

Does Reciprocity Shape Tax Morale? Evidence from Field-Experimental Variation in Tax Collection Costs

Tobias Cagala*, Johannes Rincke*, Amanda Tuset-Cueva*

October 21, 2019

Abstract

We exploit a randomized field experiment to investigate how administrative efficiency affects tax morale. The context of the experiment is a German church tax that builds on voluntary compliance. We inform taxpayers about a reduction in collection costs and find counteracting effects on compliance, depending on baseline tax morale. High tax morale types reciprocate cost reductions and increase their tax payments. In contrast, low tax morale types free-ride on cost reductions and pay even less. We conclude that reciprocity is a component of tax morale, but less motivated taxpayers are unlikely to reciprocate improvements in administrative efficiency with higher compliance.

JEL codes: C93, D64, H26, H71, K34, Z12

Keywords: Administrative efficiency; tax compliance; tax morale; reciprocity; free-riding; crowding out; field experiment

*Department of Economics, University of Erlangen-Nuremberg. Cagala: tobias.cagala@bundesbank.de; Rincke: johannes.rincke@fau.de; Tuset-Cueva: amanda.tuset.cueva@fau.de. We thank the Evangelical Lutheran Church in Bavaria and especially Johannes Bempohl for partnership on this project. We have benefited greatly from discussions with Matteo Cervelatti, Ulrich Glogowsky, Johannes Hermler, Dominika Langenmayr, Ricardo Perez-Truglia, Sebastian Sieglöcher, Joel Slemrod, and seminar participants at various locations. The paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or its staff. The experiment was approved by the ethics committee of the University of Nuremberg and pre-registered at the AEA RCT Registry (AEARCTR-0001332).

1 Introduction

The literature on the determinants of tax compliance has discussed two main motives leading taxpayers to comply with the tax law. First, as modelled by [Becker \(1968\)](#) and [Allingham and Sandmo \(1972\)](#), taxpayers may respond to deterrence parameters of the tax system, like audit probabilities and penalties, and the ability of the tax authority to collect and process third-party information on true tax liabilities ([Blumenthal *et al.*, 2001](#); [Kleven *et al.*, 2011](#); [Naritomi, 2014](#); [Pomeranz, 2015](#)). Second, the behavior of taxpayers may also reflect tax morale, i.e. non-pecuniary motivations including forms of intrinsic motivation (like a duty to comply, warm glow, or altruism) and reciprocity towards the state or the society as a whole ([Andreoni *et al.*, 1998](#); [Luttmer and Singhal, 2014](#)).¹ Although recent field studies have made considerable progress in measuring and understanding the importance of tax morale ([Slemrod *et al.*, 2001](#); [Fellner *et al.*, 2013](#); [Del Carpio, 2014](#); [Dwenger *et al.*, 2016](#); [Hallsworth *et al.*, 2017](#); [De Neve *et al.*, 2019](#)), our knowledge of how we can use non-pecuniary motivations to increase tax compliance is still rather limited.

This paper adds to the latter strand of literature and studies the role of administrative efficiency for tax morale. Specifically, we test the conjecture of [Luttmer and Singhal \(2014\)](#) that the perceived competence of the government shapes tax compliance via a reciprocal response of taxpayers who view taxes as payments in exchange for services provided by the state. Our evidence comes from a randomized field experiment on compliance responses to a reduction in tax collection costs. The context of the field experiment is a small-stakes income-dependent tax, raised by the local church districts in Germany. Two features of the local church tax stand out. First, tax collection relies exclusively on church members' tax morale. There are no audits, and church members who do not comply with the tax do not face any consequences. Second, we can link administrative data on taxes owed to records of actual local church tax payments. This allows us to exactly measure individual compliance before and after the experimental intervention and to describe the treatment responses for different baseline compliance types.

¹For recent surveys on the literature on tax compliance, see [Slemrod and Weber \(2012\)](#) and [Mascagni \(2018\)](#).

In the experiment, we implemented treatments that manipulate the district-level administrative fixed cost of producing and mailing the tax notifications to church members. Treatment assignment was randomized at the individual level, and taxpayers were informed about the cost reductions in the annual tax notification. The treatments aim to disentangle two effects: a reciprocal response, and a crowd-out of voluntary tax payments. The reciprocity effect captures a *positive* compliance response to the cost reduction as a signal of improved performance by the local administration. The crowding-out effect results from the fact that a lump-sum reduction in administrative costs is financially equivalent to a positive third-party transfer to the local public good. For altruistic taxpayers, whose utility function depends on the amount of the provided public good, the funding sources of this public good are fungible. Therefore, the implicit transfer crowds out individual contributions of altruistic taxpayers, resulting in a *negative* compliance response (Warr, 1982; Roberts, 1984; Andreoni, 1989; Ribar and Wilhelm, 2002).

We disentangle the two effects as follows. First, we identify the reciprocity effect by comparing a treatment that communicates a real reduction of tax collection costs to a treatment that highlights an equivalent shift of administrative costs to a higher level of the church administration. Technically, the cost shifting is achieved by the higher-level administration refunding the church districts for part of the tax collection costs. In both treatments, we hold the size of the implicit transfer to the local public good constant but vary the communicated source of the cost reduction, from a real reduction to a reduction that takes the form of a cost refund. The identification of the reciprocity effect, hence, rests on the assumption that a real reduction in administrative costs is a stronger signal of improved administrative performance compared to the administration receiving a partial cost refund. Second, the crowding-out effect is identified by comparing two treatments that vary only in the share of the tax collection cost that is subject to the refund, i.e., the size of the implicit transfer to the local public good.

Our main result is that administrative efficiency has counteracting effects on compliance, depending on baseline tax morale. Taxpayer types who show high tax morale in the baseline increase their tax payments in a reciprocal response to a true cost re-

duction, but do not show any crowd-out if a shifting of costs to a third party results in an implicit transfer to the public good. In contrast, among types with low baseline tax morale, cost reductions do not trigger any reciprocal response, but crowd out compliance further. We conclude that it is difficult to predict the overall compliance responses to improvements in administrative efficiency. Our findings suggest that the overall response will depend not only on the relative strength of the reciprocity effect and the crowd-out effect, but also on the distribution of tax morale in the population of taxpayers.

Relative to the previous literature, this paper provides two main and novel insights. First, we are among the first providing causal evidence that tax morale is at least partly shaped by reciprocity concerns. The finding implies that by signaling an enhanced administrative performance, tax authorities can, in principle, increase the compliance of at least some taxpayers. Previous literature has documented positive correlations between institutional quality and survey-based measures of tax morale (for a review, see [OECD 2013](#)),² or has produced related, but indirect evidence suggesting that taxpayers reciprocate good governance. For instance, [Timmons and Garfias \(2015\)](#) document that revenues from the municipal property tax in Brazil decrease if audits reveal corruption by local officials, and [Doerrenberg and Peichl \(2018\)](#) use a survey experiment to show that stated levels of tax morale can be manipulated by making salient that tax evasion negatively affects the provision of public services. We advance this literature by isolating a direct causal link between administrative efficiency and tax morale. Importantly, most previous field studies testing for reciprocity in tax morale surveyed by [Luttmer and Singhal \(2014\)](#) failed to detect any systematic link between attitudes towards the state and tax morale.³ As argued by [Luttmer and Singhal \(2014\)](#),

²An interesting example is [Bodea and LeBas \(2016\)](#), who use novel data from urban Nigeria to show that individuals with a positive experience of public service delivery are more likely to express belief in an unconditioned citizen obligation to pay taxes. [Torgler \(2003, 2005\)](#) shows that survey measures of tax morale are positively correlated with trust in public officials and in the legal system, and with institutional features of direct democracy.

³Notable exceptions include [Hallsworth *et al.* \(2017\)](#), who find that moral suasion letters emphasizing the ways in which tax revenue finances public goods lead to more timely tax payments in the United Kingdom, and [Cullen *et al.* \(2015\)](#), who use U.S. county-level data and show that measures of federal income tax evasion are negatively correlated with the degree of voters' political alignment with the president's party. In the laboratory, researchers have consistently found that when tax revenues are spent on public goods subjects value, subjects are more willing to comply ([Alm *et al.*, 1992a,b](#)). Also from the laboratory, [Cagala *et al.* \(2018\)](#) show that an administrator who decides to expropriate part

this could be due to the fact that most of those studies use ‘moral suasion’ treatments which might be too soft to trigger a significant shift in individuals’ attitudes. Instead of relying on moral suasion, our treatments implement real changes in the functioning of the administration taxpayers interact with.

A second insight relates to the strong heterogeneity in treatment responses that we observe in our experiment. While the evidence from our experiment suggests that policies reducing tax collection costs can be effective in raising tax morale, our results also highlight the risk that alluding to individuals’ tax morale may trigger adverse compliance responses among individuals with a weak baseline tax morale. This heterogeneity is complementary to [Dwenger *et al.* \(2016\)](#), who study compliance rewards in the context of the local church tax in Germany and show positive compliance responses of strongly intrinsically motivated taxpayers, whereas weakly motivated taxpayers *decrease* their compliance even further. The fact that we study the role of tax collection costs in a setting without any enforcement relates our work to the literature on the role of charities’ overhead costs for charitable giving. Consistently, this literature has shown that potential donors are averse to financing charities’ overhead cost ([Okten and Weisbrod, 2000](#); [Gneezy *et al.*, 2014](#); [Caviola *et al.*, 2014](#); [Meer, 2014](#)). We demonstrate that, in principle, an increased willingness to contribute in response to lower administrative costs extends to a setting with (legally) compulsory tax payments. However, it is noteworthy that the literature on charitable giving has focused almost exclusively on warm-list individuals (i.e., subjects who have revealed some willingness to donate in the baseline). This has prevented the literature from studying the effect of reduced overhead costs on individuals with a low motivation to give. Bridging the charitable giving and the tax setting, our findings suggest that among cold-list individuals, communicating lower administrative costs may actually crowd out contributions.

The remainder of the paper is organized as follows. Section 2 discusses the institutional background, and Section 3 presents conceptual considerations. Section 4 describes our data sources, sampling, and experimental design. Section 5 discusses our findings, and Section 6 concludes.

of contributions to non-profit organizations sharply diminishes individuals’ contributions.

2 Institutional Background

In Germany, the Catholic and Protestant churches are financed through church taxes. The main source of revenue for the churches is the state church tax, which corresponds to 8 or 9 percent (depending on the state) of church members' income tax liabilities.⁴ Due to an automated process by which state tax authorities provide the church tax office with individual income tax records for all church members, the state church tax is very well enforced. In addition to the state church tax, in some states the churches levy a local church tax. It is smaller in size than the state church tax and collected by the church administration at the level of local church districts.

