Do Taxes Crowd Out Intrinsic Motivation? Field-Experimental Evidence from Germany

Pierre C. Boyer, Nadja Dwenger, Johannes Rincke*

May 4, 2015

Abstract

This paper studies how imposing norms on contribution behavior affects individuals' intrinsic motivation. We consider the church levy, which the Catholic Church in Germany collects as a charitable donation, despite the fact that the levy is legally a tax. We design a randomized field experiment with treatments informing individuals that the levy is a tax. Guided by a theoretical model, we demonstrate that treatment effects differ across motivational types. Among weakly intrinsically motivated individuals, communicating a legal norm results in a significant crowd-out of intrinsic motivation. In contrast, strongly intrinsically motivated individuals do not show any treatment response.

JEL Codes: C93, D03, H26, H41

Keywords: intrinsic motivation, crowding out, charitable giving, taxes, public goods, randomized field experiment

*Boyer: University of Mannheim (e-mail: pierre.boyer@uni-mannheim.de); Dwenger: Max Planck Institute for Tax Law and Public Finance (e-mail: nadja.dwenger@tax.mpg.de); Rincke: University of Erlangen-Nuremberg (e-mail: johannes.rincke@fau.de). We thank Armin Falk, David Huffman, John List, Emmanuel Saez, Monica Singhal, Michèle Tertilt, Christian Traxler and numerous seminar participants for useful comments. We are grateful for financial support from the Schoeller Foundation and the Emerging Field Initiative of the University of Erlangen-Nuremberg. All errors remain our own.

1 Introduction

How do external incentives affect behavior? For a long time economists have focused on how external incentives shape individuals' *extrinsic motivation*. More recently, they have started to integrate into their models the view that external incentives can actually backfire by crowding out individuals' *intrinsic motivation* (Bénabou and Tirole, 2003, 2006; Sliwka, 2007).[1] So far, the experimental literature on the extrinsic-intrinsic crowd-out has produced mixed results (see Gneezy *et al.* 2011 for a survey). Gneezy and Rustichini (2000a) find that small monetary incentives impair individuals' intrinsic motivation, and Falk and Kosfeld (2006) demonstrate that agents tend to reduce their effort in response to a principal's decision to control their performance. In contrast, Lacetera *et al.* (2012), Ashraf *et al.* (2012), and Chetty *et al.* (2014) conclude that even in the case of small monetary rewards, any potential crowd-out is dominated by the positive effect of the external incentive.[2]

This paper adds to the existing literature in two important dimensions. First, while most previous field work on the extrinsic-intrinsic crowd-out has considered the effects of *(monetary) rewards*, we focus on the role of taxes as *externally imposed norms on contribution behavior*, as opposed to voluntary contributions. Second, we consider a setting where we can very accurately measure individuals' initial strength of intrinsic motivation. This allows us to study heterogenous responses across different motivational types when voluntary contributions are transformed into compulsory tax payments (and vice versa).

We implement our research design in the context of the local church levy in Germany, an institutional setting ideally suited to study how taxes affect individuals' intrinsic motivation. We focus on an urban area in Bavaria where the Catholic Church has always collected the local church levy as a charitable donation on a purely voluntary basis, despite the fact that the church levy is legally a tax on all church members.[3] Starting from a pure voluntary-giving baseline without any external incentives, we conduct a randomized field experiment with letter treatments informing individuals that the church levy is in fact a tax. Thereby, we can study how payment behavior changes if public goods are financed through compulsory tax payments instead of voluntary contributions. While this one-of-a-kind feature of the local church levy provides a unique opportunity to obtain

[1] In psychology and sociology, the extrinsic-intrinsic crowd-out has been discussed somewhat earlier (Deci, 1971, 1975).

[2] The role of external incentives has been extensively studied in the context of labor markets. The evidence is yet inconclusive (Rebitzer and Taylor, 2011).

[3] Dwenger *et al.* (2014) study the local church tax collected by the Protestant Church in a different metropolitan area in Bavaria. In contrast to our setting, the Protestant Church has always highlighted the fact that the local church tax is a compulsory payment. Dwenger *et al.* (2014) exploit this feature to study social and economic incentives for tax compliance.

novel insights on the crowd-out of intrinsic motivation, there is of course a potential trade-off with external validity which we discuss in the paper.

In our field experiment we randomly assign a total of almost 40,000 individuals to a control group and three different treatment groups: a compulsory tax, a voluntary tax, and a donation letter group. The *compulsory tax* letter comes close to a tax notice in other settings by highlighting the fact that the church levy is a legally binding tax. The letter encourages overpayments and explains that payments which exceed the tax owed are treated as charitable donations. The *voluntary tax* letter communicates the status quo in the baseline. It states the fact that the church levy is legally a tax, but informs recipients that the church administration considers the levy a charitable contribution on a purely voluntary basis. Both tax letters are naturally compared to the *donation* letter group. The donation letter states that the church considers the levy a purely voluntary contribution. As the tax letters mention the amount most individuals owe according to the tax law, the donation letter refers to the same amount as a suggested donation. To measure the effect of mentioning the amount and thus providing a reference point, we compare the donation letter group to the control group which receives the same letter as in previous years (voluntary contribution, no suggested amount).

The empirical analysis is guided by a simple theoretical model of contribution behavior under different collection regimes, ranging from voluntary contributions to a tax that might be implemented as a more or less binding norm. Our first prediction is that non-contributors to the charity in a voluntary contribution regime strictly increase their contributions when these are collected under a binding tax norm. For intrinsically motivated individuals, the model predicts responses depending on the strength of individuals' motivation: imposing a tax on weakly intrinsically motivated individuals crowds out their 'warm-glow' motivation, but making the tax norm more binding (partially) compensates for this effect. Highly intrinsically motivated types may also be subject to a crowd-out, but in contrast to the weakly intrinsically motivated, they do not respond if the tax becomes more binding.

It is a major advantage of our setting that we can measure the strength of intrinsic motivation in the baseline. This enables us to study the heterogeneity in treatment responses predicted by the theoretical model.[4] To do so we exploit the fact that there were no external incentives in place initially such that any contributions made prior to treatment necessarily reflect individuals' intrinsic motivation. We use administrative data on individual contributions in the eight years prior

[4]The only field-experimental studies we are aware of that look at differences between motivational types when studying the extrinsic-intrinsic crowd-out are Ashraf *et al.* (2012) and Huffman and Bognanno (2014). Ashraf *et al.* (2012) derive their measures of agents' motivation from behavior in a dictator game and from survey responses. Huffman and Bognanno (2014) analyze heterogeneous responses to incentives in a real work setting and distinguish between motivational types by means of a post-treatment survey.

to treatment to determine the relative frequency of pre-treatment contributions as a straightforward measure of individuals' baseline motivation. We distinguish between two main motivational types. A first group consists of individuals who never contributed in the baseline. Individuals of this type reveal that their intrinsic motivation is too low to trigger any financial contribution. We call these individuals the *baseline non-contributors*. A second group consists of individuals who have contributed at least once, thereby revealing *some* baseline intrinsic motivation. We call these individuals the *intrinsically motivated* and use the baseline probability of contributing as a continuous measure of their intrinsic motivation.

The findings from the field experiment reveal a distinct heterogeneity in treatment responses. First, individuals with regular baseline contributions (the strongly intrinsically motivated) on average do not show any response to the information that the church levy is a tax. This finding stands in stark contrast to the behavior of individuals who contributed only occasionally in the baseline (the weakly intrinsically motivated): individuals in this group significantly reduce their payments in response to the voluntary tax letter, but do not show any net response to the compulsory tax treatment. This behavior is consistent with the notion that imposing norms on contribution behavior crowds out intrinsic motivation, but that a sufficiently binding tax norm compensates the crowd-out. The crowd-out identified by our field experiment is economically significant: in the voluntary tax treatment, subjects from the bottom of the distribution of baseline motivation are about 20% less likely to make a positive contribution compared to the control group. Finally, baseline non-contributors significantly increase their payments if they receive the compulsory tax letter but do not respond to the voluntary tax letter that communicates the existence of a non-binding legal norm. Again this is consistent with our theoretical predictions.

The identification of the crowd-out effect rests on the assumption that baseline church levy contributions provide us with a reliable measure of intrinsic motivation. To cross-validate our measure of intrinsic motivation, we conduct an extensive post-treatment survey among the treated individuals. The survey extracts information on respondents' relation to their parish, church attendance habits, and general willingness to donate or volunteer. From this we derive alternative measures of individuals' intrinsic motivation.

Two main results emerge from the post-treatment survey. First, we show that baseline contribution behavior is strongly and positively correlated with each of the alternative measures of intrinsic motivation. This supports our behavior-based measure of intrinsic motivation in the field experiment. Second, we replicate the estimates of the crowd-out effect in the sample of survey respondents using the survey-based measures of motivation and find all results from the field experiment confirmed.

Interestingly, the heterogeneity in treatment responses results in a situation where the average treatment effects are small and insignificant, despite the fact that several subgroups of individuals show strong behavioral responses. One of our main conclusions is therefore that in order to predict how individuals respond to external incentives, it is essential to take into account the heterogeneity in motivational types.

Our contribution relates to various strands of literature. First, we present new evidence on the extrinsic-intrinsic crowd-out, relating our work to Frey and Oberholzer-Gee (1997), Gneezy and Rusticcini (2000a,b), Falk and Kosfeld (2006), and Mellström and Johannesson (2008) (who find evidence for crowding out) and Lacetera *et al.* (2012), Ashraf *et al.* (2012), and Chetty *et al.* (2014) (who conclude that it is of minor importance).[5] In contrast to most previous work on the extrinsic-intrinsic crowd-out, we study a context where social image concerns are of little or no importance: individual church levy contributions are strictly private, and the church administration collecting the payments does not inform local parishes about individual contributions. This differentiates our work conceptually from contexts where external incentives dilute the signaling value of prosocial behavior and thereby reduce individuals' social image motivation for prosocial activities (Ariely *et al.*, 2009; Friedrichsen and Engelmann, 2014).[6]

Second, by experimentally shifting the framing from donation to tax, we bridge the gap between the charitable giving and the tax compliance literature. While it is well established that intrinsic motivation in the form of 'warm glow' is important for charitable giving (Andreoni, 1989, 1990), the role of intrinsic motivation in the context of tax compliance is less clear (Andreoni *et al.*, 1998; Luttmer and Singhal, 2014). While some studies have shown that tax morale and internalized social norms can be relevant (Del Carpio, 2013; Pruckner and Sausgruber, 2013; Fellner *et al.*, 2013; Hallsworth *et al.*, 2014), most of the field-experimental literature on tax compliance pioneered by Slemrod *et al.* (2001) and Blumenthal *et al.* (2001) focuses on external incentives. Dwenger *et al.* (2014) contrast policies aiming at a stronger enforcement of taxes with reward-based approaches in a field experiment. They conclude that intrinsic motivation is important to understand tax compliance behavior, but that the positive effects of tax enforcement overcompensate any associated loss in intrinsic motivation. Taken together, our results and the findings of the tax compliance literature suggest that imposing a tax norm as such crowds out intrinsic motivation, but that the adverse effects of increasing the level of enforcement given a tax frame are modest.

