

Administrative Efficiency and Tax Compliance

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

November 1, 2017

Abstract

We investigate how administrative efficiency affects tax compliance. The context is a German church tax that builds on voluntary compliance. In a field experiment, we inform taxpayers about a reduction in collection costs. We find counteracting effects on compliance, depending on baseline tax morale: High tax morale types do not respond if costs are shifted to a third party, but increase tax payments in response to a true cost reduction. The latter finding establishes that reciprocity is a component of tax morale. Low tax morale types do not reciprocate cost reductions and free-ride on cost shifts.

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

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

1 Introduction

The recent literature on the determinants of tax compliance has discussed two main motives leading taxpayers to comply with the tax law. First, as modeled 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

*Department of Economics, University of Erlangen-Nuremberg. Cagala: tobias.cagala@fau.de; Rincke: johannes.rincke@fau.de; Tuset Cueva: amanda.tuset.cueva@fau.de. We have benefitted greatly from discussions with Matteo Cervelatti, Ulrich Glogowsky, Johannes Hermle, Dominika Langenmayr, Ricardo Perez-Truglia, Sebastian Sieglösch, Joel Slemrod, and seminar participants at various locations.

tax authority to collect and process third-party information on true tax liabilities (Blumenthal *et al.*, 2001; Kleven *et al.*, 2011; Pomeranz, 2015; Naritomi, 2014). Second, taxpayer behavior may also be driven by tax morale, i.e. by nonpecuniary motivations including forms of intrinsic motivation (like a duty to comply, warm glow, or altruism) and reciprocity concerns (Andreoni *et al.*, 1998; Luttmer and Singhal, 2014). While recent field studies have made considerable progress in measuring and understanding the importance of tax morale (Slemrod *et al.*, 2001; Fellner *et al.*, 2013; Hallsworth *et al.*, 2014; Del Carpio, 2014; Dwenger *et al.*, 2016), our knowledge of how nonpecuniary motivations can be used 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 want to test the conjecture that tax compliance is at least partly driven by a reciprocal response to the (perceived) competence of the government (Luttmer and Singhal, 2014). The evidence comes from a randomized field experiment testing for 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 not complying 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 and to describe the treatment responses for different baseline compliance types.

In the experiment, we implemented treatments manipulating the district-level administrative fixed cost of producing and mailing the tax notifications to church members. Treatment assignment was randomized at the level of the individual, and taxpayers were informed about the cost reductions in the annual tax notification. The treatments aim at disentangling two effects: a reciprocal response, and a crowd-out of voluntary tax payments caused by an implicit transfer to the local public good. The former effect captures a *positive* compliance response to the cost reduction as a signal of competence of the local administration. We identify the reciprocity effect by comparing a treatment communicating a real reduction of tax collection costs with one

referring to an equivalent shift of administrative costs to a higher level of the church administration. The crowding out effect results from the fact that reducing a fixed administrative cost by a certain amount is financially equivalent to a positive third party transfer to the local public good. The public-goods crowding-out literature suggests that this implicit transfer triggers a *negative* compliance response by crowding out voluntary tax payments (Warr, 1982; Roberts, 1984; Andreoni, 1989; Ribar and Wilhelm, 2002).

Our main result is that administrative efficiency has counteracting effects on compliance, depending on baseline tax morale. Types who show high tax morale in the baseline increase their tax payments in a reciprocal response to a true cost reduction, but do not show any crowd-out if a shifting of costs expands the pool from which the public good is financed. 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 it 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.

Relative to previous literature, this paper provides two main insights. First, we provide causal evidence that tax morale is at least partly shaped by reciprocity concerns. This implies that by signalling a concern for good governance, tax authorities can, in principle, increase compliance at least of some taxpayers. Previous literature has documented positive correlations between institutional quality and survey-based measures of tax morale (for a review, see OECD 2013),¹ or has produced related, but indirect evidence suggesting that taxpayers reciprocate good governance. For instance, Timmons and Garfias (2015) show that revenues from the municipal property tax in Brasil decrease if audits reveal corruption by local officials. Gonzalez-Navarro and Quintana-Domeque (2016) exploit exogenous variation in local public good provision and find that improved provision increases satisfaction with the local government,

¹An 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.

while [Doerrenberg and Peichl \(2017\)](#) 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 the 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.² As argued by [Luttmer and Singhal \(2014\)](#), this might have to do with those studies using ‘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 are interacting with. Moreover, we demonstrate that it is unlikely to detect taxpayers’ reciprocal responses without differentiating between different motivational types.