This study focuses exclusively on the local church tax, collected by the Protestant Church in Bavaria. As can be seen in Table A1 in the appendix, the local church tax is a progressive income tax following a step function with an exemption level of €8,652 in tax year 2016. Incomes exceeding the exemption level are taxed using six brackets, with the annual tax liability increasing from €5 to €100. The tax base includes all types of income, including wage income, business income, and capital income. To collect the annual local church tax, the local church administration mails a tax notification letter (see the appendix) to all resident church members above the age of 18. The tax notification asks church members to self-assess their income and to transfer the local church tax owed to the church's bank account within a month's time. Attached to the tax notification is a bank transfer form, pre-filled with the church's bank account information and the taxpayer's name and tax number.

An important feature characterizing the local church tax is that the church does not enforce tax payments. Although the church (through the tax office handling the state church tax) could, in principle, cross-check local church tax payments against individual income tax records, the church does not audit its members with respect to the local church tax.⁵ The church district administration does not even hand down information on local church tax payments to the parish level. As a consequence, church members who do not pay the local church tax, or underpay, do not face any legal consequences.

⁴The specifics of church finances vary somewhat across the German states and between both churches. The following description captures the case of the Protestant Church in Bavaria.

⁵Interestingly, in one of the districts participating in our field experiment, the church administration did not even keep individual records of local church tax payments prior to our experiment.

Moreover, access to church services is unrelated to the individual's compliance with the tax. The local church tax thus provides us with the rare opportunity to study individual compliance decisions in a context with zero enforcement. One implication is that any compliance we observe must be driven either by misperception, or by tax morale. [Dwenger et al. \(2016\)](#) study the same context and show that making the absence of audits explicit has only a small impact on compliance. This suggests there is little misperception on average, i.e. church members seem to be well aware of the fact that the local church tax is not enforced at all. Therefore, we can use local church tax payments as a measure of taxpayers' tax morale. Moreover, the absence of enforcement makes the local church tax a setting where tax underpayments (evasion) and donations coexist within a single tax instrument. Evasion takes the form of income underreporting in the self-assessment of the local church tax. Donations can be made by overpaying the local church tax relative to the amount owed. Such overpayments are not refunded, but treated like donations reaching the church through other channels.⁶

3 Conceptual Considerations

The main purpose of this paper is to provide evidence on reciprocity as a component of tax morale. We use the term 'reciprocity' because it is frequently used in the tax compliance literature. However, it is important to note that in our context, the term describes a motivation that differs conceptually from sequential reciprocity in the sense of [Dufwenberg and Kirchsteiger \(2004\)](#). Following [Luttmer and Singhal \(2014\)](#), we use the term reciprocity to describe a motivation to pay taxes that depends on the individual's relationship with the state. In the context of this paper, the role of the state is taken by the Protestant Church in Germany with its own administration and its tax and expenditure policies set independently from the government. Importantly, reciprocity as a component of tax morale captures motivations to pay taxes that go beyond direct personal benefits from tax payments, such as church services that are contingent on a

⁶Charitable donations and church tax payments are deductible when filing for the German income tax. Hence, the price when paying the local church tax is the same as the price of giving through donations (to the church or to other charities).

payment. In the context of church taxes, reciprocity captures the idea that individuals may view their fiscal relation to the church as a social contract: church tax payments are made in exchange for public goods provided by the church. Individuals holding this view will naturally consider the performance of the church and its administration in delivering services to its constituency as a yardstick to judge the legitimacy of their own tax payment. Holding the individual's tax payment constant, an improved performance of the church administration will then lead to a higher perceived legitimacy of church taxes, whereas a decline will decrease the legitimacy of the tax burden. As a consequence, in the presence of reciprocity concerns, an improved performance of the church administration should result in higher tax morale among church members.

In our field experiment, we test the reciprocity hypothesis by randomly varying the church's communication regarding the performance of the local church administration. The change in performance that is communicated takes the form of a reduction of the administrative cost to collect the local church tax. All other things equal, a reduced collection cost leads to tax payments resulting in higher net tax revenues and, therefore, more church-provided public goods and services. *Ceteris paribus*, we expect higher tax morale, i.e., higher tax payments, from church members who are motivated by reciprocity concerns.

It is important to note, however, that measures contributing to a better performance of a public administration might also trigger a different type of response that may counteract the positive compliance effect that is due to taxpayers' reciprocity concerns. Conceptually, this response has been extensively discussed in the public-goods crowding-out literature that goes back to, e.g., [Warr \(1982\)](#), [Roberts \(1984\)](#) and [Andreoni \(1989\)](#). The workhorse model in this literature considers individuals who contribute to a public good via two channels, lump-sum taxes and gifts. Depending on whether individuals' utility depends on the provided public good (altruistic preferences) and/or for their own gift (warm-glow preferences), the model generates predictions about how individuals optimally respond to changes in government contributions to the public good: While government contributions completely crowd out private contributions if preferences are purely altruistic, the model predicts zero crowd out in the case of pure warm glow.

To see the relevance of a possible crowding out for our context, let us consider a reduction of the cost per taxpayer to collect the local church tax. Importantly, the cost reduction is financially equivalent to a lump-sum contribution by the church (or any other third party) to the public good. If at least some taxpayers have altruistic preferences, the model discussed before predicts that the implicit transfer crowds out voluntary contributions the respective taxpayers would have made in the absence of the cost reduction. In contrast, if taxpayers are motivated only by warm glow, the cost reduction would not imply any crowd-out.

The conceptual considerations imply that our strategy to identify the reciprocal response of taxpayers to a reduction in administrative costs must carefully consider the possibility of a crowding-out effect that counteracts the reciprocity response. In the following section, we describe how the design of our treatments takes care of crowding out in order to isolate the pure reciprocity effect.

4 Experimental Design, Data, and Methods

4.1 Experimental Design

Setting of the field experiment The field experiment was implemented in two urban districts of the Protestant Church in Bavaria, comprising a total of 14 parishes.⁷ The districts are located in cities with a population of about 60,000 and 140,000, respectively. The share of Protestants in the two cities is 17.8 percent and 13.3 percent, respectively. In Bavaria as a whole, the share of Protestants is 20.9 percent.⁸

The field experiment took place in 2016. At the beginning of June, the church administration in the two districts mailed tax notifications for the local church tax to about 35,000 resident members of the Protestant Church. Both the layout and the content of the notifications was identical across both districts, and all notifications were mailed on the same day. In collaboration with the church, we manipulated the content of the tax notifications. Using a stratified randomization scheme,⁹ we assigned

⁷Dwenger *et al.* (2016) implemented their field experiment in a different region in Bavaria.

⁸The largest congregation in Bavaria is the Catholic Church, comprising 55.3 percent of the population (data from the German Census 2011).

⁹Strata were defined using age quartiles, gender, dummies for the taxpayers' tax bracket according

each taxpayer to one of four treatment groups. In the following, we describe the treatments and how they relate to the identification of the reciprocity and crowding out effects.

Control group The first treatment we implemented was a *control group*. The sole purpose of this treatment is to test the persistence of compliance behavior between baseline and treatment year absent any treatment. In the *control group*, the tax notification (see the appendix) was very similar to the letter that was mailed in previous years.¹⁰ The control group notification comprises three short paragraphs. It informs taxpayers that, based on the state law regulating the church tax, the church district raises a local church tax. It is made clear that the local church tax forms part of the general church tax, and that paying the tax is compulsory. Taxpayers are also informed that tax revenues are used to finance local public services provided by the district's parishes. The notification displays the tax schedule on the front page, and asks the recipient to settle the tax liability by bank transfer based on a self-assessment of income.

The control group letter does not mention administrative costs. Comparisons between the treatments described below and the control group are therefore confounded by differences in the salience of administrative costs. This is the reason why we do not use the *control group* for treatment comparisons.

Treatments used to identify reciprocity The tax notifications in the treatment groups are based on the control group letter. After the first paragraph, the treatment notifications include a short note printed in bold type, containing information regarding a change in the tax collection costs borne by the local church administration.

The identification of the reciprocity effect rests on comparing two treatments, the *30% cost reduction* treatment and the *30% cost refund* treatment. In the *30% cost reduction* treatment, the additional paragraph reads: *Please note: In collaboration with various service providers, this year, we were able to reduce the administrative cost asso-*

to the local church tax schedule (including a dummy for taxpayers for whom we do not observe the true tax liability), dummies for baseline payment behavior, and parish dummies.

¹⁰While the layout was identical across both districts in previous years, the content varied slightly. Hence, we had to adjust both notifications only marginally to make them identical across districts.

ciated with the local church tax notification by 30 percent. This means that your local church tax payment is now even more effective.

The additional paragraph in the 30% cost refund treatment is similar, but communicates a shift of administrative costs from the church district to the state church rather than a real cost reduction. It reads as follows: *Please note: This year, we get a refund of 30 percent of the administrative cost associated with the local church tax notification from the state church. This means that your local church tax payment is now even more effective.*

Importantly, the only difference between the 30% cost reduction and the 30% cost refund treatment is the communicated reason for the change in costs borne by the district. The underlying assumptions to interpret the resulting treatment effect as evidence of reciprocity is that taxpayers ignore the change in the fiscal situation of the state church that results from the cost shift, and that taxpayers perceive a real reduction of costs as a stronger signal of the local administration's performance compared to a pure shifting of costs to the state church as a third party. Because both treatments communicate a drop in local administrative expenses by 30%, the implicit transfer to the local public good implied by the communicated changes is held constant. In other words, from a local perspective the monetary consequences of what is communicated in the tax notifications are equal for both treatments.¹¹ In terms of the theoretical discussion in Section 3, our design makes sure that the reciprocity effect is not confounded by a crowding-out effect.

Treatments used to identify crowding out To separately identify the crowd-out effect, we need an additional treatment that varies the extent of the cost refund benefiting the local church district. To maximize our chances to identify the crowd-out, we chose this treatment to be a 100% cost refund treatment. Like the 30% cost refund treatment, it communicates a shift of administrative costs to the state church. The respective paragraph reads: *Please note: This year, we get a refund for the administrative cost associated with the local church tax notification from the state church. This means that your local church tax payment is devoted to its purpose without any deduction.*

¹¹We make the additional assumption here that taxpayers' perceptions about possible adjustments in grants from the state level administration to the districts are the same between treatments.

Hence, your local church tax is now even more effective.

To identify the crowding out effect, we compare the *100% cost refund* treatment to the *30% cost refund* treatment. Because the only difference between those treatments is the size of the implicit transfer associated with the shifting of costs to the state-level administration, this comparison characterizes the crowd out that is associated with increasing the transfer to the public good from 30% to 100% of the fixed administrative cost.

Implementation We made sure that the changes in administrative costs borne by the district administrations were implemented as communicated in the tax notifications. This was achieved as follows. In preparing the notifications, the district administrations rely on in-house service providers operating within the state-level administration of the Protestant Church in Bavaria. The service providers prepare and handle the records defining the population of taxpayers (also containing address information for the mailing of tax notifications), produce the tax notifications (including printing, franking, and inserting), and deliver them to a mail service company for mailing. The church's in-house service providers bill the local church districts for their services. In collaboration with the church's state-level administration and the participating church districts, we arranged for a partial reimbursement of the districts for the billed cost through state-level funds of the Protestant Church. For each district, the overall reimbursement reflected exactly the distribution of treatments and the changes in locally borne costs communicated in the respective notifications, i.e. each individual notification carried a change of administrative cost borne by the district administration in accordance with the content of the treatment paragraph. This reimbursement procedure was implemented in the refund treatments as well as in the cost reduction treatment. In contrast to the cost refund treatments, the tax notification in the cost reduction treatment does not mention the reimbursement and thus frames the change in collection costs as a real cost reduction.