Finally, from a methodological point of view, we add to an expanding literature using field

[5]The crowd-out studied in this literature is conceptually different from the one discussed by Andreoni (1993), who explores if government contributions towards privately provided public goods crowd out private contributions.

[6]Our design also avoids confounding factors like retaliation or loss of morale (Esteves-Sorensen *et al.*, 2013).

experiments to study charitable giving (Falk, 2007; Landry *et al.*, 2010; Huck and Rasul, 2011; DellaVigna *et al.*, 2012).

The remainder of the paper is organized as follows. Section 2 discusses the institutional background. The theoretical framework is presented in Section 3. The design of the field experiment and the data are discussed in Section 4. The results from the field are presented in Section 5, while Section 6 discusses the post-treatment survey. Section 7 concludes.

2 Institutional Background: The Local Church Levy

This study focuses on the *local church levy* (Kirchgeld) which is collected both by the Catholic and the Protestant Churches in Germany.[7] The church levy is the local component of overall church finances in Germany and is raised in addition to the general church tax, which is collected by state tax authorities on behalf of the church. In contrast to the general church tax, the collection of the local church levy falls in the responsibility of the local parishes and is therefore highly decentralized. In conjunction with the Catholic Church, we implement our field experiment in an urban area in Bavaria where the local church levy is raised jointly by 29 local parishes forming a church district.

It is of key importance for our study that the church district has always handled the local church levy as a charitable donation on a purely voluntary basis. To ask for the church levy donation, the district administration sends a solicitation letter to all full-age members once a year, typically in March/April.[8] The letter asks for a donation to the district's church levy funds and informs church members that the funds are mainly used to co-finance building measures undertaken by the local parishes.[9] The purpose of the solicitation letter is communicated in the first paragraph of the letter in a straightforward manner, stating that "[...] as every year, we kindly ask you herewith for your local church levy contribution. [...] The church district considers the church levy a contribution equivalent to a charitable donation."[10] Attached to the letter is a bank transfer form pre-filled with the church district's bank account information. In order to contribute, church

[7]Regions of Germany were the local church levy is raised include Bavaria, Saxony, Lower Saxony, and Rhineland-Palatinate.

[8]Married couples in which both spouses are members of the Catholic Church receive only one letter, with the husband figuring as addressee. We account for this mailing pattern and stratify our samples according to household type (see Section 4.1).

[9]Typical building measures are reconstruction works on existing churches, clergy houses, and parish centers. Examples of recent measures co-financed from local church levy funds are shown in a leaflet accompanying the solicitation letter. The leaflet is kept identical across all treatment groups in the experiment and is very similar to the leaflet sent out together with the solicitation letter in earlier years.

[10]Pre-treatment, the solicitation letter was very similar across years. The version cited here was used in the last pre-treatment year 2012.

members simply need to add their own bank account information together with the amount they intend to give, and initiate a regular bank transfer. Given the framing as a voluntary contribution, it is not surprising that few people pay the levy: in the baseline, about 9% of church members respond to the letter with a bank transfer.

The practice of collecting the church levy as a charitable donation stands in sharp contrast to the underlying legal framework that entitles major religious communities in Germany to raise church taxes from their members. Regarding the local church levy, the Bavarian church tax law and the corresponding regulations clearly state that the church levy is a tax and that local parishes are responsible for collecting the levy from their members. According to the statutes of the church district under consideration, the church levy is a compulsory payment depending on church members' gross income (including wages, business income, capital income, pensions, etc.). The church levy ranges from €2.5 to €15 for individuals exceeding the exemption level of €1,800 annual income.[11] Table A1 in the Appendix demonstrates that in practice, the vast majority of individuals subject to the church levy owes the maximum amount of €15 (77% of single and 66% of married taxpayers).

In our experiment, we exploit the unique feature that the church levy is handled as a charitable contribution despite the fact that it is legally a tax. In this specific institutional context, we can truthfully shift from voluntary contributions to compulsory tax payments—a variation that would be very difficult to implement in most other settings. Hence, the local church levy provides us with an ideal testing ground for studying a potential crowd-out of intrinsic motivation in a tax vs. voluntary contribution setting.

The contrast between the practice of collecting the church levy as a charitable donation and the legal tax framework is due to various reasons. First, the local church administration has no information on church members' incomes and thus cannot enforce the local church levy as an income-dependent tax.[12] Second, by framing the local church levy as a charitable donation, the district administration manages to collect average payments (conditional on paying) which exceed the maximum tax amount of €15: in 2012, the average payment (conditional on paying) was €33.82. Hence, the decision not to collect the church levy as a tax reflects the tradeoff between

[11] The full schedule is as follows (with I being annual income, and $\hat{d}$ representing the amount due):
$I <$ €1,800: $\hat{d} =$ €0.00
€1,800 $\leq I <$ €5,000: $\hat{d} =$ €2.50
€5,000 $\leq I <$ €7,500: $\hat{d} =$ €5.00
€7,500 $\leq I <$ €10,000: $\hat{d} =$ €7.50
€10,000 $\leq I <$ €12,500: $\hat{d} =$ €10.00
€12,500 $\leq I <$ €15,000: $\hat{d} =$ €12.50
$I \geq$ €15,000: $\hat{d} =$ €15.00.

[12] Even if income information was available, enforcement would hardly be cost-efficient given the modest size of tax liabilities.

individuals' intrinsic and extrinsic motivations. In the field experiment, we study precisely this margin by varying the framing of the church levy from charitable donation to tax.

What is important to note is that both charitable donations and church levy payments are tax deductible and treated equally in personal income taxation. Thus, individuals face the same costs of payment in whether the church collects the levy as a tax or as a voluntary contribution. A few more institutional details are important in our setting. First, given the donation framework used in the baseline, contributing nothing or underpaying relative to the amount legally owed has no consequences whatsoever.[13] Second, information on individual contributions remains strictly private. While the personal interaction between church members and the clergy or other church staff takes place at the level of the local parishes, the church district administration typically does not interact with individual church members. The church district collects the church levy and distributes the revenues to local parishes, but it does not provide information on individual church levy contributions to local parishes. This implies that social image concerns related to prosocial behavior (Bénabou and Tirole, 2006; Ariely *et al.*, 2009) are not pertinent in the context of the local church levy. Third, the church district uses the church levy funds to pay a fixed annual grant for each building (church, clergy house, or parish center) a parish maintains. Fourth, the local church levy is of minor importance to the Catholic Church's overall finances. As mentioned before, the main source of revenue of the Catholic Church in Germany is the general church tax, which is collected among all church members.[14] Fifth, treatment take-up in our setting is very high. In the year after our intervention, Cagala *et al.* (2014) conducted a randomized phone survey on take-up among recipients of the solicitation letter ($N = 101$). 96% of respondents acknowledged that they received the solicitation letter, and 83% stated that they opened the letter.

While the local church levy provides us with a unique opportunity to study crowding out of intrinsic motivation through taxation there is of course a potential trade-off with external validity along four dimensions. First, if some church members prefer to donate outside the context of the local church levy this leads us to underestimate their baseline intrinsic motivation, so that we potentially misclassify them as individuals with weak or no intrinsic motivation. Note, however, that this should leave our findings unaffected (and only raise standard errors) as potential

[13]The district administration does not do anything to collect church levy contributions apart from sending out solicitation letters once a year. There are no reminders for individuals who do not contribute.

[14]A system of church taxes that is similar to the one in Germany also exists in other European countries, including Austria, Denmark, Finland, Iceland and Sweden. In Germany, the general church tax amounts to 8% or 9% (depending on the state) of an individual's income tax liability. In Bavaria, the church tax is collected by a church tax administration obtaining information to levy the tax from the state tax authorities. There is no link between the church tax administration and local parishes or church districts. In particular, the church tax administration does not forward any information on individual incomes to parishes or to the church district administration.

misclassification is uncorrelated with treatment. In practice, it turns out that it is very uncommon for individuals in Germany to directly donate to the church: direct donations account for less than 3% of total church revenues, making misclassification an unlikely issue (the most important sources of funding are the general church tax and governmental transfers, which together account for about 97% of total church revenues). Survey responses (see Section 6) show that the weakly intrinsically motivated individuals, for which the crowd-out is strongest, are the least likely to attend church services. This should further alleviate concerns about individuals considering the local church levy and direct donations to be substitutes. Second, if mainly pro-social individuals become church members this would lead us to overstate baseline intrinsic motivation. Our findings should still go through as treatment is again uncorrelated with potential mismeasurement of intrinsic motivation. Notwithstanding the above, we expect our sample to rather accurately reflect the strength of intrinsic motivation in the population as a whole: i. individuals become church members by default when baptized (typically at birth) and ii. church members are very similar to the overall population in terms of giving behavior–Table A1 in the Appendix shows that if anything, donations by church members are somewhat below average. Third, if church members value church services more than the public services financed by other taxes, this could raise baseline intrinsic motivation compared to other contexts. This is very unlikely, though, as most church members make use of church services rather infrequently. For instance, in the area studied only about 11.6% of church members attend a religious service on Sundays (Catholic Church 2014). Forth, individuals could consider it particularly inappropriate to collect the local church levy as a tax. This would lead us to overestimate the crowd-out compared to other settings. However, given that churches are almost fully funded through church taxes in Germany, both nowadays and historically, this should not be a major concern.[15]

3 Theoretical Framework

This section presents a simple warm-glow model of public goods contributions (Andreoni 1989, 1990). The model highlights the role of one particular institutional feature, namely the mode by which individual contributions are collected. The collection mode varies between voluntary contributions (donations) and compulsory payments (taxes). For simplicity of exposition, we let the charity choose whether to collect the contributions as taxes or as donations. While this is exactly the choice that the Catholic Church faces in our experimental setting, most charities do not have

[15]In our post-treatment survey, we find that acceptance of the church levy is pretty high even if it is framed as a tax: in the tax treatment groups, 69% of respondents choose 'strongly agree', 'agree', or 'undecided' in response to the statement 'I consider it just that the church district collects the church levy'.

the power to raise taxes. A more general interpretation would thus take the charity's choice in our model as the institutional choice of the society to finance a given public good by taxes or by voluntary contributions. To allow for the crowding-out effects studied in the field experiment, individuals' intrinsic motivation to give to the charity may be affected by the mode of collection. We show how the profile of contributions changes when the charity switches from a donation to a taxation mode of collection. All proofs are relegated to Appendix A.