A second insight relates to the strong heterogeneity in treatment responses observed in our experiment. The stark differences in treatment responses between low-tax morale and high-tax morale types suggests that it is crucial for tax authorities to develop a precise understanding how different motivations for tax compliance translate into actual compliance responses. The heterogeneity documented in this paper is complementary to [Dwenger et al. \(2016\)](#), who also study the local church tax in Germany and show that strongly intrinsically motivated taxpayers respond positively to compliance rewards, whereas weakly motivated taxpayers *decrease* their compliance even further. Taken together, the evidence suggests that policies reducing tax collection costs or rewarding taxpayers for compliant behavior can be effective. However, it also highlights the risk that alluding to individuals’ tax morale may trigger adverse compliance responses among individuals with a weak tax morale to begin with.

The heterogeneous treatment responses also link our work to a substantial literature on the role of charities’ overhead costs for charitable giving. Consistently, this literature has shown that potential donors are averse to financing charities’ over-

²Notable exceptions include [Hallsworth et al. \(2014\)](#), 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](#)).

head cost (Okten and Weisbrod, 2000; Gneezy *et al.*, 2014; Caviola *et al.*, 2014; Meer, 2014).³ 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. It is noteworthy that the literature on charitable giving almost exclusively focuses on warm-list individuals (i.e., subjects who have revealed some willingness to donate in the baseline). This focus has prevented the literature from studying the effect of reduced overhead costs on individuals with a low motivation to give. Derived from a setting that bridges the charitable giving and the tax setting, our findings suggest that among cold-list individuals, communicating lower administrative costs may crowd out contributions.

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

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 in 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, there exists 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, it 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

³Cagala *et al.* (2016) report evidence from the laboratory showing that an administrator who decides to expropriate part of public goods contributions sharply diminishes individuals' contributions.

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.

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 (see the appendix) to all resident church members above age 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 do anything to 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 not paying the local church tax, or underpaying it, do not face any legal consequences. Moreover, access to church services is also 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. \(2016\)](#) 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. As a consequence, we can use local church tax payments as a measure of taxpayer's 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 due. Such overpayments are not refunded, but treated like donations reaching the church through other

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

channels.⁶

3 Experimental Design, Data, and Methods

3.1 Experimental Design

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 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,⁸ we assigned each taxpayer to one of four treatment groups.

In the *control group*, the tax notification (see the appendix) was very similar to the one mailed in previous years.⁹ The control group allows us to compare compliance behavior in the treatment year with behavior in baseline years. 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 transfer the appropriate tax amount based on a self-assessment of income.

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).

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

⁸Strata were defined using age quartiles, gender, dummies for the taxpayers' tax bracket according 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.

⁹While 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.

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 informing the recipient about a reduction of tax collection costs born by the local church administration. The additional paragraph in the *30% cost reduction* treatment reads:

Please note: In collaboration with various service providers, this year we were able to reduce the administrative cost associated 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.

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

We made sure that the changes in administrative costs born 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 services 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 born costs communicated in the respective notifications, i.e. each individual notification carried a change of administrative cost born by the district administration in accordance with the content of the treatment paragraph.

Two more details about the cost reductions in our design are worth noting. First, the cost reduction communicated in the *30% cost reduction* treatment could actually be achieved by the district administrations 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 competitive 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 cost reduction by €0.72 per notification amounts to an average increase of net-of-cost revenue by 20.6 percent.

3.2 Data and Sampling

3.2.1 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. In the field experiment, we did not consider 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.¹⁰

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 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.

3.2.2 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 competence. 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. This is individuals who

¹⁰We also excluded employees of the Protestant Church from the experiment.

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. This is 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.¹¹

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

3.2.3 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.¹² As a consequence, we do not observe the individual's income in church tax records. The church tax records also have missing values due to late filers.¹³ 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

¹¹All 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.

¹²The church collects the state church tax from wage earners through automatic tax withholding, very similar to withholding of the federal income tax.