Two more details about the cost reductions in our design are worth noting. First, the district administrations could actually achieve the cost reduction communicated in the *30% cost reduction* treatment by contracting with service providers outside the

church administration at current market prices. While the church's in-house service provider charges the district administrations a piece rate of €0.72 per mailed notification, the market price is closer to about €0.50, implying a possible cost reduction of about 30 percent. Second, in relative terms, the communicated cost reductions imply quite significant effects on net-of-cost tax revenue: in the control group, the average local church tax payment is €4.22, implying that a cost reduction by 30 percent (or €0.22 per notification) results in an increase of net-of-cost revenue by 6.3 percent on average. In the *100% cost refund* treatment, the reduction of local expenses by €0.72 per notification amounts to an average increase of net-of-cost revenue by 20.6 percent.

4.2 Data and Sampling

4.2.1 Data Sources and Overall Sample in the Experiment

The two church districts comprise about 35,000 adult church members, all of which received a tax notification. The samples used to derive our results do not include individuals living in households with more than one taxpayer being liable for the local church tax of the Protestant Church. The reason is that in such households, family members liable for the local church tax often combine their tax payments into a single bank transfer, without disaggregating the overall amount paid into individual payments. This introduces error in our measurement of compliance with the local church tax. After excluding individuals from such households, we were left with 24,365 individual observations.¹²

Our data links three different sources. First, the state-level administration of the Protestant Church provided us with individual characteristics on all taxpayers residing in either of the two districts, including age, gender, and the parish the taxpayer belongs to. Second, we obtained individual-level records of local church tax payments from the district administrations for a total of three tax years, 2014 to 2016. The data comprise the individual's tax number, the payment date and the amount paid. While the data for the two baseline years (2014 and 2015) are used to construct measures

¹²We also excluded employees of the Protestant Church from the experiment. In households with more than one household member being liable for the local church tax, all taxpayers received the same notifications. If we include those individuals in our estimations, all our results remain unchanged.

of baseline compliance with the local church tax, the data for the treatment year 2016 is used to construct our outcomes of interest. The third data source is individual-level records of the state church tax, comprising information on taxable income for all members of the Protestant Church who file for the federal income tax and reside in the participating church districts. This data allows us to derive the individual-level tax base of the local church tax for the baseline years and the treatment year. For income tax filers, we can thus determine the true local church tax liabilities (in baseline years and in the treatment year) and compare it to actual tax payments. This provides us with an objective measure of compliance with the local church tax.

4.2.2 Measuring Baseline Compliance

We exploit the possibility to measure compliance in the baseline to differentiate between different compliance types, i.e. taxpayers with different levels of baseline tax morale. This is very important for our study because we cannot expect purely rational taxpayers to respond to a reciprocity treatment: given zero deterrence, taxpayers behaving according to the Allingham-Sandmo framework (i.e., low tax morale types) will always maximize their utility by fully evading the local church tax, irrespective of whether the local administration signals low or high performance. The ability to determine individuals' baseline tax morale should thus greatly improve our chances of identifying a reciprocity response.

We distinguish two main taxpayer types in terms of baseline compliance behavior. As a first type, we consider baseline donors and compliers. These are individuals who have revealed a high level of tax morale by either complying with the local church tax in all baseline years (provided that we observe the local church tax payment and the true tax liability, and the tax liability is strictly positive), or by overpaying (and thus donating) at least once. The second taxpayer type is baseline evaders. Baseline evaders are all individuals who have revealed a low level of tax morale by underpaying the local church tax at least once in the baseline, and by not overpaying in any year.¹³

The baseline data show that the local church tax is a setting where pervasive tax

¹³All our results are robust to alternative type definitions as regards individuals who underpay in one baseline year, but comply (or overpay) in the other baseline year.

evasion coexists with voluntary compliance. In our sample, 17.7 percent of taxpayers whose type can be determined are of the baseline-donor-or-complier type, and 82.3 percent are classified as baseline evaders. The dominance of evasion results in timely tax payments amounting to only 24.4 percent of income tax filers' true local church tax liabilities. Among all taxpayers (irrespective of whether we can observe their type or not), only 15 percent make a local church tax payment in the baseline, generating on average € 3.99 of revenue per notification.

4.2.3 Samples for Estimation

It is important to note that we observe true tax liabilities for some, but not all individuals. This is because for most taxpayers, it is not mandatory to file an income tax declaration. If individuals do not file, no data is transmitted from the state tax authority to the church tax office.¹⁴ As a consequence, for non-filers we do not observe the individual's income in church tax records. The church tax records also have missing values due to late filers.¹⁵ Finally, we had to match church tax records with records for local church tax payments based on the individual's name and address. Imperfections in this matching also lead to missing values for the tax base of the local church tax (and, hence, the true tax liability).

Due to missing values in the true tax liability, our sample size varies, depending on which outcome is considered and whether we condition on baseline compliance behavior or not. We consider outcome variables that allow us to assess both the total effect, i.e., the combination of extensive and intensive margin responses, and variables that capture the extensive margin responses to our treatment. For the extensive margin we use four outcome variables: the probability of payment, the probability of donation, the probability of exact compliance, and the probability of evasion. The combined effect of extensive and intensive margin responses is captured using the payment amount as the dependent variable.¹⁶ When using either the probability of

¹⁴The church collects the state church tax from wage earners through automatic tax withholding, very similar to withholding of the federal income tax.

¹⁵Of course, late filing is less of a problem for the baseline years. As a result, we have fewer missing values for the true tax liability in baseline years as compared to the treatment year.

¹⁶We follow standard procedures in the tax compliance literature and consider only timely payments to define the outcome variables. Among all payments arriving until December 31 of the treatment year,

payment or the payment amount, and provided that we do not condition on past compliance behavior, we can exploit the full sample, comprising all taxpayers in the experiment. When considering the remaining outcomes, or when conditioning on past compliance, we have to rely on smaller samples. Table 1 shows the various samples used in our analysis, and corresponding distributions across treatments. As can be seen from the table, our sampling procedure ensures a symmetric size of treatment groups in all samples. Column (1) shows the full sample (all taxpayers in the experiment). As can be seen from Column (2) of Table 1, we observe the true tax liability in the treatment year for a total of 5282 individuals. The remaining columns split the full sample according to baseline compliance behavior. In total, we are able to classify 8617 individuals in terms of their baseline compliance behavior, 1525 as baseline donors or compliers (Column 3), and 7092 as baseline evaders (Column 5). Columns (4) and (6) display the respective samples if we condition on the true tax liability being observed both in the baseline and in the treatment year.

Among baseline donors and compliers, we observe the true tax liability in about 63 percent of the cases (Columns (3) and (4)). Among baseline evaders, this rate is only about 55 percent (Columns (5) and (6)). One possible reason for the difference is that baseline payment behavior could proxy for the unobserved tendency of taxpayers to keep their tax matters in good order. More orderly types would not only be more compliant with the local church tax, but also more likely to file their federal income tax return before the standard deadline (May 31st), making it more likely that the tax base of the local church tax for the treatment year is observed in our data.

4.2.4 **Balancedness and Persistence of Compliance Over Time**

Table 2 shows taxpayers' baseline characteristics for our main treatments. Panel A reports the respective means for baseline donors and compliers, Panel B for baseline evaders, and Panel C for the full sample. The table also displays, for both the *30% cost reduction* and the *100% cost refund* treatment groups, the estimated differences in means relative to the *30% cost refund* group and the corresponding standard errors. None of the differences is statistically different from zero at conventional levels. Table

80.2 percent were timely payments.

[A2](#) in the appendix reports more detailed baseline characteristics (including also the control group) and confirms that all treatment groups are balanced in observable characteristics. The same holds true for all treatment groups if we consider the (smaller) samples of taxpayers that are used to estimate the extensive margin responses requiring information about the tax base in the treatment year.

Given that we use baseline payment behavior to measure individuals' tax morale, it is important that overall payment behavior absent any treatment did not change significantly between the baseline and the treatment year. Using the sample of taxpayers assigned to the control group, we can compare average compliance behavior across tax years. As documented in [Table A3](#) in the appendix, compliance behavior is very consistent over time. Among taxpayers who received a tax notification in 2015 and 2016, the share of individuals making a payment was 14.2 and 14.8 percent, respectively. The difference is significant at the 10 percent level (p -value = 0.06). For average payment amounts, the figures are € 3.95 and € 4.28, respectively (p -value = 0.01). The fact that the payment probability and the average payment amount slightly increased between 2015 and 2016 is likely driven by income growth. Among individuals for whom we have income information for both years, taxable income growth was on average 3.7 percent, leading to an increase in the average local church tax liability of € 1.16, or 4.2 percent. This interpretation is corroborated by the fact that the distribution of the different compliance types did not change between 2015 and 2016: In 2015, 83.7 percent of income tax filers with a strictly positive tax liability (partly) evaded the local church tax, while 16.3 percent either paid the amount due or overpaid. In 2016, the shares were 82.8 percent and 17.2 percent, respectively (p -value = 0.30).

4.2.5 Representativeness

A possible concern regarding the external validity of our study could be that church members might differ from the more general population in important respects. For instance, it is often claimed that Protestants have a special work ethic, which could translate into higher incomes and specific attitudes towards taxes. It could also be that church members are more pro-social. However, [Table A4](#) in the online appendix

demonstrates that, at least in observable characteristics, Protestants are very similar to the overall population in Germany, and also similar to non-church members. The data reported in Table A4 are from the personal income tax statistics for 2007 and show mean characteristics for single and for joint filers (couples). If anything, members of the Protestant church earn slightly lower incomes, most likely reflecting a higher tendency among high-income church members to quit their membership and avoid the church tax. Interestingly, the average amount of charitable donations is very similar across the different groups, suggesting that, in general, Protestants are not exceptional in terms of pro-sociality.