Consider an economy with a continuum of individuals of mass 1. Each individual has an initial income of I and decides to allocate this income between two goods: a private consumption good c and a contribution to a charity d. We assume that the decision to contribute is driven only by warm-glow (Andreoni, 1990). The utility function of an individual is given by

$$U = u(c) + \Theta v(d),$$

where the functions $u(.)$ and $v(.)$ are increasing and strictly concave, and $v(0) = u(0) = 0$. The type Θ denotes the intrinsic motivation of the individual for contributing to the charity. The individual's budget constraint is given by

$$c + d \leq I. \tag{BC}$$

The charity operates in an institutional environment where it can decide on the mode of collection of individual contributions: the charity can accept donations, but it is also entitled to claim a mandatory contribution from all individuals of value $\hat{d}$. As discussed in Section 2, this mirrors exactly the situation in Germany, where major religious communities can collect donations and at the same time are allowed to raise church taxes from their members. We therefore consider two modes of collection: a *donation* and a *taxation* mode. Under the donation mode, the charity does not try to recover the mandatory contribution from all individuals and let individuals freely decide whether and how much they want to donate. Under the taxation mode, the charity imposes a compliance constraint. This constraint represents the minimal level of contribution and captures the legal norm implied by a taxation mode so that

$$0 \leq \tau \hat{d} \leq d. \tag{CC}$$

The parameter $\tau \in [0, \bar{\tau}]$, with $\bar{\tau} \leq 1$, reflects the degree to which the legal norm is binding.[16] In

[16] In a standard tax compliance model, the parameter τ would correspond to the level of enforcement of the mandatory contribution $\hat{d}$.

our field experiment, we induce variation in τ by treatments communicating the existence of a tax law requiring individuals to make certain minimum payments.

We now turn to the key ingredient of our model: the idea that the warm-glow utility from giving might depend on the collection mode. For simplicity of exposition, we consider an economy with three groups of individuals. Each group is characterized by a level of intrinsic motivation that is either zero, intermediate, or high. When the charity uses a donation (resp. taxation) frame, the individual's intrinsic motivation is given by $\Theta_D \in \{0, \underline{\theta}, \overline{\theta}\}$ (resp. $\Theta_T \in \{0, \underline{\theta}', \overline{\theta}'\}$). To capture the idea that individuals' intrinsic motivation might decrease when switching from a donation to a taxation mode, we assume[17]

$$0 < \underline{\theta}' < \underline{\theta} < \overline{\theta}' \leq \overline{\theta}. \tag{1}$$

The following proposition presents the schedule of contributions when the charity uses a donation mode.

Proposition 1: Contributions under donation mode

In the donation mode, the contribution of individuals with intrinsic motivation Θ_D *equal to* 0, $\underline{\theta}$, *and* $\overline{\theta}$ *is given by* 0, $\underline{d}$, *and* $\overline{d}$, *respectively, with* $0 < \underline{d} < \overline{d}$.

Proof: See Appendix A.

Proposition 1 shows that individuals with intrinsic motivation separate themselves from the non-motivated ones and donate some positive contributions even if there is no enforcement of the contributions to the charity.

We assume that $\underline{\theta}$ is sufficiently large so that

$$\bar{\tau}\hat{d} \leq \underline{d}. \tag{A1}$$

This assumption implies that, in the donation mode, individuals with an intermediate level of intrinsic motivation would contribute more than their contribution under the most binding legal norm in the taxation mode.[18] We next turn to the schedule of contributions when the charity uses a taxation mode.

Proposition 2: Contributions under taxation mode

In the taxation mode, the contribution of individuals with intrinsic motivation Θ_T *equal to* 0, $\underline{\theta}'$, *and* $\overline{\theta}'$ *is given by* $\tau\hat{d}$, $\max\{\underline{d}', \tau\hat{d}\}$, *and* $\overline{d}'$, *respectively, with* $0 < \max\{\underline{d}', \tau\hat{d}\} < \overline{d}'$ *for any* τ.

[17] Allowing $\underline{\theta}'$ to be equal to $\underline{\theta}$ would make the presentation of the results more lengthy without changing our main messages.

[18] Since the assumption is fulfilled in the church levy context, relaxing it would increase the number of cases to cover in the theoretical study without developing interesting insights for the empirical analysis.

Proof: See Appendix A.

Proposition 2 shows that individuals with no intrinsic motivation give (weakly) positive contributions in the taxation mode. These individuals are affected by the mandatory nature of the contribution through the compliance constraint. In particular, our model predicts that individuals with no intrinsic motivation contribute the minimal possible level. Individuals with a high intrinsic motivation do not get affected by the compliance constraint since they contribute strictly more than the mandatory requirement anyway. Finally, we assume that the mandatory contribution with the most binding legal norm $\bar{\tau}\hat{d}$ is such that

$$0 < \underline{d}' \leq \bar{\tau}\hat{d}. \tag{A2}$$

Assumption $(A2)$ implies that, in the taxation mode, individuals with an intermediate level of intrinsic motivation may be affected by the compliance constraint, depending on the degree to which the legal norm imposed by the charity binds.[19]

The following corollary characterizes the configuration of parameters leading to some bunching of individuals with different intrinsic motivation at the mandatory contribution level.

Corollary 1 *When the legal norm is sufficiently binding so that* $\tau\hat{d} = \max\{\underline{d}', \tau\hat{d}\}$*, individuals with type* $\Theta_T \in \{0, \underline{\theta}'\}$ *bunch at the contribution level* $\tau\hat{d}$*.*

The following proposition establishes our results regarding the crowding out of intrinsic motivation.

Proposition 3: Crowding out of intrinsic motivation

(I.) Weak legal norm: *When the legal norm is not very binding so that* $\underline{d}' = \max\{\underline{d}', \tau\hat{d}\}$*, the crowding out of intrinsic motivation when switching from donation to taxation mode leads to the following schedule of contributions:*

$$\underline{d}' < \underline{d} < \overline{d}' \leq \overline{d}.$$

(II.) Strong legal norm: *When the legal norm is sufficiently binding so that* $\tau\hat{d} = \max\{\underline{d}', \tau\hat{d}\}$*, the crowding out of intrinsic motivation when switching from donation to taxation mode can be partially compensated by enforced compliance. The schedule of contributions then becomes:*

$$\tau\hat{d} \leq \underline{d} < \overline{d}' \leq \overline{d}.$$

[19] This assumption can be relaxed without changing the main results of our model. However, assumption $(A2)$ holds in our setting (see Section 5).

Proof: See Appendix A.

Proposition 3 establishes the crowding out of intrinsic motivation by taxes by demonstrating that a switch from the donation to the taxation mode may trigger a decrease in contributions made by individuals with types $\underline{\theta}'$ and $\overline{\theta}'$ relative to their initial donations $\underline{d}$ and $\overline{d}$. For individuals with intermediate intrinsic motivation, the crowding out of intrinsic motivation under taxation mode can be partially compensated by making the legal norm more binding, i.e. increasing τ. In contrast, contributions of highly motivated individuals do not respond to making the legal norm more binding.

4 Experimental Design and Data

4.1 Randomized Natural Field Experiment

We exploit the institutional setting described in Section 2 to design a field experiment which shifts the mode of financing of a public good from purely voluntary contributions to compulsory tax payments. In conjunction with the Catholic Church we manipulated the content of the cover letter of the mail-out in April 2013 and varied the framing for contributions to the church levy funds. Recipients were randomly assigned into a control group and three treatment groups: a donation treatment, a voluntary tax treatment, and a compulsory tax treatment.[20] In the following, we discuss each of the four letters.

Control letter. The content of the control letter corresponds to the letter which was sent out in earlier years. The exact wording and format of the control letter is shown in the Appendix. The control letter emphasizes that the church levy is considered a charitable donation. Accordingly, the letter specifies neither the amount church members might contribute nor a payment deadline. The front page of the letter highlights the good cause and explains that the church levy is necessary to provide local public goods ("the parishes need your church levy"). The second page of the letter informs recipients about institutional details of the church levy. The letter provides no external incentive for contributing to the public good.

Donation letter. Apart from shortening the first paragraph, the main difference to the control group letter is that the donation letter mentions the amount of €15 (the amount that most individuals legally owe). The first paragraph of the letter reads as follows:

"As every year, we kindly ask you herewith for your local church levy contribution (see overleaf

[20] 'Voluntary tax' is an established term in the public finance literature. Cooper (1979) and Slemrod (1998) use and discuss the term, pointing to the fact that when designing tax systems, governments strongly rely on 'voluntary compliance'.

for legal background). The church levy is staggered according to income and equal to €15 for the highest income bracket. The church district considers the church levy a contribution equivalent to a charitable donation. You decide how much you wish to contribute."

To determine the effect of providing a reference point of €15, we compare the donation letter to the control letter. All else equal, we might reasonably expect that the treatment increases the probability of contributing among baseline non-contributors: some non-contributors might be uncertain about how much to give in the baseline and potentially start contributing once they receive guidance on that issue. An increase in the probability of contributing is also anticipated if baseline non-contributors incur a mental cost for disappointing articulated expectations on which they only had vague information before treatment (Charness and Dufwenberg, 2006; Batigalli and Dufwenberg, 2007). If the reference amount mentioned in the letter serves as an anchor we should additionally see a treatment response of baseline contributors at the intensive margin: we expect individuals with baseline contributions above €15 to reduce and individuals with baseline contributions below €15 to increase their contribution.

