30.
- Ariel, B. (2012). Deterrence and moral persuasion effects on corporate tax compliance: findings from a randomized controlled trial. *Criminology* 50(1), 27–69.
- Basu, K. (2013). The method of randomisation, economic policy, and reasoned intuition. Policy Research Working Paper 6722, World Bank.
- Beck, P. J., J. S. Davis, and W.-O. Jung (1991). Experimental evidence on taxpayer reporting under uncertainty. *The Accounting review* 66(3).
- Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76, 169.
- Becker, W., H.-J. Buchner, and S. Sleeking (1987). The impact of public transfer expenditures on tax evasion : An experimental approach. *Journal of Public Economics* 34(2), 243–252.
- Benzarti, Y. (2014). How taxing is tax filing? Leaving money on the table because of compliance costs. Working paper, University fo California, Berkeley.
- Blumenthal, M., C. Christian, and J. Slemrod (2001). Do normative appeals affect tax compliance? Evidence from a controlled experiment in minnesota. *National Tax Journal* 54(1), 125–138.
- Bosco, L. and L. Mittone (1997). Tax Evasion and Moral Constraints: Some Experimental Evidence. *Kyklos* 50(3), 297–324.

- Bott, K., A. W. Cappelen, and E. Ø. Sørensen (2014). You've got mail: A randomised field experiment on tax evasion. Discussion paper 26/2014, Department of Economics, Norwegian School of Economics.
- Calvet Christian, R. and J. Alm (2014). Empathy, sympathy, and tax compliance. *Journal of Economic Psychology* 40(C), 62–82.
- Carrillo, P., D. Pomeranz, and M. Singhal (2014). Dodging the taxman: Firm misreporting and limits to tax enforcement. NBER Working Paper 20624, NBER.
- Castro, L. and C. Scartascini (2013). Tax compliance and enforcement in the Pampas: evidence from a field experiment. *Journal of Economic Behavior and Organization* 116, 65–82.
- Coolidge, J. (2012). Findings of tax compliance cost surveys in developing countries. *eJournal of Tax Research* 10(2), 250–287.
- Coricelli, G., M. Joffily, C. Montmarquette, and M. Villeval (2010). Cheating, emotions, and rationality: an experiment on tax evasion. *Experimental Economics* 13(2), 226–247.
- Coricelli, G., E. Rusconi, and M. C. Villeval (2014). Tax evasion and emotions: An empirical test of re-integrative shaming theory. *Journal of Economic Psychology* 40(C), 49–61.
- Cummings, R. G., J. Martinez-Vazquez, M. McKee, and B. Torgler (2009, June). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior & Organization* 70(3), 447–457.
- Deaton, A. (2009). Instruments of development: randomization in the tropics, and the search for the elusive keys to economic development. Working Paper 14690, NBER.
- Del Carpio, L. (2013). Are the neighbors cheating? Evidence from a social norm experiment on property taxes in peru. Mimeo, Princeton University.
- Dunning, T., F. Monestier, R. Piñeiro, F. Rosenblatt, and G. Tuñón (2014). Positive vs. negative incentives for compliance: Evaluating a randomized tax holiday. mimeo.
- Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2014). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. mimeo.
- Eissa, N. and A. Zeitlin (2014). Using mobile technologies to increase VAT compliance in Rwanda. mimeo, McCourt School of Public Policy, Georgetown University.
- Feldman, N., P. Katuščák, and L. Kawano (forthcoming). Taxpayer confusion: Evidence from the child tax credit. *American Economic Review*.
- Fellner, G., R. Sausgruber, and C. Traxler (2013). Testing enforcement strategies in the field: Legal threat, moral appeal and social information. *Journal of the European Economic Association* 11(3).
- Fjeldstad, O.-H., C. Schulz-Herzenberg, and I. H. Sjursen (2012). People's views of taxation in africa: A review of research on determinants of tax compliance. Working Paper 8, International Centre for Taxation and Development.
- Fjeldstad, O.-H. and J. Semboja (2001). Why people pay taxes: The case of the development levy in tanzania. *World Development* 29(2059-2074).
- Friedland, N., S. Maital, and A. Rutenberg (1978). A simulation study of income tax evasion. *Journal of Public Economics* 10(1), 107–116.