5 Empirical Results

5.1 Estimation

We present estimates of the cost reduction (reciprocity) and increased cost refund (crowding out) effects separately for our two baseline compliance types (baseline donors and compliers vs. baseline evaders) and for the full sample. Importantly, the focus on the heterogeneity of the treatment effects in terms of baseline compliance was pre-registered and follows the example of [Dwenger *et al.* \(2016\)](#).¹⁷

All our results are derived from simple OLS regressions. For the total response, we estimate the specification

$$y_i = \alpha + \beta_1 REDUC_i + \beta_2 REFUND_i + X_i \gamma + \epsilon_i, \quad (1)$$

where y_i denotes the amount paid by individual i , $REDUC_i$ denotes an indicator for the 30% COST REDUCTION treatment, $REFUND_i$ denotes an indicator for the 100% COST REFUND treatment, X_i is a vector of taxpayer characteristics used for stratification, and ϵ_i is an error term. X_{ip} consists of a series of indicators for age quartiles, gender, the taxpayers' tax bracket according to the local church tax schedule, including a dummy

¹⁷See the AEA RCT Registry entry AEARCTR-0001332. In the registry entry, we stated that we planned to “study the heterogeneity of treatment effects by interacting treatments with information on tax payments prior to the experiment.” Instead of running regressions with interactions, we decided to report regressions with sample splits according to baseline compliance types. We report the results of estimating treatment effects with interaction terms in the appendix (Table A5).

for taxpayers for whom we do not observe the true tax liability, and the parish the taxpayer belongs to. Most importantly, the inclusion of tax bracket indicators flexibly controls for income effects. We estimate (1) while using the *30% cost refund* treatment as the omitted reference group. The reciprocity effect is given by the estimate of β_1 , measuring the intent-to-treat effect of a 30% cost reduction relative to an equivalent shift of administrative costs to a third party, i.e. the state-level administration of the church. The crowd-out effect is given by β_2 .

For the extensive margin responses (probability of payment, probability of donation, probability of exact compliance, and probability of evasion), we estimate linear probability models.¹⁸ The left-hand side of equation (1) then becomes $Prob(\text{payment}_i > 0)$, $Prob(\text{payment}_i > \text{tax liability}_i)$, $Prob(\text{payment}_i = \text{tax liability}_i)$, or $Prob(\text{payment}_i < \text{tax liability}_i)$, respectively.

Before turning to the results, we would like to point out that a key identifying assumption is that there are no spillovers across treatments. Arguably, such spillovers would most likely occur within households comprising more than one adult member of the Protestant church. We reiterate that we excluded such households from the experiment. As stated before, we also excluded all church employees and priests. Furthermore, there was no media coverage of the field experiment. Finally, evidence supporting the credibility of our experimental design with respect to the no-spillovers assumption comes from [Dwenger et al. \(2016\)](#), who set up a telephone enquiry line for recipients of the local church tax notification when implementing their field experiment in a different urban area in Bavaria. They found that only 0.34 percent of taxpayers called the enquiry line, most of them with questions relating to the tax base.

5.2 High Tax Morale Types: Baseline Donors and Compliers

We begin the discussion of our empirical results with the treatment responses of baseline donors and compliers. Panel A in Table 3 reports the corresponding findings. In addition, Figure 1 depicts treatment effects derived from estimations without controls

¹⁸All our results are unchanged if we fit probit models instead.

relative to the mean outcome in the *30% cost refund* treatment group.¹⁹

Starting with the effects of the *30% cost reduction* treatment, we note that the total effect is positive and significant. Baseline donors and compliers receiving the tax notice communicating a real reduction in administrative costs pay €4.21 more on average compared to high tax morale types receiving the notice pointing to an equivalent shift of costs to a higher level of the church administration. Relative to the average payment amount of €24.63 in the *30% cost refund* treatment group, this is an increase by about 17 percent. The probability of payment is not affected: the coefficient is small and not statistically different from zero. This implies that the total effect is completely driven by a positive intensive margin response. With the remaining estimates in columns (3) to (5), we can trace out the reciprocal treatment response of the high tax morale types in more detail. The results show that the total effect is mostly driven by a strong increase in the probability of donating, i.e. overpaying the local church tax. Relative to the mean of 28.9 percent in the *30% cost refund* treatment group, the probability of donating increases by 13.4 percentage points, or 46 percent. In contrast, both for the probability of exact compliance and the probability of evading the point estimates are negative. The reduction in the probability of exact compliance is statistically significant at the 10 percent level. This means that the reduction of administrative costs has led a significant share of highly motivated taxpayers to switch from exact compliance to overpaying the church tax. Taken together, the effects of the *30% cost reduction* treatment in the sample of baseline donors and compliers show a significant reciprocal response to a reduction of administrative costs. The high tax morale types respond to the cost reduction by a strongly increased willingness to overpay the local church tax relative to their true tax liability, translating into a significant increase in average tax payments from this group.

Turning to the effects of the *100% cost refund* treatment in Panel A, we find none of the estimates for the five different outcomes to be significantly different from zero. We conclude from Panel A that high tax morale types reciprocate cost reductions, but they do not show any crowd-out in response to an implicit transfer to the local

¹⁹As documented in part 5.5, none of our results depends on whether or not we use controls for strata.

public good.²⁰ In terms of the underlying preferences discussed by the public-goods crowding-out literature, our findings imply that at the margin, high tax morale types are motivated by joy-of-giving, or warm glow, rather than altruism. This is because, for a taxpayer motivated by altruism, a third-party transfer to the public good is a perfect substitute for her own private contribution, and would thus trigger a reduction in voluntary tax payments. We cannot say, however, whether the joy-of-giving part in taxpayers' utility is separated from the reciprocity part, or whether reciprocity and joy-of-giving are interlinked. From a policy perspective, the crucial insight from our findings on high tax morale types is that a positive signal of the administration's performance leads to a positive reciprocal response and zero crowd-out. This suggests a potential for policy makers to improve tax compliance among high tax morale individuals by highlighting improvements in how tax payments are transformed into public goods and services.

5.3 Low Tax Morale Types: Evaders

We next turn to the treatment responses of the low tax morale types (i.e., the baseline evaders) in our sample. Panel B in Table 3 collects the corresponding estimation results. We report exactly the same estimations as for baseline donors and compliers. Figure 2 depicts the treatment effects relative to the mean of the outcome in the 30% *cost refund* treatment group as the omitted reference group, as in Figure 1 based on estimations without controls.

A first observation is that, for both the cost reduction and the increased cost refund treatment, the effects on the payment amount are weakly significant and negative, suggesting that both treatments result in a decrease of overall compliance relative to the 30% *cost refund* treatment. Given that in the 30% *cost refund* treatment the average payment amount among evaders is only €2.52, the absolute size of the total responses is necessarily small. In relative terms, the treatments diminish average tax payments by about 25 percent and 28 percent, respectively. The probability of payment is also

²⁰In columns (3) and (4), the estimates of the two treatment effects are significantly different from each other (p -values < 0.01). In column (1), the estimates are not significantly different from each other (p -value = 0.304).

negatively affected, but only the coefficient of the *100% cost refund* treatment is statistically significant. While the probability of payment is 8.0 percent in the *30% cost refund* group, it is only 5.8 percent in the *100% cost refund* treatment. Turning to the effects on the probability of donation, exact compliance, and evasion, we note that both treatments have significantly shifted tax compliance decisions among baseline evaders from exact compliance towards evasion. Specifically, while 4.6 percent of baseline evaders exactly comply with the local church tax in our reference group (*30% cost refund*), this share is reduced by 1.9 percentage points (*cost reduction*) and 2.2 percentage points (*increased cost refund*), respectively. The diminished rate of exact compliance translates almost exclusively into a corresponding increase in the probability of evasion.

In summary, the results in Panel B demonstrate that both the *cost reduction* and the *increased cost refund* treatment have crowded out voluntary compliance among baseline evaders. As regards the *increased cost refund* treatment, this is consistent with a free-riding behavior of some low tax morale types who reduce their tax payment in response to the implicit transfer of resources to the local public good that goes along with the *increased cost refund*. However, interpreting this finding in terms of the underlying preferences is difficult for two reasons. First, in the *30% cost refund* treatment group, the payment probability is only 8.0 percent, meaning that 92.0 percent of baseline evaders are in a corner solution. We cannot say whether or not the *increased cost refund* has further crowded out the willingness of those taxpayers to pay the local church tax. Second, taxpayers with low tax morale who end up in the *100% cost refund* treatment group and want to reduce their tax payment relative to the amount they would have paid if they were in the *30% cost refund* treatment group face a choice set with focal points from the tax schedule. It is likely that those focal points, including the option to reduce the tax payment to zero, distort the crowd-out relative to a situation without focal points. We therefore abstain from interpreting to what extent the behavior of low tax morale types is motivated by altruism or warm glow.

Another observation is that the responses to the *cost reduction* treatment among evaders are quite similar (but somewhat weaker) compared to the *increased cost re-*

fund treatment. This suggests that the implicit transfer associated with the increased cost refund treatment is not the only force reducing compliance among low tax morale types.²¹ One possible interpretation is that some taxpayers may perceive the cost reduction not as a one-time effect, but as a permanent reduction of the church district's expenses. This would lead to a higher perceived present value of the 30% cost reduction relative to the 30% cost shift in the reference group. The implicit transfer to the local public good associated with the difference in present values could trigger a crowding out of voluntary tax payments even in the cost reduction group. Most importantly, however, we learn from the responses of low tax morale types that the cost reduction treatment does *not* lead to improved compliance among evaders. We conclude that reciprocity does not uniformly affect tax compliance: it is a factor that can affect compliance of high tax morale types quite strongly, but it does not seem to shape compliance behavior of less motivated taxpayers.

5.4 Full Sample

For completeness, Panel C in Table 3 also reports the results for the full sample. Despite the small share of baseline donors and compliers, the strong positive effect of the cost reduction treatment on the probability of donating among highly motivated taxpayers (and the corresponding drop in the probability of exact compliance) translates into a small increase in the full sample. In contrast, the effects of the increased cost refund treatment are all small and insignificant.

The fact that the full sample results mask most of the heterogeneous responses discussed before highlights the importance of being able to distinguish between different baseline motivations for understanding tax compliance behavior. While the German church tax setting is certainly specific, it offers favorable conditions to objectively measure taxpayers' baseline compliance behavior, something that would be difficult to achieve in other contexts. It is this feature of our data that allows us to uncover the counteracting responses of the different motivational types.

²¹None of the differences in the estimated coefficients between treatments in Panel B of Table 3 is significantly different from zero.

5.5 Discussion and Robustness

In the following, we briefly summarize the results of some further analyses and robustness checks.

Income heterogeneity As discussed before, the treatment groups are balanced in observables, including income (see Table A2). However, the diverging responses by subjects with high and low baseline tax morale could still essentially reflect different treatment responses by high and low income individuals. For instance, it could be that high-income individuals are more likely to pay the local church tax, and at the same time behave more reciprocal. In contrast, low-income individuals could show lower compliance in the baseline and at the same time be more likely to free-ride on administrative cost reductions. Table A7 in the online appendix explores this issue by splitting the samples into taxpayers with a tax liability smaller or equal to €10 (annual income up to €25,000) and taxpayers with a tax liability between €25 and €100 (annual income of more than €25,000). We chose this split because it results in comparable sample sizes in lower and higher brackets.

For baseline donors and compliers, we do not observe any striking differences between taxpayers in lower and higher brackets. Most importantly, the effects of the cost reduction treatment on the probability of donating are relatively similar: For the lower brackets, the point estimate is 0.107, with a p -value of 0.113. For the higher brackets, the point estimate is 0.145 and is significant at the 1% level. We conclude that the main component of the reciprocity effect identified in our study is similar across the different parts of the income distribution, but more precisely estimated for taxpayers in higher tax brackets.