The two subsequent treatment letters communicate the existence of a legal norm by varying the first paragraph of the donation letter.

Compulsory tax letter. The compulsory tax letter communicates a binding legal norm to contribute to the public good by informing individuals that the church levy forms part of the church tax. The first paragraph of the compulsory tax letter reads as follows:

"As every year, we kindly ask you herewith for your local church levy contribution (see overleaf for legal background). The church levy is part of the church tax and is therefore a compulsory payment. The levy is staggered according to income and equal to €15 for the highest income bracket. The church district considers any church levy payment that exceeds the compulsory amount a contribution equivalent to a charitable donation. You decide how much you wish to contribute."

Voluntary tax letter. The voluntary tax letter communicates the existence of the legal norm by informing individuals that the church levy forms part of the church tax, but frames the norm as being non-binding by letting recipients know that the church refrains from collecting the church levy as a tax. This is communicated as follows:

"As every year, we kindly ask you herewith for your local church levy contribution (see overleaf for legal background). The church levy is part of the church tax and is therefore a compulsory payment. The levy is staggered according to income and equal to €15 for the highest income bracket. We abstain, however, from collecting the church levy as a compulsory payment. Instead, the church district considers the church levy a contribution equivalent to a charitable donation. You decide how much you wish to contribute."

The natural comparison group for both tax treatments is the donation letter. All else equal, Proposition 3 suggests that baseline contributors should reduce their contributions if their intrinsic motivation is crowded out by the tax framing. We expect the drop in contributions to be most pronounced among recipients of the voluntary tax letter as the crowd-out among recipients of the compulsory tax letter might be (partly) compensated by the incentive to comply with a binding norm. For baseline non-contributors, making the legal norm binding is predicted to increase compliance at no cost. For this group, we thus expect the compulsory tax letter to dominate (in terms of contributions) any other letter.

4.2 Data and Measure of Intrinsic Motivation

Data. Our empirical analysis rests on administrative records of church levy payments made in years 2005-2013. The data is collected by the church district administration and records the amount and date of each payment together with individual characteristics such as marital status, sex, and age.

Our sample consists of 39,788 individuals.[21] In 2012, the year before the experiment, 11.5% of individuals in the sample made a strictly positive contribution to the church levy funds, compared to 88.5% who did not contribute. Figure 1 shows the distribution of strictly positive contributions in 2012. The mode of the distribution is €20 (23% of contributors). The vast majority of individuals contribute between €5 and €50. Less than 8% give strictly more than €50 (with 6% contributing €100).

Table A1 in the Appendix provides evidence on the representativeness of our sample by comparing average characteristics of individuals in our sample to those of the full population living in the urban area we study.[22] The table shows that Catholic Church members are very similar to the overall population in terms of age, distribution of income, and charitable giving behavior.

Our sampling procedure uses strata defined by past contribution behavior, household type and age.[23] Table A2a in the Appendix reports randomization checks. On average, individuals are 52

[21] In total, the church district mailed the solicitation letter to 63,177 individuals. To prevent spillovers (Rincke and Traxler, 2011), we excluded church employees interacting with members and individuals who share the same name and address with at least two other individuals. We also excluded individuals who appeared on the mailing list for the very first time (as we cannot measure their intrinsic motivation by past contribution behavior) and a few individuals who donated more than 300 Euro in previous years. This left us with 56,750 individuals for treatment, of which 16,962 individuals were assigned to treatments belonging to a separate research agenda on gift exchange (Cagala *et al.*, 2014). To preserve power for the identification of crowding out, we assigned to the gift exchange treatments only individuals who did not make any positive contribution in the two years prior to treatment. We correct for the resulting differences in sampling ratios for baseline contributors and baseline non-contributors by using weighted regressions for population effects (only applies to results on donation letter in the Appendix).

[22] The data on the full population stems from personal income tax records 2007, the most recent year of available data for filers and non-filers.

[23] The stratification frame uses three bins for baseline contribution behavior in 2011 and 2012 as well as bins for

years old and 49% of them are men (single or married). The average probability of contributing in baseline year 2012 is 11%, with an (unconditional) average contribution of €3.9. An F-test of joint significance of the covariates reveals no difference in observable characteristics across treatment groups (p-values displayed in column (10)). Hence, differences in behavior across treatment groups reflect the causal effect of treatment.[24] Table A2b repeats the randomization checks for the subsample of baseline contributors (at least one positive contribution in years 2005 to 2012), who comprise 17.5% of the sample.[25] The table demonstrates that baseline characteristics are again balanced. Compared to the full sample, baseline contributors are older (69 years), less likely to be male (29%), and more likely to contribute in baseline year 2012 (66%), with a much larger (unconditional) average contribution in 2012 (€22.2).

Figure A1 in the Appendix shows the time pattern of payments made. The figure reveals that more than 80% of all contributions are made within the first five weeks after the mail-out of the letter. In our empirical analysis we include all payments received within the first 20 weeks of the experiment (corresponding to 94% of all payments effected until December 31, 2013).

Measure of intrinsic motivation. It is essential for our design to accurately distinguish between different types regarding the strength of intrinsic motivation. To do so, we exploit the fact that we have access to individual-level panel data on contribution behavior in up to 8 years prior to treatment. Given the absence of external incentives in the baseline, we use this data to derive a continuous measure of intrinsic motivation. It serves two different purposes. First, we employ a basic distinction between baseline non-contributors and baseline contributors to perform sample splits corresponding to the distinction between individuals with zero intrinsic motivation and intermediate or high intrinsic motivation from the theory model. Second, we use the *relative frequency of contributing* in the baseline as a continuous measure of intrinsic motivation within the group of baseline contributors to flexibly estimate the interaction between the tax treatments and motivation.[26]

age quartiles and four household types (single Catholic female, single Catholic male, Catholic female married to non-Catholic spouse, married Catholic male). This gives a total of $3 \times 4 \times 4 = 48$ stratification bins.

[24]The other key identifying assumption is that there are no spillovers across treatments. For several reasons, spillovers are not likely to be a central issue here. First, each household only received one letter. Second, there was no media coverage of the experiment and, therefore, no general public awareness about it. Third, tax information is strictly private in Germany and it is very uncommon to talk about tax issues outside the family. Fourth, only a small share of about 11.6% of church members actively participate in church activities, by attending church services (Catholic Church 2014).

[25]We note that the level of intrinsic motivation revealed by baseline contribution behavior in our setting is in the same range as in DellaVigna *et al.* (2012), who study a charitable giving context and find that 25% of individuals are intrinsically motivated to give.

[26]Across eight pre-treatment years used to measure baseline motivation, the average individual is observed in 6.7 years. In our estimations, we restrict attention to individuals observed in at least three pre-treatment years. Our measure of intrinsic motivation is robust towards excluding individuals who are observed in less than eight pre-treatment years (17% of the population): dropping them leaves the distribution of motivational types unchanged.

An alternative measure of intrinsic motivation could be based on *amounts contributed* in the baseline. However, as the relative frequency of contributions is much less likely to pick up income effects, we prefer a frequency-based measure of intrinsic motivation over amount-based measures. In Section 6, we check the validity of our approach by relating the frequency of past contributions to survey-based measures of intrinsic motivation.

5 Results From the Field Experiment

This section presents the evidence from the field experiment. We focus on the effects of the tax treatments relative to the donation treatment. To study the crowding out of intrinsic motivation, we consider the sample of baseline contributors. The compliance effects are identified from the sample of baseline non-contributors. In the Appendix, we compare the donation treatment to the control group and demonstrate that providing a reference point alone has the predicted effects: it increases the probability of contributing among baseline non-contributors and serves as an anchor when individuals decide about how much to give (for details, see Appendix B and Tables A3a and A3b).

Throughout the paper, we report relative treatment effects from linear regressions (estimated treatment effects in levels divided by mean outcome in the omitted reference group) while controlling for strata variables and parish fixed effects[27] and calculate robust standard errors.

5.1 Crowding Out of Intrinsic Motivation

Table 1 documents responses of baseline contributors to the tax treatments relative to the donation treatment. Following Proposition 3, we expect any crowding out of intrinsic motivation to be most pronounced among the weakly intrinsically motivated, and to be smaller for the strongly intrinsically motivated. To account for this heterogeneity, we estimate the treatment effects including interaction terms between treatment indicators and baseline motivation.

Note that we tend to underestimate the size of the crowd out when looking at average effects evaluated at the sample mean: baseline probabilities of contributing for the weakly intrinsically motivated are well below the overall average in the sample, so that the relative drop in intrinsic motivation for the weakly intrinsically motivated is in fact much more pronounced than what is suggested by average effects evaluated at the sample mean. To dig deeper, Figure 2 presents evidence on the heterogeneity in treatment effects across motivational types.

[27] Point estimates with and without controls are very similar (with slightly smaller standard errors when using controls). See Tables A3a and A3b in the Appendix for a comparison of estimates with and without controls.

Columns (1) to (4) refer to treatment effects in the treatment year. Panel A displays our findings if we pool the voluntary and the compulsory tax treatments. On the extensive margin, column (1) shows that communicating the existence of a legal norm significantly reduces the probability of contributing: if intrinsic motivation is very weak, the tax framing reduces the likelihood of contributing significantly. Evaluating the relative treatment effect for the pooled sample at the minimum of our measure for motivation of 0.125 (one year with strictly positive contribution out of eight baseline years), we estimate the tax letters to reduce the probability of contributing by 7.3% (p-value: 0.049).[28] The crowding out effect disappears, however, if baseline intrinsic motivation is strong: then, the negative baseline effect is fully offset by the positive interaction term. Columns (2) and (3) of Panel A display the pooled treatment effect on the probability of contributing weakly less and strictly more than the reference point of €15. The estimates show that crowding out of intrinsic motivation significantly reduces the probability of contributing larger amounts (again among the weakly intrinsically motivated), while we do not see any effect on the probability of contributing small amounts. The latter observation is particularly interesting. It suggests that individuals with large baseline contributions mainly respond at the *extensive* margin by ceasing to contribute, instead of reducing their contribution at the intensive margin. Column (4) presents results on the total effect, summarizing extensive and intensive margin responses. The estimates confirm that informing individuals about the legal norm significantly reduces contributions by the weakly intrinsically motivated, but not by the strongly intrinsically motivated.