¹³Of 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.

behavior or not. We use a total of four different outcomes: the extensive margin responses to our treatments are identified using the probability of payment, the probability of donation, and the probability of evasion. The combined effect of extensive and intensive margin responses is captured by estimations using the payment amount as dependent variable.¹⁴ When using either the probability of 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 8,617 individuals in terms of their baseline compliance behavior, 1,525 as baseline donors or compliers (Column 3), and 7,092 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.

3.2.4 Balancedness and Persistence of Compliance over Time

Table A2 in the appendix shows descriptive statistics for the full sample (Panel A), baseline donors and compliers (Panel B), and baseline evaders (Panel C). The p -values for an F -test of joint significance of taxpayer characteristics in explaining treatment assignment show that the treatments are balanced.

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 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.

¹⁴We 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, 80.2 percent were timely payments.

It turns out that 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. For average payment amounts, the figures are € 3.95 and € 4.28, respectively. 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.

3.3 Separating Reciprocity from Crowding Out

We derive our main results from simple OLS regressions. For the total response, we estimate the specification

$$y_i = \alpha + \sum_k \beta^k T_i^k + x_i \gamma + \epsilon_i, \quad (1)$$

where y_i denotes the amount paid, T_i^k denotes the indicator for treatment k , x_i is a vector comprising the strata variables and parish fixed effects, and ϵ_i is an error term. For the extensive margin responses (probability of payment, probability of donation, and probability of evasion), we estimate a linear probability model. The left-hand side of equation (1) then becomes $Prob(\text{payment}_i > 0)$, $Prob(\text{payment}_i < \text{tax liability}_i)$, or $Prob(\text{payment}_i > \text{tax liability}_i)$, respectively.

The main difficulty of estimating economically meaningful treatment effects lies in the fact that if a fixed administrative cost per taxpayer is reduced by a certain amount, this is financially equivalent to a positive lump-sum transfer from a third party to the local public good. This complicates the identification of taxpayers' reciprocal response, because the public-goods crowding-out literature going back to [Warr \(1982\)](#), [Roberts \(1984\)](#) and [Andreoni \(1989\)](#) suggests that conceptually, such an implicit transfer crowds out voluntary payments to the public good. In our context, any crowd-out implies a negative compliance response that counteracts the reciprocal response. To separate the two effects, we use (1) and perform two main treatment comparisons. While the first comparison aims at identifying the reciprocal response, the second captures the crowd-out. We discuss both estimations in turn.

To identify the reciprocity effect, we restrict attention to the sample of individuals who received either the *30% reduction* or the *30% refund* treatment. Using *30% refund* as the omitted reference category, β measures the average treatment 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. Importantly, in this treatment comparison the implicit transfer to the local public good that is associated with lowering the local administration's expenses for preparing and mailing the tax notifications is held constant. In other words, from a local perspective the material consequences of what is communicated in the tax notifications are constant between treatments.¹⁵ The only difference between the *30% reduction* and the *30% refund* treatments is thus the communicated reason for the change in costs born by the district. Our interpretation of β as a measure of taxpayers' reciprocity rests on the assumption 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 competence compared to a pure shifting of costs to a third party.

To identify the crowd-out effect, we compare the *100% refund* to the *30% refund* treatment. In this case, the only difference between treatments is the size of the implicit transfer associated with the shifting of costs to the state-level administration. Using again *30% refund* as the omitted reference category, β identifies the crowd out that is associated with increasing the transfer to the public good from 30% to 100% of the fixed administrative cost.

4 Empirical Results

4.1 Reciprocal Response to Administrative Cost Reduction

We begin the discussion of our empirical results with taxpayers' reciprocal response to the administrative cost reduction (30% reduction vs. 30% refund). Table 2 reports the corresponding findings. We consider four outcomes. The payment amount captures the total effect and is displayed in column (1). Columns (2) to (4) report three

¹⁵We make the assumption here that taxpayers' perceptions about possible adjustments in grants from the state level administration to the districts are the same between treatments.

different extensive margin responses: the probability of payment, the probability of donation, and the probability of evasion. The estimations reported do not include any control variables.¹⁶ In the appendix, we report all estimations shown in the main tables including controls for strata variables.