For baseline evaders, Table A7 reveals that the crowding out of tax payments is larger and more precisely estimated for taxpayers in higher income brackets. Our main conclusion is thus that the diverging responses by individuals with high and low baseline tax morale identified in our main analysis do not reflect diverging effects by high and low income individuals. The evidence rather suggests that taxpayers in higher tax brackets have responded more strongly to our treatments: Among high tax morale individuals, the positive reciprocal response to the cost reduction is somewhat

more pronounced among higher incomes, and among low tax morale types the negative response to the increased cost refund (crowding out) is mostly driven by taxpayers from higher tax brackets.

An alternative way to insure ourselves against the possibility that our main findings may reflect an underlying income heterogeneity is presented in Table A8. It shows entropy-reweighted estimates of our treatment effects following the methodology of Hainmueller (2012). The idea behind the analysis is to construct weights for the individual observations in a given subsample ensuring that certain characteristics in the reweighted subsample show the same distribution as in the overall sample. In our case, considering the sample of baseline donors and compliers, the entropy weights are chosen such that in the last pre-treatment year, the reweighted sample of baseline donors and compliers displays the same income distribution as the overall sample (i.e., all individuals in the experiment for whom we observe the baseline income). Regarding the sample of baseline evaders, the entropy weights make sure that in the last pre-treatment year, the reweighted sample of baseline evaders follows the same income distribution as the overall sample. If the treatment effects from estimations using the entropy weights are similar to the unweighted estimates, we can conclude that it is unlikely that income heterogeneity is driving the differences in treatment responses between high and low tax morale types.

We implement the reweighting using a series of six indicators for the different income brackets of the local church tax. Hence, when we estimate the treatment effects for either baseline donors and compliers or baseline evaders, all six indicators have means (after reweighting) that are equal to the respective means in the overall sample.

Table A8 documents that all our previous findings are robust to entropy reweighting. This bolsters our confidence that the treatment effect heterogeneity between the different baseline compliance types is not an artifact driven by unobserved forms of income heterogeneity.

Estimations without controls As it is common practice, all our main estimations control for strata variables. As described before, this means we account for age, gen-

der, baseline income brackets, and parish fixed effects. Table A6 in the online appendix additionally reports the regressions that are used to construct Figures 1 and 2. In these regressions, we do not include any controls.²² The results show that our findings do not depend on whether or not we control for strata variables.

6 Conclusion

This paper asks how administrative efficiency affects tax compliance. We report on a field experiment that randomly varied the collection costs of the local church tax in Germany, and study how individual compliance decisions respond to the improved efficiency of the tax administration. Because the church does not enforce the local church tax, this is a context where compliance reveals individuals' tax morale. Another valuable feature of the local church tax is that we can objectively describe individual compliance by linking data that allow us to compare tax payments to taxes owed. Using the linked data, we classify individuals according to their baseline tax morale. Our main type distinction is between baseline donors and compliers (high tax morale types) and baseline evaders (low tax morale types). We note that a reduction in administrative costs is associated with an implicit transfer to the local public good. This requires to disentangle a potential reciprocal response from a crowding out of voluntary compliance that occurs if taxpayers perceive the implicit transfer as a substitute for their voluntary tax payment.

To identify reciprocity, we compare a treatment that communicates a reduction in tax collection costs with one that refers to an equivalent shift of collection costs to a third party. In this comparison, the implicit transfer is held constant. We find that high tax morale types reciprocate a reduction in tax collection costs by increasing their tax payments. The effect is mostly driven by a significant increase in the share of taxpayers acting as donors by paying more than the tax amount due. The crowding out of voluntary tax payments is identified by comparing treatment groups with different degrees of cost shifting to a third party. We find zero crowding out of voluntary com-

²²Note that in Table A6, the estimations using the payment amount as dependent variable include a church district dummy to account for the difference in the tax amount due for the income bracket between € 40,000 and € 54,999.

pliance among high tax morale types if the implicit transfer to the local public good is increased. Taken together, those findings suggest a potential for governments to improve tax compliance among high tax morale individuals via improved administrative performance. Our most important finding on low tax morale types is the absence of any reciprocal response. In fact, we observe that baseline evaders further decrease their compliance (which is already low in the baseline) in response to both treatments. This suggests that signaling enhanced administrative performance is likely to crowd out tax compliance among low tax morale types.

In terms of external validity, taking into account all aspects of the local church tax in Germany, the context studied in this paper seems rather specific. However, in our view this does *not* hold with respect to the local church tax not being enforced: around the world, and in particular in developing countries, there is an abundance of examples where important taxes are very poorly enforced. Hence, we believe one can draw some general conclusions from our analysis, at least for such environments. First, given the counteracting responses in our experiment, it is a rather difficult exercise to predict the overall compliance responses to improvements in administrative efficiency. Our findings suggest that the overall response will depend not only on the relative strength of the reciprocity effect and the crowd-out, but also on the distribution of tax morale in the population of taxpayers. Second, the prevalence of tax evasion in settings with weak enforcement and little or no third-party reporting documented in the recent literature suggests that most taxpayers have little intrinsic motivation to comply with the tax law. Based on our results, in such settings we would not expect improvements in administrative efficiency alone to be very effective in raising average tax compliance.

While our contribution speaks first and foremost to the literature on tax morale, it also links to research on charitable giving. This is because the church does not enforce the local church tax and relies exclusively on individuals' tax morale. Consistently, the charitable giving literature has shown that potential donors have a strong aversion against financing charities' overhead cost (Gneezy *et al.*, 2014; Meer, 2014). This paper shows that, in principle, an increased willingness to contribute in response to lower administrative costs extends to a setting with (legally) compulsory tax pay-

ments. However, it is noteworthy that the literature on charitable giving focuses on warm-list individuals. Compared to this narrow focus on a small subset of potential donors, when studying a tax setting it is natural to consider compliance responses of individuals across the whole range of motivations to voluntarily contribute to the common good. Our results suggest that the findings on charitable giving derived from warm-list individuals do not extend to less motivated subjects: in fact, we document that communicating lower administrative costs may actually crowd out contributions among cold-list individuals.

References

- ALLINGHAM, M. G. and SANDMO, A. (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, **1**, 323–338.
- ALM, J., JACKSON, B. R. and MCKEE, M. (1992a). Estimating the Determinants of Taxpayer Compliance with Experimental Data. *National Tax Journal*, **45**, 107–114.
- , MCCLELLAND, G. H. and SCHULZE, W. D. (1992b). Why Do People Pay Taxes? *Journal of Public Economics*, **48** (1), 21–38.
- ANDREONI, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy*, **97**, 1447–1458.
- , ERARD, B. and FEINSTEIN, J. (1998). Tax Compliance. *Journal of Economic Literature*, **36**, 818–860.
- BECKER, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, **76**, 169–217.
- BLUMENTHAL, M., CHRISTIAN, C. W. and SLEMROD, J. (2001). Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota. *National Tax Journal*, **54**, 125–138.
- BODEA, C. and LEBAS, A. (2016). The Origins of Voluntary Compliance: Attitudes

- toward Taxation in Urban Nigeria. *British Journal of Political Science*, **46** (01), 215–238.
- CAGALA, T., GLOGOWSKY, U., GRIMM, V., RINCKE, J. and TUSET CUEVA, A. (2018). Rent Extraction and Prosocial Behavior, mimeo, University of Erlangen-Nuremberg.
- CAVIOLA, L., FAULMÜLLER, N., EVERETT, J. A., SAVULESCU, J. and KAHANE, G. (2014). The Evaluability Bias in Charitable Giving: Saving Administration Costs or Saving Lives? *Judgment and Decision Making*, **9** (4), 303–316.
- CULLEN, J. B., TURNER, N. and WASHINGTON, E. (2015). Political Alignment and Tax Evasion, mimeo.
- DE NEVE, J., IMBERT, C., SPINNEWIJN, J., TSANKOVA, T. and LUTS, M. (2019). How to Improve Tax Compliance? Evidence from Population-Wide Experiments in Belgium, Saïd Business School Working Paper 2019-07.
- DEL CARPIO, L. (2014). Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru, mimeo, Princeton University.
- DOERRENBURG, P. and PEICHL, A. (2018). Tax Morale and the Role of Social Norms and Reciprocity: Evidence from a Randomized Survey Experiment, CESifo Working Paper No. 7149.
- DUFWENBERG, M. and KIRCHSTEIGER, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, **47** (2), 268–298.
- DWENGER, N., KLEVEN, H., RASUL, I. and RINCKE, J. (2016). Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany. *American Economic Journal: Economic Policy*, **8** (3), 203–232.
- FELLNER, G., SAUSGRUBER, R. and TRAXLER, C. (2013). Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information. *Journal of the European Economic Association*, **11**, 634–660.
- GNEEZY, U., KEENAN, E. A. and GNEEZY, A. (2014). Avoiding Overhead Aversion in Charity. *Science*, **346** (6209), 632–635.

- HAINMUELLER, J. (2012). Entropy Balancing: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis*, **20** (1), 25–46.
- HALLSWORTH, M., LIST, J. A., METCALFE, R. D. and VLAEV, I. (2017). The Behavioralist As Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance. *Journal of Public Economics*, **148**, 14–31.
- KLEVEN, H. J., KNUDSEN, M. B., KREINER, C. T., PEDERSEN, S. and SAEZ, E. (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, **79**, 651–692.
- LUTTMER, E. F. and SINGHAL, M. (2014). Tax Morale. *Journal of Economic Perspectives*, **28**, 149–168.
- MASCAGNI, G. (2018). From the Lab to the Field: A Review of Tax Experiments. *Journal of Economic Surveys*, **32** (2), 273–301.
- MEER, J. (2014). Effects of the Price of Charitable Giving: Evidence from an Online Crowdfunding Platform. *Journal of Economic Behavior & Organization*, **103**, 113–124.
- NARITOMI, J. (2014). Consumers as Tax Auditors, forthcoming in: *American Economic Review*.
- OECD (2013). *Tax and Development: What Drives Tax Morale?* Organization for Economic Cooperation and Development.
- OKTEN, C. and WEISBROD, B. A. (2000). Determinants of Donations in Private Non-profit Markets. *Journal of Public Economics*, **75** (2), 255–272.
- POMERANZ, D. (2015). No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. *American Economic Review*, **105**, 2539–2569.
- RIBAR, D. C. and WILHELM, M. O. (2002). Altruistic and Joy-of-Giving Motivations in Charitable Behavior. *Journal of Political Economy*, **110**, 425–457.