Having shown pooled treatment effects, we now turn to the individual effects of the two tax letters (Table 1, Panel B). The estimates reveal that the pooled crowd-out effect is driven by the voluntary tax letter: the least intrinsically motivated among baseline contributors are significantly less likely to contribute (-8.9% (p-value: 0.035), column (1)) under the voluntary tax treatment. Again, this effect is mostly driven by a decline in the probability of contributing more than the reference point of €15 (-14.7% (p-value: 0.013), column (3)). In total, the least intrinsically motivated individuals reduce their contribution by 43.6% (column (4), p-value: 0.030) on average.

The findings for the voluntary tax letter stand in sharp contrast to the results for the compulsory tax letter (see lower part of Panel B). Across all margins, we find that the effect of the compulsory tax letter is insignificant. This is in line with the theoretical prediction of a compliance effect counteracting the crowding-out of intrinsic motivation among the weakly intrinsically

[28]Note that the relative effects in Table 1 are based on the average probability of contributing in the sample as a whole. This necessarily leads us to underestimate the size of the crowd-out as the probability of contributing among the weakly intrinsically motivated is much lower compared to the overall population. Our discussion of Figure 2 below will take up this issue.

motivated.

Note that we tend to underestimate the size of the crowd out when looking at average effects evaluated at the sample mean: baseline probabilities of contributing for the weakly intrinsically motivated are well below the overall average in the sample, so that the relative drop in intrinsic motivation for the weakly intrinsically motivated is in fact much more pronounced than what is suggested by average effects evaluated at the sample mean. To dig deeper, Figure 2 presents evidence on the heterogeneity in treatment effects across motivational types. The figure is based on the sample of all individuals who have received the solicitation letter in at least three years prior to treatment and who have contributed between 20 and 100 Euro at least once ($N = 2283$). Each panel plots the relative difference in the probability of contributing between the tax letters and the donation letter for subsamples based on the strength of baseline motivation. In each panel, the left-most bar depicts the difference in the probability of contributing for those contributing in up to 25 percent of baseline years (weak intrinsic motivation).[29] The second (third) bar shows the difference in probability for those with a frequency of contributing larger than 25 and weakly less than 50 percent (larger than 50 and weakly less than 75 percent), respectively, while the right-most bar is for those contributing in more than 75 percent of baseline years (strong intrinsic motivation). The figure thus flexibly accounts for heterogenous treatment responses without imposing the restriction of a *linear* interaction (as in Table 1).

Panel A of Figure 2 considers the pooled effects of the tax letters relative to the donation letter. It demonstrates that the likelihood of contributing is reduced by about 14% among the weakly intrinsically motivated. In contrast, the strongly intrinsically motivated do not seem to respond to the tax treatments.

Panels B and C of Figure 2 show the relative differences in the probability of contributing for the two tax treatments separately. For the voluntary tax treatment, we note a distinct reduction in the probability of contributing by almost 20% among the weakly intrinsically motivated. Moreover, Panel B establishes a monotonic relation between baseline motivation and the strength of the treatment effect. As regards the compulsory tax letter, Panel C reveals much smaller treatment effects, consistent with the notion of compliance effects offsetting the motivational crowd-out. It is only for the weakly intrinsically motivated that we find a negative treatment effect of the compulsory tax letter (-8%).

[29] Recall that we focus on baseline contributors here. Individuals who have never contributed in the baseline are not considered.

5.2 Persistency of Crowding Out

Having established the presence of a short-term crowding-out effect, we now turn to its persistency. In the year after treatment, we sent out the donation letter (exact same layout and wording as in the treatment year) to all individuals in the donation and tax treatment groups. That is, we fully removed any reference to the legal norm from the letters, which may restore the initial level of intrinsic motivation. It is possible, though, that even a one-time intervention which points individuals to the fact that the church levy is a tax has a persistent negative impact on contribution behavior (if individuals remember the legal norm from the previous letter).

Studying the payment data from the year 2014, we indeed find evidence suggesting a persistent, but attenuated, crowd-out. When we repeat the estimations from Table 1, columns (1) to (4), for the year after treatment, we obtain the same pattern of coefficients as before, but with much smaller point estimates. With similar standard errors as before, the point estimates become insignificant where they were significant for the treatment year (results not reported). Columns (5) and (6) in Table 1 display the results if we consider the probability of a reduced contribution *relative to the baseline year 2012* for the year of treatment and the year after treatment, respectively. Column (6) demonstrates that even after removing the information on the legal norm, the weakly intrinsically motivated in the voluntary tax group are significantly more likely to pay less compared to the baseline year 2012.

To summarize, we find strong evidence of a crowding out of intrinsic motivation if voluntary contributions are turned into compulsory tax payments. However, we identify a crowd-out effect only among the weakly intrinsically motivated. The crowd-out of intrinsic motivation is (partly) compensated by a more binding legal norm, which is in line with the predictions of our theoretical model.

5.3 Compliance Responses

We next consider baseline non-contributors and study compliance responses to the tax treatments. As revealed by baseline contribution behavior, there is no potential for a crowd-out effect on contributions in this group. Proposition 2, however, predicts that imposing a legal norm *increases* contributions among baseline non-contributors if the norm is sufficiently binding.

We first consider the effect of the voluntary tax letter. As it communicates a legal norm that is not made effective, we expect a small compliance effect (if any) among baseline non-contributors. This is confirmed in Table 2, which reports the treatment effects of the voluntary tax and the compulsory tax letters, again evaluated relative to the donation letter. The table shows that the voluntary tax letter has no impact on behavior among baseline non-contributors: individuals are

no more likely to contribute in the presence of a non-binding norm (column (1)) nor more likely to increase their average contribution (column (4)). Given the sample size, the insignificance of these effects is unlikely to be driven by lack of power.

Second, we take a closer look at the treatment effects of the compulsory tax letter. Our model predicts that, if perceived as legally binding, the compulsory tax letter should increase the contribution among baseline non-contributors. More specifically, we expect this effect to be confined to the probability of contributing weakly less than the amount owed (€15). This is exactly what we find. As Table 2 demonstrates, the compulsory tax letter increases the probability of contributing among baseline non-contributors by 26% (column (1)). This effect is entirely driven by an increase in the probability of contributing weakly below the requested amount (+38%, column (2)). As regards the total response, the compulsory tax letter increases the average contribution of baseline non-contributors by about 3.6% (column (4)).

6 The Post-Treatment Survey

6.1 Survey Design

This section describes cross-validation tests regarding the crowd-out effect.[30] They are based on a post-treatment survey that elicits alternative measures of intrinsic motivation. We stopped the collection of payment data at the end of week 20 after the mail-out of the church levy notice. Shortly thereafter survey questionnaires were sent out to all individuals who received either the donation letter or one of the tax letters ($N = 29,841$). The mailings comprised a short cover letter, a one-page questionnaire (see the Appendix for both documents), and a return envelope pre-filled with the church district's postal address for postage-free return of the questionnaire. The cover letter explained that the church district seeks advice on how to frame the church levy notice in future mail-outs and explicitly mentions that participation was voluntary, anonymous, and costless.

The questionnaire covers a total of 11 items. In each item, respondents could choose between five ordered response options (Likert scale). The items refer to attitudes towards the church levy, willingness to contribute, relation to the Catholic Church, relation to the local parish, church attendance habits, and the willingness to donate and volunteer in other than church contexts.

To facilitate the cross-validation of the field-experiment, the framing of the survey questionnaire is treatment-specific. First, a short header repeats the treatment from the church levy notice by

[30] We do not consider compliance effects here as it seems natural to interpret the decision to respond to the survey itself as evidence of *some* intrinsic motivation. Hence, based on the survey data, it seems inappropriate to define a group of individuals representing agents with *zero intrinsic motivation* in the theoretical model.

reiterating the information regarding the church levy being a voluntary contribution, a voluntary tax, or a compulsory tax. Second, the questionnaire asks respondents about the change in their willingness to contribute if the collection mode changed *relative to what was communicated in the treatment letters*. The questionnaire going to individuals in the donation treatment group can be found in the Appendix. For the tax treatment groups, the wording of the treatment-specific parts is as follows:

Compulsory tax treatment group. The questionnaire header states that *"In mid-April, you received the church levy notice. The notice has informed you that the church levy forms part of the church tax and is therefore a compulsory payment".* The willingness-to-contribute question is formulated accordingly: *"The church levy is a compulsory payment. If the church levy was instead a completely voluntary contribution, I would pay..."*, with response options ranging from *much less* to *much more*.

Voluntary tax treatment group. The questionnaire header reads as follows: *"[...] The notice has informed you that the church levy forms part of the church tax and is therefore a compulsory payment. As stated in the notice, however, we abstain from collecting the church levy as a compulsory payment. Instead, the church district of [location] considers the church levy a contribution equivalent to a charitable donation".* The willingness-to-contribute question reads: *"The Catholic Church treats the church levy as a voluntary contribution, despite the fact that it is legally a compulsory payment. If the church levy was instead a completely voluntary contribution, I would pay...".*

In order to be able to relate individual survey responses to a set of key individual characteristics, including church levy contributions, we pre-coded the questionnaires prior to the mail-out.[31] The following information is captured by the code: household type, age, the respondent's local parish, church levy contribution in baseline year 2012, and church levy payment in 2013.

We exploit the pre-coded information in three distinctive ways. First, we test whether respondents' observable characteristics are balanced across treatments. Second, the pre-coded information allows us to link baseline contribution behavior to several survey-based proxies for intrinsic motivation. This allows us to check how well baseline contribution behavior captures individuals' intrinsic motivation and to replicate estimations of the crowd-out effect from the field experiment using survey-based motivational measures (instead of the frequency of baseline contributions). Third, the pre-coded information allows us to estimate the crowd-out effect while conditioning on the exact same set of control variables as in the field experiment.

[31] The code allows us to recover individual characteristics from the incoming survey questionnaires while protecting the privacy of respondents.