Panel A displays estimations for the sample of baseline donors and compliers. From column (1) we see that the total effect is positive and significant. Baseline donors and compliers receiving the tax notice communicating a real reduction in administrative costs pay € 5.51 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% refund treatment group, this is an increase by 22.4 percent. From column (2) we note that the probability of payment is not affected: the coefficient is small and not statistically different from zero. This implies that the total effect is driven by a positive intensive margin response. Columns (3) and (4) add further details: 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 30% refund treatment group, the probability of donating increases by 13.5 percentage points, or 46.7 percent (column 3). In contrast, the probability of evading is not affected (column 4). The latter finding is explained by the fact that among baseline donors and compliers, the bulk of evasion decisions (83.4 percent) take the form of full evasion. Given that we do not find any effect on the probability of payment, there remains little room for a treatment effect on the probability of evading as the broader measure of non-compliance. Taken together, Panel A shows a significant reciprocal response to a reduction of administrative costs among high tax morale types.

Panel B uses the sample of baseline evaders and reports the same estimations as before. Column (1) shows a weakly significant *negative* treatment effect on the payment amount. In contrast, the probability of payment and the probability of donating are unaffected (columns 2 and 3). Finally, the estimation in column (4) shows a weakly significant *positive* effect on the probability of evading. We conclude that in contrast

¹⁶There is one exception: 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. See Table A1 for details.

to high tax morale types, taxpayers with low baseline tax morale do not reciprocate administrative cost reductions. If anything, the cost reduction signal triggers a further decrease in tax compliance in this group. One possible interpretation of a negative compliance response 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 cost reduction relative to the cost shift. The implicit transfer to the local public good associated with the difference in present values could trigger a crowding out of voluntary tax payments.

Panel C in Table 2 reports the findings for the full sample. For the total effect, it turns out that the positive reciprocity effect among the relatively few high tax morale types gets diluted among the majority of taxpayers with low tax morale and is no longer visible (column 1). In contrast, the particularly strong effect on the probability of donating among baseline donors and compliers is still discernable in the full sample. In total, Table 2 highlights the importance of being able to distinguish between different baseline compliance types when studying taxpayers' reciprocal response to signals of good governance.

4.2 Crowding Out of Voluntary Tax Payments

We proceed with the findings on the crowding out of voluntary tax payments (100% refund vs. 30% refund). Table 3 displays the estimation results. From Panel A we see that for high tax morale types, none of the estimates for the four different outcomes is significantly different from zero. Taken together with Table 2, this means 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 public good. In the light of the public-goods crowding-out literature discussing impure altruism as a motivation for voluntary contributions ([Andreoni, 1989](#); [Ribar and Wilhelm, 2002](#)), this implies 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 separate 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 signal of the administration's competence leads to a positive reciprocity response and zero crowd-out. This suggests a potential for policy makers to improve tax compliance among high tax morale individuals through signalling improvements in governance.

From Panel B, we see that low tax morale types again behave very differently from high tax morale types. Column (1) shows a weakly significant negative total effect. The small size of the effect should be seen in relation to the average payment amount in the 30% refund treatment group of just €2.52. Column (2) demonstrates that part of this response is a negative extensive margin response: the average probability of payment is reduced by 2.2 percentage points, a reduction of 27.5% relative to the payment probability of 8 percent in the 30% refund treatment group. We do not see any treatment effect on the probability of donating, a finding that is not surprising given that in the 30% refund treatment group, only 1.2 percent of baseline evaders behave as donors. The coefficient in column (4) confirms the finding that the increased cost refund has reduced compliance in the group of baseline evaders: the probability of evasion is increased by 2.2 percent. Although this estimate comes from a narrower sample (only taxpayers for whom we observe the tax amount owed in the baseline *and* in the treatment year), the coefficient mirrors the one from column (2). All in all, Panel B demonstrates that an implicit third-party transfer to the public good further reduces tax compliance among taxpayers with low tax morale to begin with. Combining this insight with the fact that low tax morale types do not reciprocate true cost reductions, we arrive at the conclusion that signalling improvements in governance is likely to *reduce* tax compliance among individuals with weak tax morale.