- Gadenne, L. (2013). Tax me, but spend wisely: sources of public finance and government accountability. Mimeo, University College London and Institute for Fiscal Studies.
- Gillitzer, C. and P. Skov (2013). Evidence on unclaimed charitable contributions from the introduction of third-party information reporting in Denmark. EPRU Working Paper Series 2013-04, Economic Policy Research Unit (EPRU), University of Copenhagen. Department of Economics.
- Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2014). The behavioralist as tax collector: using natural field experiments to enhance tax compliance. Working paper 20007, NBER.
- Khan, A. Q., A. I. Khwaja, and B. A. Olken (2014). Tax farming redux: Experimental evidence on performance pay for tax collectors. NBER Working Paper 20627, NBER.
- Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica* 79(3), 651–692.
- Lamberton, C., J.-E. De Neve, and M. I. Norton (2014). Eliciting taxpayer preferences increases tax compliance. Working Paper 14-106, Harvard Business School.
- Luttmer, E. and M. Singhal (2014). Tax morale. *Journal of Economic Perspectives* 28(4), 149–68.
- Manoli, D. S. and N. Turner (2014). Nudges and learning: evidence from informational interventions for low-income taxpayers. Working paper 20718, NBER.
- Martin, L. (2014). Taxation, loss aversion, and accountability: Theory and experimental evidence for taxation’s effect on citizen behavior. mimeo, Yale University.
- Mittone, L. (1997). Subjective versus objective probability: results from seven experiments on fiscal evasion. CEEL Working Papers 4, CEEL.
- Mossin, J. (1968). Taxation and risk-taking: An expected utility approach. *Economica* 35(137), 74 – 82.
- Myles, G. D. and R. A. Naylor (1996, April). A model of tax evasion with group conformity and social customs. *European Journal of Political Economy* 12(1), 49–66.
- Paler, L. (2013). Keeping the public purse: An experiment in windfalls, taxes, and the incentives to restrain government. *American Political Science Review*.
- Pomeranz, D. (2013). No taxation without information: deterrence and self-enforcement in the value added tax. NBER Working Paper 19199, NBER.
- Roth, A. E. (1995). Introduction to experimental economics. In J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*. Princeton University Press.
- Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives* 21(1), 25–48.
- Slemrod, J., M. Blumenthal, and C. Christian (2001). Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics* 79(3), 455–483.
- Slemrod, J., B. Collins, J. Hoopes, D. Reck, and M. Sebastiani (2015). Does credit-card information reporting improve small-business tax compliance? NBER Working Paper 21412, NBER.

- Slemrod, J. and C. Weber (2012). Evidence of the invisible: toward a credibility revolution in the empirical analysis of tax evasion and the informal economy. *International Tax and Public Finance* 19(1), 25–53.
- Spicer, M. W. and R. E. Hero (1985). Tax evasion and heuristics : A research note. *Journal of Public Economics* 26(2), 263–267.
- Spicer, M. W. and J. E. Thomas (1982). Audit probabilities and the tax evasion decision: An experimental approach. *Journal of Economic Psychology* 2(3), 241–245.
- Torgler, B. (2002). Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments. *Journal of Economic Surveys* 16(5), 657–83.
- Torgler, B. (2003). Beyond Punishment: a tax compliance experiment with taxpayers in Costa Rica. *Revista de Analisis Economico - Economic Analysis Review* 18(1), 27–56.
- Torgler, B. (2004). Moral suasion: An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. *Economics of Governance* 5(3), 235–253.
- Torgler, B., M. Schaffner, and A. Macintyre (2007). Tax Compliance, Tax Morale and Governance Quality. CREMA Working Paper Series 2007-17, Center for Research in Economics, Management and the Arts (CREMA).
- Traxler, C. (2010). Social norms and conditional cooperative taxpayers. *European Journal of Political Economy* 26(1), 89–103.



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