6.2 Characteristics of Survey Responders

The mailing lists for the questionnaires were identical to the corresponding mailing lists of the church levy notice. The randomization of treatment assignment in the field experiment thus ensures that observable characteristics of survey *recipients* are balanced across treatments. However, selective response behavior might lead to different average characteristics of survey *respondents* across treatment groups. Table A4 in the Appendix demonstrates that all observable individual characteristics of survey respondents are balanced across treatments, but that survey respondents differ in observable characteristics from the average survey recipient. Column (2) reports the survey response rates (8.3% to 9.3%). Columns (3) to (8) report means and 95% confidence intervals for respondents' age, three out of four household type dummies (single female being the omitted reference category), and contribution behavior in baseline year 2012. Compared to the full sample of survey recipients, survey respondents are, on average, about 10 years older and more likely to be married. The probability of contributing in the baseline year 2012 is more than four times larger among respondents than in the population covered by the survey, translating into higher unconditional average contributions.[32] Column (9) reports p-values for F-tests suggesting that characteristics are jointly insignificant in predicting assignment to treatments. We conclude that, although survey respondents differ in observables from the overall population of survey recipients (i.e., all treated individuals in the field experiment), there is no evidence of differences in the selectivity of survey respondents between treatment groups.

6.3 Validity of Field-Experimental Measure of Intrinsic Motivation

We next check the validity of our field-experimental measure of intrinsic motivation. For this purpose, we consider the correlation between survey-based measures of intrinsic motivation and the willingness to contribute revealed by baseline contribution behavior. The survey includes three questions aiming at different indicators of intrinsic motivation: relation to local parish, church attendance, and charitable giving and volunteering in other contexts. The wording of the questions is as follows:

Relation to local parish. *"My relation to my local parish is best described as..."*, with response options ranging from "very weak" to "very close". As revenues of the church levy remain at the local level, we consider individuals who care more about their local parish to be the more intrinsically motivated to pay the church levy.

Church attendance. *"I attend church services or other religious events..."*, with response

[32] Conditional on contributing, average contributions in both samples are almost identical (results not shown).

options ranging from "never" to "daily". The church levy funds are used to provide public goods within parishes. We thus consider individuals who attend church more regularly (and thus use the public goods provided more intensely) to be more strongly intrinsically motivated to contribute to the church levy funds.

Charitable giving and volunteering in other contexts. *"I engage as a volunteer or a donor..."*, with response options varying between "very rarely" and "very frequently". This question provides us with a measure of intrinsic motivation which goes beyond the church context and captures an individual's general propensity to engage as a donor or volunteer.

Figure 3 depicts how responses to these three survey questions relate to baseline contribution behavior and reports (polychoric) correlations (ρ, with standard errors in parentheses). The panels on the left-hand side (Panels A, C, and E) report the average probability of contributing for each response category, while the panels on the right-hand side (Panels B, D, and F) report the average amount contributed for each response category. Except for Panel E, we observe a strong positive correlation between the survey-based measures of intrinsic motivation and baseline contribution behavior. For instance, the probability of contributing in 2012 among survey respondents who consider their relation to their local parish as very weak was 26.4%, while the corresponding figure for those who consider their relation to be very strong is 67.7%. 39.5% of individuals who report never to go to church made a strictly positive contribution in 2012, whereas 76.5% of daily church attendees contributed. A very similar pattern is observed when we look at charitable giving and volunteering in other contexts (44.2% vs. 60.2%). For all three survey questions the probability of contributing is monotonically increasing in intrinsic motivation. Figure 3 thus establishes a strong correlation between baseline contribution behavior and survey-based behavioral measures of intrinsic motivation.

6.4 Survey-Based Cross-Validation of Crowding Out

We now turn to the cross-validation of the crowd-out effect identified in the field experiment. Our first exercise is straightforward and consists of replicating the estimations of the crowd-out effect from Table 1 in the sample of survey respondents. We recover post-treatment contribution behavior from the information pre-coded on questionnaires and derive the same dependent variables as in the section on our randomized field experiment. Unlike before, we do not rely on baseline contribution behavior as a proxy for intrinsic motivation but use survey responses instead. Table 3a reports effects on the probability of contributing (columns (1) to (3)) and on contributions (column (4) to (6)). As in Table 1, the explanatory variables of interest are the tax treatment indicators and

the interactions between treatment indicators and measures of motivation.[33] The measures of motivation take values from 1 to 5 (higher values indicating stronger motivation), corresponding to the five ordered response categories for each of the motivational survey questions.

Table 3a confirms our earlier finding of a significant crowding out of intrinsic motivation: among weakly intrinsically motivated individuals, the voluntary tax treatment has a negative effect both at the extensive and the intensive margin.[34] To give an example, survey respondents who never attend church (motivational measure 'Church Attendance' takes value one) are 14.0 percent less likely to contribute in the voluntary tax group, relative to the donation letter group. At the intensive margin, the effect is even more pronounced (minus 30.9 percent). Again we find the effect of the compulsory tax treatment to be insignificant, which confirms the finding that compliance compensates the revenue loss caused by crowding out of intrinsic motivation.[35]

Our second cross-validation test focuses only on the voluntary and the compulsory tax letter groups. It exploits between-treatment differences in responses to the survey question on changes in the willingness to pay in case of an institutional switch from tax to donation mode. Our test takes the form of a simple linear probability model, with the voluntary tax treatment indicator as the explanatory variable of interest (i.e., the compulsory tax treatment group serves as reference category). The dependent variable is a dummy variable taking value one if the respondent states that she *would pay more* if the church levy, instead of being a legally binding tax, was a completely voluntary contribution.[36] The model thus tests if respondents in the voluntary tax group differ from their counterparts in the compulsory tax group regarding their willingness to *increase* their contribution in case of an institutional reform making the church levy a pure donation.[37]

Table 3b displays the results for our second cross-validation test. Column (1) shows that for

[33]Just like in the regressions using the field data, all regressions reported in this section include strata variables and parish fixed effects as controls.

[34]The fact that we do not find significant effects in column (3) is in line with Figure 3, Panel E, showing that the correlation between the baseline probability of contributing in 2012 and motivation measured by the frequency of charitable giving and volunteering in other contexts is rather weak.

[35]Table 3a also suggests that strongly intrinsically motivated survey respondents tend to respond *positively* to the voluntary tax treatment. One possible interpretation is that in this specific group of church members, the voluntary tax treatment is interpreted as a signal of trust.

[36]The dummy combines two response categories, "would pay much more" and "would pay more". We do not consider individuals in the donation group here as the questionnaire for this group asks individuals about the change in their willingness to pay in case of the *reverse* institutional change, i.e. from donation to tax mode. Hence, for this group, the dependent variable in the linear probability model is not defined.

[37]This test of the crowd-out hypothesis is related to, but conceptually different from studies testing for the crowd-out effect by exposing subjects to an external incentive and then removing it (Deci, 1971). While external incentives like piece rates entail a signal that might crowd out intrinsic motivation even when the incentive has been removed, the wording of our survey question regarding the change from taxation to donation mode aims at individuals' willingness to contribute *in a different institutional setting*. We thus interpret a respondent's statement of a higher willingness to pay if the setting changed from taxation to donation as evidence of crowing out of intrinsic motivation under taxation mode.

the full sample, the coefficient of the voluntary tax indicator is positive and weakly significant, implying that on average, respondents in the voluntary tax treatment would be more likely to increase their payment if the church levy was collected as a pure donation. Thus, the potential gain in revenues if the legal norm is removed is larger for the voluntary than for the compulsory tax letter group. This is consistent with the theoretical prediction that choosing the taxation mode in the voluntary and compulsory tax letters crowds out intrinsic motivation, but that stronger enforcement in the compulsory tax letter has a (partially) compensating effect.

Our next step is to check whether the survey data display the predicted heterogeneity in the crowd-out for various measures of intrinsic motivation. Columns (2) to (7) demonstrate that this is indeed the case. Across all three motivational measures, we find strong evidence for crowding-out effects among weakly intrinsically motivated respondents, while we do not find any significant effects among individuals with strong intrinsic motivation.[38] For instance, we split the sample between regular church goers (respondents saying they attend church at least once a month, strongly intrinsically motivated) and individuals less inclined to attend church (weakly intrinsically motivated). The weakly intrinsically motivated are 86.7% more likely in the voluntary tax group (relative to the compulsory tax group) to indicate that they would pay more if the church levy was collected as a pure donation. As in the field experiment, we find very small and insignificant effects for the strongly intrinsically motivated.

7 Conclusion

This paper studies how taxes as externally defined legal norms on contribution behavior affect the willingness to contribute to public goods provision. We implement our field experiment in an urban area in Germany where the Catholic Church collects the local church levy. The setting is ideally suited to shed light on how taxes affect the willingness to contribute because the levy is collected as a voluntary contribution, despite the fact that it is legally a tax. Starting from this baseline, we implement treatments that aim at two distinct effects: crowding out of intrinsic motivation among those who previously contributed, and compliance responses among those who did not contribute in the first place.

Building on a simple theoretical model, we compare the contribution behavior of different motivational types between treatments that frame the church levy as a tax and a control letter asking for a voluntary contribution. Several novel empirical findings emerge. First, individuals

[38] Across all three motivational measures, we split the sample into weakly and strongly intrinsically motivated respondents according to the five ordered response categories such that the resulting subsamples are as similar as possible to each other in terms of sample sizes. This ensures that differences in treatment effects between subsamples are not driven by differences in statistical power.

with regular baseline contributions (the strongly intrinsically motivated) do not show any response to the information that the church levy is a tax. Second, individuals contributing only occasionally in the baseline (the weakly intrinsically motivated) reduce their payments significantly in response to a treatment framing the church levy as a voluntary tax, but do not respond to a treatment saying that the tax is compulsory. This is consistent with the notion that imposing externally defined norms on contribution behavior crowds out intrinsic motivation, but that a sufficiently strong compliance incentive can compensate the revenue effects of the crowd-out. Third, baseline non-contributors are more compliant if the communicated norm is binding (compulsory tax) but do not respond if it is non-binding (voluntary tax). Our findings on crowding out are cross-validated by an extensive post-treatment survey.

Two main conclusions can be drawn from our findings. First, imposing external rules on contribution behavior crowds out individuals' intrinsic motivation to voluntarily contribute to public goods provision. Raising taxes thus entails a hidden cost. The finding of a significant crowding-out of intrinsic motivation complements recent evidence on tax compliance suggesting that the positive effects of a better enforcement of taxes overcompensate any associated loss in intrinsic motivation (Kleven *et al.*, 2011; Pomeranz, 2013; Dwenger *et al.*, 2014). We conclude that imposing a tax norm as such crowds out intrinsic motivation, but once a tax frame is in place, the adverse effects of increasing the level of enforcement on individuals' intrinsic motivation seem to be modest. The finding of detrimental effects of imposing norms on contribution behavior also relates our study to findings of a hidden costs of control in the context of principal-agent relations (Falk and Kosfeld, 2006).