- ROBERTS, R. D. (1984). A Positive Model of Private Charity and Public Transfers. *Journal of Political Economy*, **92**, 156–148.
- SLEMROD, J., BLUMENTHAL, M. and CHRISTIAN, C. (2001). Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota. *Journal of Public Economics*, **79**, 455–483.
- and WEBER, C. (2012). Evidence of the Invisible: Toward a Credibility Revolution in the Empirical Analysis of Tax Evasion and the Informal Economy. *International Tax and Public Finance*, **19**, 25–53.
- TIMMONS, J. F. and GARFIAS, F. (2015). Revealed Corruption, Taxation, and Fiscal Accountability: Evidence from Brazil. *World Development*, **70**, 13 – 27.
- TORGLER, B. (2003). Tax Morale, Rule-Governed Behaviour and Trust. *Constitutional Political Economy*, **14** (2), 119–140.
- (2005). Tax Morale and Direct Democracy. *European Journal of Political Economy*, **21** (2), 525–531.
- WARR, P. G. (1982). Pareto Optimal Redistribution and Private Charity. *Journal of Public Economics*, **19**, 131–138.

Table 1: Overview of Treatments and Sample

Treatment	Full Sample		Baseline Donors & Compliers		Baseline Evaders	
	All	Tax Base 2016 Observed	All	Tax Base 2016 Observed	All	Tax Base 2016 Observed
	(1)	(2)	(3)	(4)	(5)	(6)
Control	6126	1328	396	246	1788	948
30% Cost Reduction	6066	1341	365	236	1765	992
30% Cost Refund	6090	1310	375	232	1781	972
100% Cost Refund	6083	1303	389	241	1758	952
Total Sample Size	24365	5282	1525	955	7092	3864

Notes: This table shows the various samples used in our analysis, and corresponding distributions across treatments. Column (1) shows the full sample (all taxpayers in the experiment). Column (2) displays the full sample, conditional on income (i.e., the tax base) being observed in the treatment year 2016. Column (3) shows the sample of baseline donors and compliers. This comprises all individuals from the full sample for whom we observe baseline compliance behavior (i.e., baseline income and baseline tax payment), and who have complied or overcomplied in the baseline. Column (4) displays the sample of baseline donors and compliers, conditional on income being observed in the treatment year 2016. Columns (5) and (6) show the respective samples for baseline evaders, i.e. individuals for whom we observe baseline compliance behavior, and who have (fully or partly) evaded the tax in the baseline.

Table 2: Baseline Characteristics by Treatment Group

	T1: 30% Cost Reduction	T2: 100% Cost Refund	T3: 30% Cost Refund	Difference T1 - T3	Difference T2 - T3
A: Baseline Donors and Compliers					
Female	0.54	0.517	0.544	0.004 (0.037)	0.027 (0.036)
Age	60.4	58.6	60.4	-0.1 (1.330)	1.86 (1.320)
Tax Liability \leq € 10	0.504	0.509	0.541	-0.037 (0.037)	0.032 (0.036)
€ 25 \leq Tax Liability \leq € 50	0.378	0.36	0.344	0.034 (0.035)	-0.016 (0.035)
Tax Liability \geq € 70	0.118	0.131	0.115	0.003 (0.024)	-0.016 (0.024)
Number of Observations	365	389	375	740	764
B: Baseline Evaders					
Female	0.484	0.48	0.466	-0.002 (0.017)	0.006 (0.017)
Age	46.2	46.6	46.3	-0.12 (0.490)	0.3 (0.490)
Tax Liability \leq € 10	0.45	0.465	0.454	-0.003 (0.017)	-0.012 (0.017)
€ 25 \leq Tax Liability \leq € 50	0.429	0.417	0.426	0.004 (0.017)	0.009 (0.017)
Tax Liability \geq € 70	0.12	0.118	0.121	0.000 (0.011)	0.003 (0.011)
Number of Observations	1765	1758	1781	3546	3539
C: Full Sample					
Female	0.568	0.566	0.567	0.000 (0.001)	0.001 (0.009)
Age	48.5	48.7	48.6	-0.12 (0.370)	-0.06 (0.370)
Tax Liability \leq € 10	0.247	0.25	0.254	-0.007 (0.008)	0.004 (0.008)
€ 25 \leq Tax Liability \leq € 50	0.155	0.152	0.153	0.002 (0.007)	0.001 (0.006)
Tax Liability \geq € 70	0.598	0.598	0.593	0.005 (0.009)	-0.005 (0.009)
Number of Observations	6066	6083	6090	12156	12173

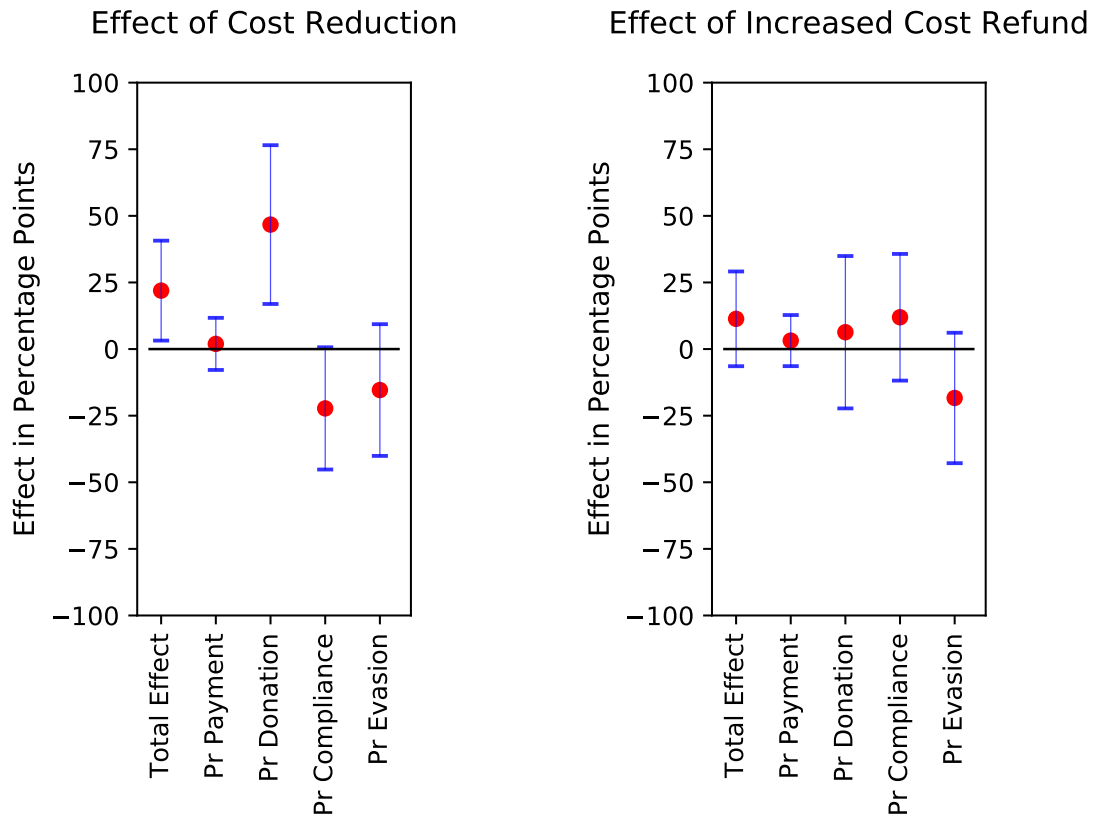
Notes: The table displays means of taxpayers' baseline characteristics for the main treatment groups in the experiment, together with estimated differences in means and corresponding standard errors in parentheses.

Table 3: Treatment Effects

	Payment	Probability of			
	Amount (1)	Payment (2)	Donating (3)	Compliance (4)	Evading (5)
A: Baseline Donors and Compliers					
Effect of Cost Reduction	4.207** (1.905)	0.015 (0.034)	0.134*** (0.042)	-0.083* (0.042)	-0.051 (0.043)
Effect of Increased Cost Refund	2.210 (1.887)	0.032 (0.033)	0.025 (0.041)	0.041 (0.044)	-0.066 (0.042)
Number of Observations	1129	1129	709	709	709
Average in Reference Group	€ 24.63	0.683	0.289	0.371	0.341
B: Baseline Evaders					
Effect of Cost Reduction	-0.627* (0.347)	-0.013 (0.009)	0.002 (0.005)	-0.019** (0.008)	0.017* (0.010)
Effect of Increased Cost Refund	-0.699** (0.350)	-0.022*** (0.008)	0.000 (0.005)	-0.022*** (0.008)	0.021** (0.010)
Number of Observations	5304	5304	2916	2916	2916
Average in Reference Group	€ 2.52	0.08	0.012	0.046	0.941
C: Full Sample					
Effect of Cost Reduction	0.259 (0.231)	-0.001 (0.006)	0.022** (0.010)	-0.029*** (0.011)	0.006 (0.014)
Effect of Increased Cost Refund	0.345 (0.231)	-0.000 (0.006)	0.003 (0.010)	-0.001 (0.012)	-0.002 (0.014)
Number of Observations	18239	18239	3954	3954	3954
Average in Reference Group	€ 3.51	0.139	0.069	0.108	0.823

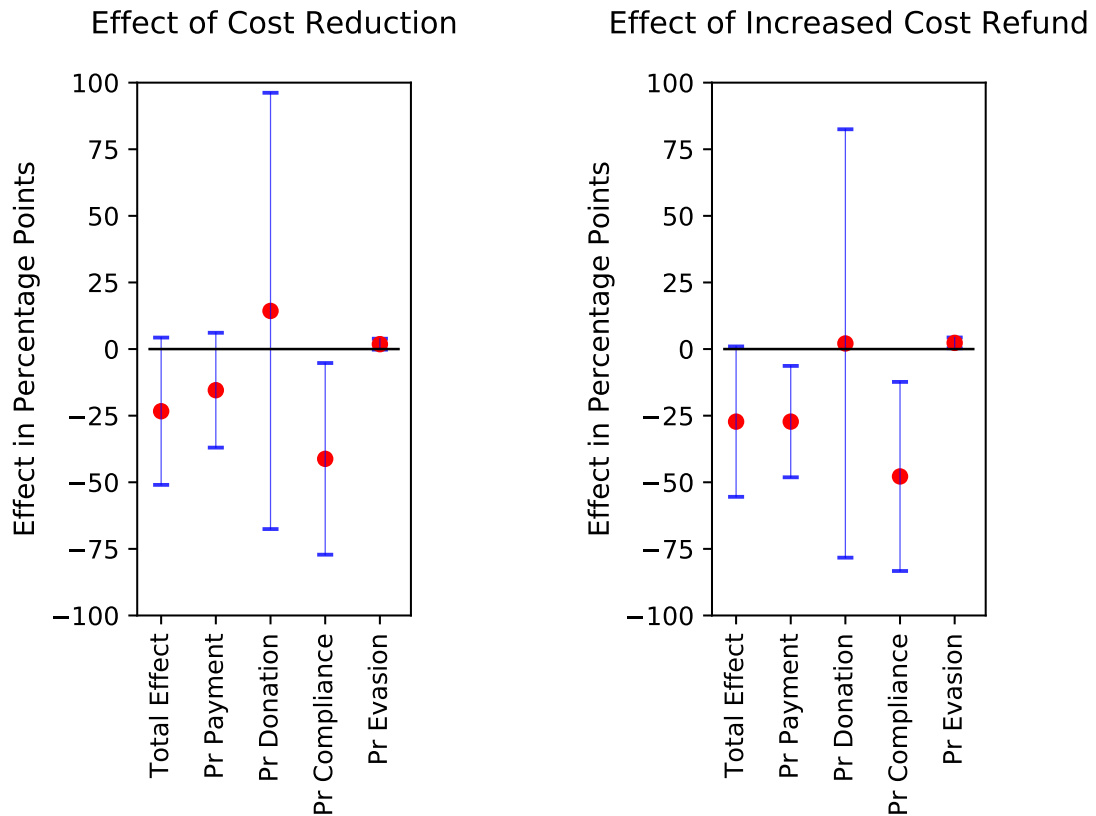
Notes: The table reports treatment effects of reducing administrative costs (30% Cost Reduction Treatment Group) and an increased refund of administrative costs (100% Cost Refund Treatment Group) relative to the 30% Cost Refund Treatment Group. Each column shows separate OLS regressions: Panel A displays the treatment effects for the sample of baseline donors and compliers, Panel B shows estimations for the sample of baseline evaders, and Panel C shows results for the full sample. Columns (3) to (5) use a smaller sample as compared to columns (1) and (2), because the dependent variables can only be constructed for the subpopulation of taxpayers for whom we observe the true tax liability in the treatment year. All regressions include a full set of controls for strata variables (based on age, gender, the tax amount due in the baseline including an indicator for missing values, and parish fixed effects). Robust standard errors in parentheses. ***, **, * denote significance level at 1, 5, 10 percent level, respectively.