Our results demonstrate that tax payments of baseline evaders are crowded out, but interpreting this finding in terms of the underlying preferences is difficult, for two reasons. First, in the 30% refund treatment group, the payment probability is only 8 percent, meaning that 92 percent of baseline evaders are in a corner solution. We simply cannot say whether or not the increased cost refund has further crowded out their willingness to pay the local church tax. Second, taxpayers with low tax morale who end up in the 100% refund treatment group and want to reduce their tax payment

relative to the amount they would have paid if they were in the 30% 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. We therefore abstain from interpretations to what extent the behavior of low tax morale types is motivated by altruism or warm glow.

Panel C completes the description of findings for the full sample. We note that none of the treatment effects is significantly different from zero. However, as in Table 2, the insignificance of the full-sample results masks a strong heterogeneity in treatment responses between baseline compliance types.

To sum up, the evidence from our field experiment establishes the following main findings. First, we provide causal evidence that tax morale is shaped by reciprocity concerns. However, only individuals with high tax morale to begin with reciprocate signals of good governance. Given that the implicit transfer to the public good that is associated with reductions of administrative costs does not crowd out voluntary compliance among high tax morale types, the reciprocal response implies that governments can improve tax compliance among highly motivated taxpayers by signalling improvements in governance. Second, we also identify an effect that counteracts the reciprocity effect among high tax morale types. The negative compliance response comes from low tax morale types who do not reciprocate cost reductions, but decrease their compliance in response to implicit transfers to the public good.

5 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 reciprocity 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 the reciprocity response, 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 compliance 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 through signalling improvements in governance. Our results on high tax morale types stand in marked contrast with the those on low tax morale types: baseline evaders do not reciprocate a reduction in tax collection costs, and they decrease their compliance further in response to a pure shift of collection costs to a third party. This suggests that signalling improvements in governance is likely to crowd out tax compliance among low tax morale types. To the best of our knowledge, our paper is the first to provide causal evidence on the heterogeneity in treatment responses with respect to baseline tax morale.

Given the counteracting responses in our experiment, we conclude that it is difficult to predict the overall compliance responses to improvements in administrative efficiency. Our findings suggest that it 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. We conclude that policy makers should be cautious when using signals of good governance to improve tax compliance: the policy is likely

to be effective among high tax morale taxpayers, but might backfire when applied to taxpayers with low tax morale.

While our contribution speaks first and foremost to the literature on tax morale, there is also a close link 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 payments. 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, in 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 findings suggest that the findings on charitable giving motivations 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. (2016). Does Corruption Affect the Private Provision of Public Goods?, mimeo.
- 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.
- DEL CARPIO, L. (2014). Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru, mimeo.
- DOERRENBERG, P. and PEICHL, A. (2017). Tax Morale and the Role of Social Norms and Reciprocity: Evidence from a Randomized Survey Experiment, mimeo.
- 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.

- GONZALEZ-NAVARRO, M. and QUINTANA-DOMEQUE, C. (2016). Paving Streets for the Poor: Experimental Analysis of Infrastructure Effects. *Review of Economics and Statistics*, **98** (2), 254–267.
- HALLSWORTH, M., LIST, J. A., METCALFE, R. D. and VLAEV, I. (2014). The Behavioralist As Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance, NBER Working Paper No. 20007.
- 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.
- 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, mimeo.
- 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.
- 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 and Compliers		Baseline Evaders	
	All (1)	Conditional on Tax Base 2016 Being Observed (2)	All (3)	Conditional on Tax Base 2016 Being Observed (4)	All (5)	Conditional on Tax Base 2016 Being Observed (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: Treatment Effect of Administrative Cost Reduction
(30% Reduction vs. 30% Refund)**

	Payment Amount (1)	Probability of Payment (2)	Probability of Donating (3)	Probability of Evading (4)
Panel A: Baseline Donors and Compliers				
Effect of Cost Reduction	5.511** (2.359)	0.013 (0.034)	0.135*** (0.044)	-0.052 (0.043)
Number of Observations	740	740	468	468
Average Outcome in Omitted Reference Group	€24.63	0.683	0.289	0.341
Panel B: Baseline Evaders				
Effect of Cost Reduction	-0.588* (0.355)	-0.012 (0.009)	0.002 (0.005)	0.017* (0.010)
Number of Observations	3546	3546	1964	1964
Average Outcome in Omitted Reference Group	€2.52	0.080	0.012	0.941
Panel C: Full Sample				
Effect of Cost Reduction	0.273 (0.242)	-0.001 (0.006)	0.021** (0.011)	0.009 (0.015)
Number of Observations	12156	12156	2651	2651
Average Outcome in Omitted Reference Group	€3.51	0.139	0.069	0.823