Second, the distinct heterogeneity in treatment responses uncovered by our analysis suggests that baseline motivation is an important factor that determines how subjects respond to external incentives. Depending on baseline motivation, incentives might thus lead to higher or lower individual effort. This is consistent with the findings in Huffman and Bognanno (2014), who show that workers respond very heterogeneously to incentives and conclude that the distribution of individual characteristics like worker personalities and preferences determines the overall effect of external incentives. In fact, our finding of a strongly heterogeneous treatment response could help to explain why results from previous studies on the net impact of external incentives on prosocial activities were mixed (Gneezy and Rustichini, 2000a; Ashraf *et al.*, 2012; Chetty *et al.*, 2014). One lesson to be drawn for future research on the extrinsic-intrinsic crowd-out would thus be to include the measurement of individuals' baseline motivation in the design of experimental work whenever possible. The insight that baseline motivation shapes individuals' responses to incentives and norms might also lead to fruitful extensions in the literature discussing tax-driven distortions

more generally. For instance, it would be interesting to know how individuals' baseline motivation interacts with the labor supply response identified in the literature on income taxation (Blundell and MaCurdy, 1999; Mirrlees *et al.*, 2010).

References

Andreoni, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy*, **97**, 1447–1458.

— (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *Economic Journal*, **100**, 464–477.

— (1993). An Experimental Test of the Public-Goods Crowding-Out Hypothesis. *American Economic Review*, **83**, 1317–1327.

—, Erard, B. and Feinstein, J. (1998). Tax Compliance. *Journal of Economic Literature*, **36** (2), 818–860.

Ariely, D., Bracha, A. and Meier, S. (2009). Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, **99**, 544–555.

Ashraf, N., Bandiera, O. and Jack, K. (2012). No Margin, No Mission? A Field Experiment on Incentives for Pro-Social Tasks, CEPR Discussion Paper No. 8834.

Batigalli, P. and Dufwenberg, M. (2007). Guilt in Games. *American Economic Review*, **97**, 170–176.

Bénabou, R. and Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies*, **70**, 489–520.

— and — (2006). Incentives and Prosocial Behavior. *American Economic Review*, **96**, 1652–1678.

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.

Blundell, R. and MaCurdy, T. (1999). Labor Supply: A Review of Alternative Approaches. In O. Ashenfelter and D. E. Card (eds.), *Handbook of Labor Economics*, vol. 3, Amsterdam: Elsevier, pp. 1559–1695.

Cagala, T., Glogowsky, U. and Rincke, J. (2014). The Role of Intrinsic Motivation for Reciprocal Behavior: Evidence from a Field Experiment on Gift Exchange, mimeo.

Catholic Church (2014). The Catholic Church in Germany. Facts & Figures 2013/14.

Charness, G. and Dufwenberg, M. (2006). Promises and Partnership. *Econometrica*, **74**, 1579–1601.

Chetty, R., Saez, E. and Sandor, L. (2014). What Policies Increase Prosocial Behavior? An Experiment with Referees at the Journal of Public Economics. *Journal of Economic Perspectives*, **28**, 169–188.

Cooper, G. (1979). *A Voluntary Tax? New Perspectives on Sophisticated Estate Tax Avoidance*. Washington, D.C.: The Brookings Institution.

Deci, E. L. (1971). Effects of Externally Mediated Rewards on Intrinsic Motivation. *Journal of Personality and Social Psychology*, **18**, 105–115.

— (1975). *Intrinsic Motivation*. New York: Plenum Press.

Del Carpio, L. (2013). Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru, mimeo.

DellaVigna, S., List, J. A. and Malmendier, U. (2012). Testing for Altruism and Social Pressure in Charitable Giving. *Quarterly Journal of Economics*, **127**, 1–56.

Dwenger, N., Kleven, H. J., Rasul, I. and Rincke, J. (2014). Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany, mimeo.

Esteves-Sorensen, C., Macera, R. and Broce, R. (2013). Do Monetary Incentives Crowd Out Intrinsic Motivation? A Field Test in the Workplace, mimeo.

Falk, A. (2007). Gift Exchange in the Field. *Econometrica*, **75**, 1501–1511.

— and Kosfeld, M. (2006). The Hidden Costs of Control. *American Economic Review*, **96**, 1611–1630.

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** (3), 634–660.

Frey, B. S. and Oberholzer-Gee, F. (1997). The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out. *American Economic Review*, **87**, 746–755.

Friedrichsen, J., Engelmann, D. (2014). Who Cares for Social Image?, mimeo.

Gneezy, U., Meier, S. and Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, **25**, 191–209.

— and — (2000a). Pay Enough or Don't Pay at All. *Quarterly Journal of Economics*, **115**, 791–810.

— and Rustichini, A. (2000b). A Fine Is a Price. *Journal of Legal Studies*, **29**, 1–17.

Hallsworth, M., List, J., Metcalfe, R. and Vlaev, I. (2014). The Behavioralist As Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance, NBER Working Paper No. 20007.

Huck, S. and Rasul, I. (2011). Matched Fundraising: Evidence from a Natural Field Experiment. *Journal of Public Economics*, **95**, 351–362.

Huffman, D. and Bognanno, M. (2014). Does Performance Pay Crowd Out Worker Non-Monetary Motivations? Evidence from a Real Work Setting, mimeo.

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.

Lacetera, N., Macis, M. and Slonim, R. (2012). Will There Be Blood? Incentives and Displacement Effects in Pro-Social Behavior. *American Economic Journal: Economic Policy*, **4**, 186–223.

Landry, C. E., Lange, A., List, J. A., Price, M. K. and Rupp, N. G. (2010). Is a Donor in Hand Better than Two in the Bush? Evidence from a Natural Field Experiment. *American Economic Review*, **100**, 958–983.

Luttmer, E. F. P. and Singhal, M. (2014). Tax Morale. *Journal of Economic Perspectives*, **28**, 149–168.

Mellström, C. and Johannesson, M. (2008). Crowding out in Blood Donation: Was Titmuss Right? *Journal of the European Economic Association*, **4**, 845–863.

Mirrlees, J., Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., and Poterba, J. (2010). *Dimensions of Tax Design. The Mirrlees Review*, Oxford University Press.

Pomeranz, D. (2013). No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax, NBER Working Paper No. 19199.

Pruckner, G. J. and Sausgruber, R. (2013). Honesty on the Streets: A Field Study On Newspaper Purchasing. *Journal of the European Economic Association*, **11**, 661–679.

Rebitzer, J. B. and Taylor, L. J. (2011). Extrinsic Rewards and Intrinsic Motives: Standard and Behavioral Approaches to Agency and Labor Markets. In O. Ashenfelter and D. E. Card (eds.), *Handbook of Labor Economics*, vol. 4A, Amsterdam: Elsevier, pp. 701–772.

Rincke, J. and Traxler, C. (2011). Enforcement Spillovers. *Review of Economics and Statistics*, **93**, 1224–1234.

Slemrod, J. (1998). On Voluntary Compliance, Voluntary Taxes, and Social Capital. *National Tax Journal*, **51**, 485–491.

—, 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.

Sliwka, D. (2007). Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes. *American Economic Review*, **97**, 999–1012.

Table 1: Crowding Out of Intrinsic Motivation
Sample: Baseline Contributors (Intrinsically Motivated)

	Treatment Responses in …					
	Year of Treatment					Year After Treatment
	Effect on Probability of Contributing (%)	Effect on Probability of Contribution … ≤ Ref. Point (%)	Effect on Probability of Contribution … > Ref. Point (%)	Effect on Contribution (%)	Effect on Prob. of Reduced Contribution (%)	Effect on Prob. of Reduced Contribution (%)
	(1)	(2)	(3)	(4)	(5)	(6)
Crowding Out Effects (Tax Letters vs. Donation Letter)						
A. Tax Letters, Pooled Effect						
Tax Treatments	-8.66**	-4.51	-11.48*	-40.36**	11.24	19.51**
	(4.29)	(9.17)	(6.28)	(20.45)	(8.95)	(8.30)
Tax Treatments x Baseline Probability of Contributing	11.24**	1.12	18.08*	49.32*	-21.61	-28.39**
	(5.59)	(14.46)	(10.17)	(27.07)	(13.91)	(13.08)
B. Tax Letters, Individual Effects						
Voluntary Tax	-10.92**	-.40	-18.04**	-52.33**	19.94*	21.69**
	(4.93)	(10.68)	(7.08)	(23.42)	(10.62)	(9.81)
Voluntary Tax x Baseline Probability of Contributing	15.88**	.04	26.60**	70.11**	-35.73**	-33.06**
	(6.40)	(16.81)	(11.61)	(30.89)	(16.27)	(15.26)
Compulsory Tax	-6.50	-8.45	-5.19	-28.90	2.91	17.45*
	(4.99)	(10.49)	(7.34)	(23.84)	(10.22)	(9.61)
Compulsory Tax x Baseline Probability of Contributing	6.73	1.96	9.96	29.25	-8.01	-23.84
	(6.51)	(16.55)	(11.81)	(31.59)	(15.93)	(15.10)
Outcome in Omitted Reference Group	62.53%	25.24%	37.29%	20.55€	27.60%	31.03%
Baseline Probability of Contributing	59.54%	59.54%	59.54%	59.54%	59.54%	59.65%
Number of Observations	5096	5096	5096	5096	5096	4777

Notes: OLS estimations at the individual level. *** denotes significance at 1%, ** at 5%, and * at 10% level. Robust standard errors in parentheses. All estimations account for the strata variables (age and household type) and include parish fixed effects. Baseline contributors made a strictly positive contribution at least once in pre-treatment years 2005-2012. The sample is restricted to all individuals who have received a solicitation letter in at least three years prior to treatment. Baseline probability of contribution given by the number of years with strictly positive payment as percentage of total pre-treatment years. In columns (2) and (3), "reference point" refers to the amount of 15€. The dependent variable in column (4) is contribution in logs. In columns (5) and (6), the outcome is the probability of a reduction in the contribution relative to the baseline year 2012.