Figure 1: Relative Treatment Effects for Baseline Donors and Compliers



Note: For each outcome, the figure shows the estimated treatment effect of the 30% cost reduction treatment (left panel) and the 100% cost refund treatment (right panel) relative to the mean in the omitted reference group (30% cost refund). The figure reports relative point estimates as red dots, and 95% confidence interval as spikes. The underlying estimations do not include further controls. For further details, see Table A6.

Figure 2: Relative Treatment Effects for Baseline Evaders



Note: For each outcome, the figure shows the estimated treatment effect of the 30% cost reduction treatment (left panel) and the 100% cost refund treatment (right panel) relative to the mean in the omitted reference group (30% cost refund). The figure reports relative point estimates as red dots, and 95% confidence interval as spikes. The underlying estimations do not include further controls. For further details, see Table A6.

Online Appendix (Not For Publication)

Table A1: Tax Schedule

Annual Income	Annual Tax Liability	% of Sample in Tax Bracket, 2016 (Conditional on Income Being Observed)
€ 8,005 to € 9,999	€ 5	3.8
€ 10,000 to € 24,999	€ 10	39.8
€ 25,000 to € 39,999	€ 25	33.8
€ 40,000 to € 54,999	€ 45 / € 50	12.1
€ 55,000 to € 69,999	€ 70	5.0
€ 70,000 and above	€ 100	5.5

Notes: This table shows the schedule of the local church tax in the two districts where the field experiment was implemented. In one of the districts, taxpayers falling into the bracket between € 40,000 and € 54,999 face a tax liability of € 45, while in the other district the respective tax liability is € 50.

Table A2: Balancedness Checks

	N	Female (2)	Age (3)	Indicators for Baseline Tax Liability										F-Test p-value (13)
				€0 (4)	€5 (5)	€10 (6)	€25 (7)	€45/€50 (8)	€70 (10)	€100 (11)	No info (12)			
Panel A: Baseline Donors and Compliers														
Control Group	396	0.52 [0.47;0.56]	60.0 [58.2; 61.8]	0.025 [0.01; 0.04]	0.083 [0.06; 0.11]	0.41 [0.36; 0.46]	0.22 [0.18; 0.26]	0.14 [0.10; 0.17]	0.051 [0.03; 0.07]	0.073 [0.05; 0.10]	-	0.999		
30% Cost Reduction	365	0.54 [0.49;0.59]	60.4 [58.5; 62.2]	0.011 [0.00; 0.02]	0.082 [0.05; 0.11]	0.41 [0.36; 0.46]	0.24 [0.20; 0.29]	0.13 [0.10; 0.17]	0.058 [0.03; 0.08]	0.060 [0.04; 0.08]	-	0.807		
100% Cost Refund	389	0.52 [0.47;0.57]	58.6 [56.8; 60.4]	0.028 [0.01; 0.04]	0.067 [0.04; 0.09]	0.41 [0.36; 0.46]	0.25 [0.20; 0.29]	0.11 [0.08; 0.14]	0.064 [0.04; 0.09]	0.067 [0.04; 0.09]	-	0.993		
30% Cost Refund	375	0.54 [0.49;0.59]	60.4 [58.6; 62.3]	0.035 [0.02; 0.05]	0.085 [0.06; 0.11]	0.42 [0.37; 0.47]	0.22 [0.18; 0.27]	0.12 [0.09; 0.15]	0.053 [0.03; 0.08]	0.061 [0.04; 0.09]	-	-		
Panel B: Baseline Evaders														
Control Group	1788	0.48 [0.46;0.51]	46.4 [45.7; 47.1]	0.034 [0.03; 0.04]	0.040 [0.03; 0.05]	0.38 [0.36; 0.41]	0.30 [0.28; 0.32]	0.11 [0.10; 0.13]	0.060 [0.05; 0.07]	0.070 [0.06; 0.08]	-	0.999		
30% Cost Reduction	1765	0.48 [0.46;0.51]	46.2 [45.5; 46.9]	0.027 [0.02; 0.03]	0.037 [0.03; 0.05]	0.39 [0.36; 0.41]	0.31 [0.28; 0.33]	0.12 [0.11; 0.14]	0.051 [0.04; 0.06]	0.069 [0.06; 0.08]	-	0.999		
100% Cost Refund	1758	0.48 [0.46;0.50]	46.6 [45.9; 47.3]	0.028 [0.02; 0.04]	0.048 [0.04; 0.06]	0.39 [0.37; 0.41]	0.30 [0.28; 0.32]	0.11 [0.10; 0.13]	0.049 [0.04; 0.06]	0.068 [0.06; 0.08]	-	0.999		
30% Cost Refund	1781	0.49 [0.46;0.51]	46.3 [45.6; 47.0]	0.028 [0.02; 0.04]	0.039 [0.03; 0.05]	0.39 [0.36; 0.41]	0.31 [0.29; 0.33]	0.12 [0.10; 0.13]	0.052 [0.04; 0.06]	0.069 [0.06; 0.08]	-	-		

Table A2: Balancedness Checks (Continued)

	N	Female	Age	Indicators for Baseline Tax Liability										F-Test p-value
				€0	€5	€10	€25	€45/€50	€70	€100	No info			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(10)	(11)	(12)	(13)		
Panel C: Full Sample														
Control Group	6126	0.56 [0.55;0.58]	48.7 [48.2; 49.2]	0.086 [0.08; 0.09]	0.018 [0.01; 0.02]	0.147 [0.14; 0.16]	0.107 [0.10; 0.11]	0.043 [0.04; 0.05]	0.021 [0.02; 0.02]	0.026 [0.02; 0.03]	0.552 [0.54; 0.56]	-		
30% Cost Reduction	6066	0.57 [0.56;0.58]	48.5 [47.9; 49.0]	0.088 [0.08; 0.09]	0.016 [0.01; 0.02]	0.143 [0.13; 0.15]	0.109 [0.10; 0.12]	0.046 [0.04; 0.05]	0.019 [0.02; 0.02]	0.025 [0.02; 0.03]	0.553 [0.54; 0.57]	1.000		
30% Cost Refund	6090	0.57 [0.56;0.58]	48.2 [48.1; 49.1]	0.090 [0.08; 0.10]	0.017 [0.01; 0.02]	0.147 [0.14; 0.16]	0.110 [0.10; 0.12]	0.425 [0.04; 0.05]	0.020 [0.02; 0.02]	0.025 [0.02; 0.03]	0.549 [0.54; 0.56]	1.000		
100% Cost Refund	6083	0.57 [0.55;0.58]	48.7 [48.2; 49.2]	0.086 [0.08; 0.09]	0.019 [0.01; 0.02]	0.145 [0.14; 0.15]	0.110 [0.10; 0.12]	0.042 [0.04; 0.05]	0.019 [0.02; 0.02]	0.024 [0.02; 0.03]	0.554 [0.54; 0.57]	1.000		

Notes: The table presents balancedness checks. Panel A refers to the sample of taxpayers who have paid the amount due or who have overpaid in the baseline (baseline donors and compliers), Panel B refers to the sample of taxpayers who have underpaid in the baseline (baseline evaders), and Panel C refers to the full sample. Column (1) displays the number of individuals, column (2) the share of females and (3) the average age. Columns (4) to (11) display means for a series of indicators for individuals' baseline tax liability according to the tax schedule. We use the latest of the two baseline years with available income information to determine baseline tax liabilities. Column (12) displays the share of individuals for whom we do not observe the true tax liability in the baseline. Column (13) shows p-values of an F-Test, testing whether the observable characteristics are jointly significant in predicting assignment to treatment relative to the group serving as omitted reference group in our estimations of treatment effects, i.e. the 30% Cost Refund treatment group. 95% confidence intervals in squared brackets.

Table A3: Persistence of Compliance Over Time

	Control Group 2015 (Baseline)	Control Group 2016 (Treatment Year)	Observations	Difference 2016 - 2015
Payment Amount	3.95	4.28	7979	0.34** (0.14)
Payment > 0	0.142	0.148	7979	0.006* (0.003)
Payment \geq Amount Due (Donors & Compliers)	0.163	0.172	1342	0.009 (0.009)
Payment < Amount Due (Evaders)	0.837	0.828	1342	-0.009 (0.009)

Notes: The table documents the persistence of compliance behavior over time. The sample consists of all taxpayers in the control group who are present in the last baseline year (2015) and in the treatment year (2016). The table displays the means for the different outcomes describing compliance behavior, together with estimated differences in means between 2016 and 2015, and corresponding standard errors in parentheses. ** and * denote significance level at 5 and 10 percent level, respectively.

Table A4: Comparability of Protestants to General Population

	Single Filers			Joint Filers		
	All	Protestants	Non-Church Members	All	Protestants	Non-Church Members
Female	47.4%	52.6%	42.8%	50.0%	50.0%	50.0%
Number of Children with Child Allowances	0.3	0.3	0.4	1.1	1.1	0.9
Age	42.8	44.0	42.4	48.2	48.8	47.4
Taxable Income (€)	30425	28153	33482	54795	54527	55064
Declarations with Wage Income	85.9%	85.8%	85.0%	93.1%	92.8%	94.5%
Declarations with Capital Income	19.0%	21.9%	15.3%	19.2%	21.1%	14.9%
Declarations with Business Income	3.9%	3.3%	4.6%	6.6%	6.6%	5.9%
Charitable Giving (€)	140.0	136.5	144.0	286.8	289.8	305.7

Notes: This table shows the mean characteristics for three groups of income tax filers in Germany, separately for single and joint filing: All filers, Protestants, and non-church members. The source of data are personal income statistics for 2007.