Notes: The table reports treatment effects of reducing administrative costs (30% Reduction Treatment Group vs. 30% Refund Treatment Group). Each column shows three 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) and (4) 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 3: Treatment Effect of Increased Cost Refund
(100% Refund vs. 30% Refund)**

	Payment Amount (1)	Probability of Payment (2)	Probability of Donating (3)	Probability of Evading (4)
Panel A: Baseline Donors and Compliers				
Effect of Increased Cost Refund	2.761 (2.233)	0.022 (0.033)	0.018 (0.042)	-0.063 (0.043)
Number of Observations	764	764	473	473
Average Outcome in Omitted Reference Group	€24.63	0.683	0.289	0.341
Panel B: Baseline Evaders				
Effect of Increased Cost Refund	-0.684* (0.363)	-0.022** (0.009)	0.000 (0.005)	0.022** (0.010)
Number of Observations	3539	3539	1924	1924
Average Outcome in Omitted Reference Group	€2.52	0.080	0.012	0.941
Panel C: Full Sample				
Effect of Increased Cost Refund	0.324 (0.243)	-0.000 (0.006)	0.003 (0.010)	-0.001 (0.015)
Number of Observations	12173	12173	2613	2613
Average Outcome in Omitted Reference Group	€3.51	0.139	0.069	0.823

Notes: The table reports treatment effects of an increased refund of administrative costs (100% Refund Treatment Group vs. 30% Refund Treatment Group). Each column shows three 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) and (4) 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 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 (Part I)

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

Notes: The table presents balancedness checks. Panel A 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 control group. 95% confidence intervals in squared brackets.

Table A2: Balancedness Checks (Part II)

	N	Female	Age	Indicators for Baseline Tax Liability							No info	F-Test <i>p-value</i>
				€ 0	€ 5	€ 10	€ 25	€ 45/€ 50	€ 70	€ 100		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(10)	(11)	(12)	(13)
Panel B: Subsample: 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]	0.407 [0.36; 0.46]	0.222 [0.18; 0.26]	0.139 [0.10; 0.17]	0.051 [0.03; 0.07]	0.073 [0.05; 0.10]	-	-
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]	0.411 [0.36; 0.46]	0.244 [0.20; 0.29]	0.134 [0.10; 0.17]	0.058 [0.03; 0.08]	0.060 [0.04; 0.08]	-	0.858
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]	0.421 [0.37; 0.47]	0.224 [0.18; 0.27]	0.120 [0.09; 0.15]	0.053 [0.03; 0.08]	0.061 [0.04; 0.09]	-	0.999
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]	0.414 [0.36; 0.46]	0.247 [0.20; 0.29]	0.113 [0.08; 0.14]	0.064 [0.04; 0.09]	0.067 [0.04; 0.09]	-	0.962
Panel C: Subsample: 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]	0.383 [0.36; 0.41]	0.300 [0.28; 0.32]	0.113 [0.10; 0.13]	0.060 [0.05; 0.07]	0.070 [0.06; 0.08]	-	-
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]	0.386 [0.36; 0.41]	0.306 [0.28; 0.33]	0.124 [0.11; 0.14]	0.051 [0.04; 0.06]	0.069 [0.06; 0.08]	-	0.997
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]	0.387 [0.36; 0.41]	0.310 [0.29; 0.33]	0.115 [0.10; 0.13]	0.052 [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]	0.390 [0.37; 0.41]	0.303 [0.28; 0.32]	0.114 [0.10; 0.13]	0.049 [0.04; 0.06]	0.068 [0.06; 0.08]	-	0.995

Notes: The table presents balancedness checks. Panel B refers to the sample of taxpayers who have paid the amount due or who have overpaid in the baseline (baseline donors and compliers), and Panel C refers to the sample of taxpayers who have underpaid in the baseline (baseline evaders). 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 control group. 95% confidence intervals in squared brackets.