Table A5: Treatment Effects of Model with Interaction Terms

	Payment	Probability of			
	Amount (1)	Payment (2)	Donating (3)	Compliance (4)	Evading (5)
A: Effects for Baseline Donors and Compliers					
Donors & Compliers × Cost Reduction	5.49** (2.201)	0.023 (0.035)	0.131*** (0.043)	-0.065 (0.044)	-0.067 (0.044)
Donors & Compliers × Increased Cost Refund	3.147 (2.096)	0.048 (0.034)	0.021 (0.041)	0.062 (0.045)	-0.083* (0.043)
B: Effects for Baseline Evaders					
Cost Reduction	-0.635* (0.346)	-0.013 (0.009)	0.002 (0.005)	-0.019** (0.008)	0.017* (0.010)
Increased Cost Refund	-0.699** (0.350)	-0.022*** (0.008)	0.00 (0.005)	-0.022*** (0.008)	0.021** (0.010)
Donors & Compliers	21.52*** (1.53)	0.558*** (0.026)	0.239*** (0.029)	0.324*** (0.033)	-0.563*** (0.033)
Intercept	-0.69 (0.971)	0.038 (0.026)	-0.016 (0.029)	0.049 (0.044)	0.967*** (0.043)
Number of Observations	6,433	6,433	3,625	3,625	3,625

Notes: The table reports the results of a model with interactions. Each column shows separate OLS regressions: Panel A displays the treatment effects for baseline donors and compliers, and Panel B shows the results for baseline evaders. The interaction terms show the difference in the reaction of donors and compliers to the treatment as compared to evaders. Columns (3) to (5) use a smaller sample as compared to columns (1) and (2) because the dependent variables can only be constructed for the subpopulation of taxpayers for whom we observe the true tax liability in the treatment year. All regressions include a full set of controls for strata variables (based on age, gender, the tax amount due in the baseline including an indicator for missing values, and parish fixed effects). Robust standard errors in parentheses. ***, **, * denote significance level at 1, 5, 10 percent level, respectively.

Table A6: Treatment Effects Without Controls

	Payment	Probability of			
	Amount (1)	Payment (2)	Donating (3)	Compliance (4)	Evading (5)
A: Baseline Donors and Compliers					
Effect of Cost Reduction	5.403** (2.356)	0.013 (0.034)	0.135*** (0.044)	-0.083* (0.043)	-0.052 (0.043)
Effect of Increased Cost Refund	2.793 (2.234)	0.022 (0.033)	0.018 (0.042)	0.044 (0.045)	-0.063 (0.043)
Number of Observations	1129	1129	709	709	709
Average in Reference Group	€ 24.63	0.683	0.289	0.371	0.341
B: Baseline Evaders					
Effect of Cost Reduction	-0.588* (0.355)	-0.012 (0.009)	0.002 (0.005)	-0.020** (0.008)	0.017* (0.010)
Effect of Increased Cost Refund	-0.686* (0.363)	-0.022** (0.009)	0.000 (0.005)	-0.022*** (0.008)	0.022** (0.010)
Number of Observations	5304	5304	2916	2916	2916
Average in Reference Group	€ 2.52	0.080	0.012	0.046	0.941
C: Full Sample					
Effect of Cost Reduction	0.273 (0.243)	-0.001 (0.006)	0.022** (0.011)	-0.030*** (0.011)	0.009 (0.015)
Effect of Increased Cost Refund	0.323 (0.243)	-0.000 (0.006)	0.003 (0.010)	-0.002 (0.012)	-0.001 (0.015)
Number of Observations	18239	18239	3954	3954	3954
Average in Reference Group	€ 3.51	0.139	0.069	0.108	0.823

Notes: The table reports treatment effects of reducing administrative costs (30% Cost Reduction Treatment Group) and an increased refund of administrative costs (100% Cost Refund Treatment Group) relative to the 30% Cost Refund Treatment Group. Each column shows separate OLS regressions: Panel A displays the treatment effects for the sample of baseline donors and compliers, Panel B shows estimations for the sample of baseline evaders, and Panel C shows results for the full sample. Columns (3) to (5) use a smaller sample as compared to columns (1) and (2), because the dependent variables can only be constructed for the subpopulation of taxpayers for whom we observe the true tax liability in the treatment year. In Column (1), regressions include a church district dummy (no other control variables included). Robust standard errors in parentheses. ***, **, * denote significance level at 1, 5, 10 percent level, respectively.

Table A7: Treatment Effects by Income Level

	Low Income: Taxable Income < €25,000					High Income: Taxable Income ≥ €25,000				
	Payment Amount (1)	Payment (2)	Probability of		Evading (5)	Payment Amount (6)	Payment (7)	Probability of		Evading (10)
			Donating (3)	Compliance (4)				Donating (8)	Compliance (9)	
A: Baseline Donors and Compliers										
Effect of Cost Reduction	0.008 (2.176)	-0.051 (0.059)	0.107 (0.067)	-0.097 (0.063)	-0.011 (0.065)	5.619 (3.772)	0.063 (0.055)	0.145*** (0.055)	-0.075 (0.060)	-0.070 (0.060)
Effect of Increased Cost Refund	1.437 (2.363)	0.038 (0.054)	0.076 (0.065)	0.011 (0.064)	-0.087 (0.060)	-1.252 (3.777)	0.018 (0.056)	-0.027 (0.054)	0.056 (0.062)	-0.029 (0.060)
Number of Observations	366	366	328	328	328	381	381	381	381	381
Average in Reference Group	€14.92	0.730	0.336	0.336	0.327	€37.80	0.728	0.248	0.400	0.352
B: Baseline Evaders										
Effect of Cost Reduction	-0.328 (0.269)	-0.009 (0.016)	-0.004 (0.010)	-0.014 (0.014)	0.017 (0.016)	-1.004 (0.688)	-0.018 (0.017)	0.003 (0.006)	-0.021* (0.011)	0.018 (0.012)
Effect of Increased Cost Refund	-0.216 (0.306)	-0.010 (0.017)	-0.002 (0.010)	-0.010 (0.014)	0.012 (0.017)	-1.331* (0.728)	-0.047*** (0.016)	0.001 (0.005)	-0.029*** (0.010)	0.028** (0.012)
Number of Observations	1,292	1,292	1,126	1,126	1,126	1,790	1,790	1,790	1,790	1,790
Average in Reference Group	€0.94	0.062	0.019	0.040	0.942	€3.79	0.113	0.008	0.050	0.941

Notes: The table reports treatment effects of reducing administrative costs (30% Cost Reduction Treatment Group) and an increased refund of administrative costs (100% Cost Refund Treatment Group) relative to the 30% Cost Refund Treatment Group by income level for the subpopulation of taxpayers for whom we observe the true tax liability in the treatment year. Each column shows separate OLS regressions: Panel A displays the treatment effects for the sample of baseline donors and compliers, while Panel B shows estimations for the sample of baseline evaders. All regressions include a full set of controls for strata variables (based on age, gender, the tax amount due in the baseline including an indicator for missing values, and parish fixed effects). Robust standard errors in parentheses. ***, **, * denote significance level at 1, 5, 10 percent level, respectively.

Table A8: Entropy Reweighted Estimations of Treatment Effects

	Payment	Probability of			
	Amount (1)	Payment (2)	Donating (3)	Compliance (4)	Evading (5)
A: Baseline Donors and Compliers					
Effect of Cost Reduction	4.322** (1.918)	0.021 (0.034)	0.122*** (0.042)	-0.070 (0.043)	-0.052 (0.043)
Effect of Increased Cost Refund	2.717 (1.893)	0.044 (0.033)	0.020 (0.041)	0.057 (0.044)	-0.077* (0.042)
Number of Observations	1,129	1,129	709	709	709
Average in Reference Group	€ 24.63	0.683	0.289	0.371	0.341
B: Baseline Evaders					
Effect of Cost Reduction	-0.635* (0.346)	-0.013 (0.009)	0.002 (0.005)	-0.019** (0.008)	0.017* (0.010)
Effect of Increased Cost Refund	-0.699** (0.350)	-0.022*** (0.008)	0.001 (0.005)	-0.022*** (0.008)	0.021** (0.010)
Number of Observations	5,304	5,304	2,916	2,916	2,916
Average in Reference Group	€ 2.52	0.08	0.012	0.046	0.941

Notes: The table reports entropy reweighted estimations of treatment effects of reducing administrative costs (30% Cost Reduction Treatment Group) and an increased refund of administrative costs (100% Cost Refund Treatment Group) relative to the 30% Cost Refund Treatment Group. Each column shows separate OLS regressions: Panel A displays the treatment effects for the sample of baseline donors and compliers, while Panel B shows estimations for the sample of baseline evaders. Columns (3) to (5) use a smaller sample as compared to columns (1) and (2), because the dependent variables can only be constructed for the subpopulation of taxpayers for whom we observe the true tax liability in the treatment year. The entropy weights ensure that the subsamples in each panel have the same income distribution as the overall sample of individuals in the experiment. All regressions include a full set of controls for strata variables (based on age, gender, the tax amount due in the baseline including an indicator for missing values, and parish fixed effects). Robust standard errors in parentheses. ***, **, * denote significance level at 1, 5, 10 percent level, respectively.

Evangelical-Lutheran Church District of < Place >

Evangelical-Lutheran Church District of < Place >, Address

Mr/Mrs
First Name and Family Name
Street
Zip Code and City

< Place >, Date

LOCAL CHURCH TAX NOTICE 2016

Dear Mr/Mrs [addressee's family name],

Based on the state law regulating the church tax, the Evangelical-Lutheran Church District of < Place > raises a local church tax for the year 2016. The local church tax forms part of the general church tax and is therefore a compulsory payment. It serves as a local tax to finance parish expenditures.

< In Treatments 1 to 3: Additional paragraph on administrative cost reductions >

The local church tax is staggered according to income. Please self-assess your income using the adjoining schedule and transfer the tax owed within a month's time after receiving this tax notice. Please use the attached bank transfer form when making your payment. This helps us to identify your payment.

You will find further information on how the revenues of the local church tax are spent in the accompanying leaflet. We very much appreciate your cooperation.

With kind regards,
Your Church District Administration in < Place >

Yearly Income or Benefits in Euro	Local Church Tax in Euro
Up to Exemption Level (8,652)	-
8,653 to 9,999	5
10,000 to 24,999	10
25,000 to 39,999	25
40,000 to 54,999	45
55,000 to 69,999	70
70,000 and above	100

See back of this page for legal advice. In case of questions, please contact < phone number >.

< Bank transfer form, pre-filled with church's bank account number and taxpayer's name and tax identifier >