Table 2: Compliance Effects

Sample: Baseline Non-Contributors

	Effect on Probability of Contributing (%)	Effect on Probability of Contribution …		Effect on Contribution (%)
		… Below Ref. Point (%)	… Above Ref. Point (%)	
	(1)	(2)	(3)	(4)
Compliance Effects (Tax Letters vs. Donation Letter)				
Voluntary Tax	-3.24	-15.33	33.05	-.29
	(10.69)	(12.00)	(23.54)	(1.62)
Compulsory Tax	25.78**	37.52***	-9.46	3.56**
	(11.47)	(13.64)	(21.24)	(1.71)
Outcome in Omitted Reference Group	2.05%	1.53%	.51%	.37€
Number of Observations	24631	24631	24631	24631

Notes: OLS estimations at the individual level. *** denotes significance at 1%, ** at 5%, and * at 10% level. Robust standard errors in parentheses. All estimations account for the strata variables (age and household type) and include parish fixed effects. Baseline non-contributors did not make any contribution in pre-treatment years 2005-2012. "Contribution below (above) reference point" means contribution weakly below (strictly above) 15€. The dependent variable in column (4) is contribution in logs.

Table 3a: Crowding Out of Intrinsic Motivation: Combining Payment Data with Survey-Based Measures of Motivation

Sample: All Survey Respondents

	Effect on Probability of Contributing (%)			Effect on Contribution (%)		
	Motivation Measured by…			Motivation Measured by…		
	…Relation to Local Parish	...Frequency of Church Attendance	...Charitable Giving and Volunteering in other Contexts	…Relation to Local Parish	...Frequency of Church Attendance	...Charitable Giving and Volunteering in Other Contexts
	(1)	(2)	(3)	(4)	(5)	(6)
Crowding Out Effects, Survey (Tax Letters vs. Donation Letter)						
Voluntary Tax	-19.09*	-24.43**	-12.96	-38.72**	-51.42***	-32.32**
	(11.25)	(11.28)	(10.60)	(17.75)	(16.66)	(14.89)
Voluntary Tax x Motivation	6.60**	10.41***	4.54	12.41**	20.50***	9.97**
	(3.24)	(3.94)	(2.93)	(5.42)	(6.52)	(4.72)
Compulsory Tax	13.80	8.17	1.64	3.60	1.57	5.17
	(11.23)	(11.28)	(10.52)	(17.23)	(17.56)	(14.90)
Compulsory Tax x Motivation	-2.27	.49	1.38	.23	1.27	-.48
	(3.21)	(3.95)	(2.90)	(5.11)	(6.47)	(4.42)
Outcome in Omitted Reference Group		54.81%			17.16€	
Number of Observations		2321			2321	

Notes: OLS estimations at the individual level. *** denotes significance at 1%, ** at 5%, and * at 10% level. Robust standard errors in parentheses. All estimations account for the variables used to define strata in the experiment (age and household type) as well as parish fixed effects. The explanatory variables are treatment indicators and interactions between treatment indicators and measures for motivation taking values from 1 to 5, corresponding to the five ordered response categories for each of the motivational survey questions. Motivation is measured by individuals' stated relationship to their local parish (very weak = 1, weak = 2, undetermined = 3, close = 4, very close = 5), the stated frequency of church attendance (never = 1, less than once a month = 2, at least once a month = 3, at least once a week = 4, daily = 5), and individuals' stated charitable giving and volunteering in other contexts (very rarely = 1, rarely = 2, undetermined = 3, frequently = 4, very frequently = 5). The sample consists of all survey respondents, excluding those with missing values in either of the following variables: relation to local parish, church attendance, and charitable giving and volunteering in other contexts. Information on individual contributions was pre-coded on questionnaires prior to mail-out.

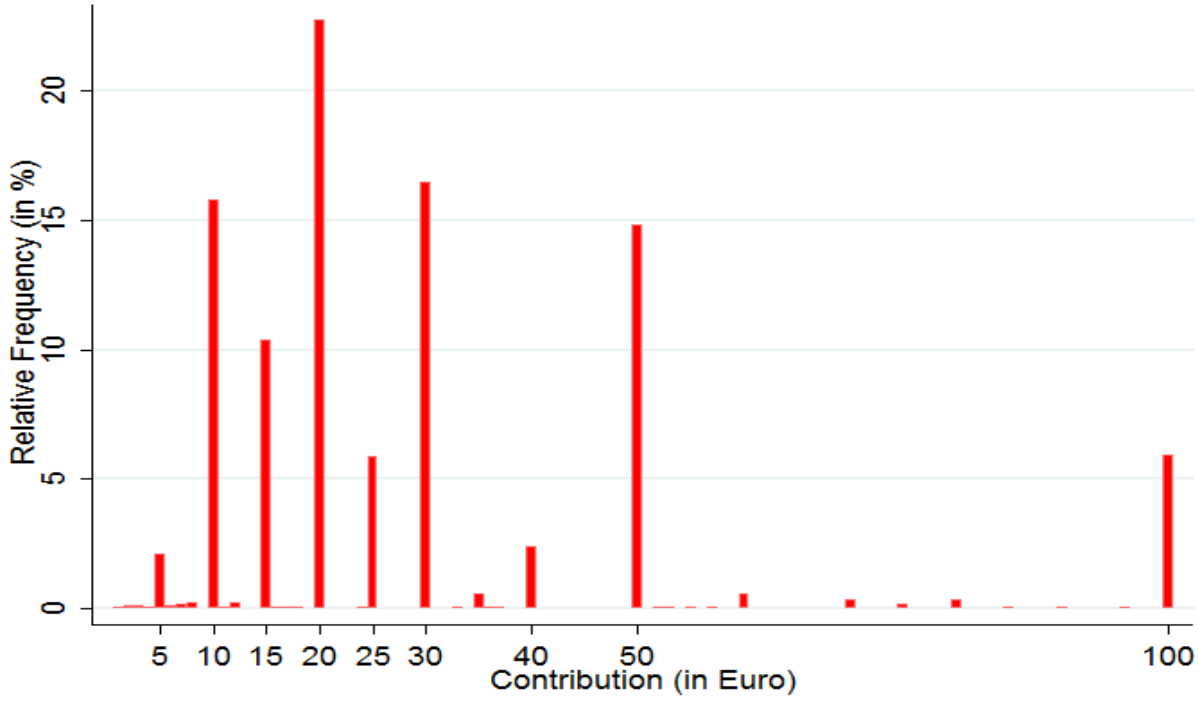
Table 3b: Crowding Out of Intrinsic Motivation: Evidence from Survey Responses

Sample: Survey Respondents from Voluntary Tax and Compulsory Tax Treatment Groups

	Effect on Probability for Response "Would Pay More" (%)						
		Motivation Measured by…					
		…Relation to Local Parish		...Frequency of Church Attendance		...Charitable Giving and Volunteering in Other Contexts	
	Full sample	Weak Intrinsic Motivation	Strong Intrinsic Motivation	Weak Intrinsic Motivation	Strong Intrinsic Motivation	Weak Intrinsic Motivation	Strong Intrinsic Motivation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Crowding Out Effect, Survey (Voluntary Tax vs. Compulsory Tax Letter)							
Voluntary Tax	32.69*	60.44**	3.30	86.70**	-1.86	107.26**	9.67
	(19.74)	(29.55)	(26.87)	(35.52)	(23.88)	(47.58)	(21.62)
Outcome in Omitted Reference Group	7.40%	6.61%	8.40%	5.36%	9.69%	4.20%	9.64%
Number of Observations	1525	855	670	823	702	650	875

Notes: OLS estimations at the individual level. *** denotes significance at 1%, ** at 5%, and * at 10% level. Robust standard errors in parentheses. All estimations account for the variables used to define strata in the experiment (age and household type) as well as parish fixed effects. The dependent variable is equal to 1 for individuals who state they would make a "somewhat higher" or "much higher" payment if the church levy was completely voluntary, and 0 otherwise. Intrinsic motivation is measured in various ways: columns (2) and (3) differentiate according to individuals' stated relationship to their local parish. The "weak intrinsic motivation" group comprises individuals who describe the relationship to their parish as "very weak, "weak" or "undetermined", whereas the "strong intrinsic motivation" group comprises those with a "close" or "very close" relationship. Columns (4) and (5) use individuals' stated frequency of church attendance. The "weak intrinsic motivation" group comprises individuals who say they attend church "less than once a month" or "never", whereas the "strong intrinsic motivation" group comprises those attending church "at least once a month", "at least once a week", or "daily". Columns (6) and (7) use individuals' charitable giving and volunteering in other contexts. The "weak intrinsic motivation" group comprises individuals who describe their charitable givings/volunteering as "very rarely", "rarely" or "undetermined", whereas the "strong intrinsic motivation" group comprises those with "frequent" or "very frequent" charitable givings/volunteering.

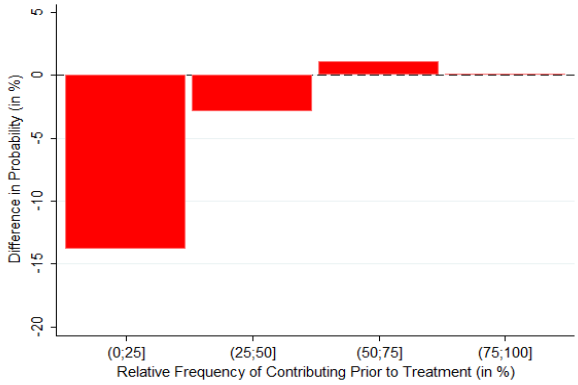
Figure 1: Baseline Distribution of Contributions in 2012



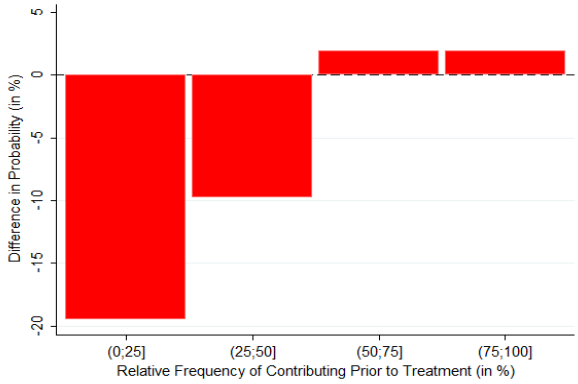
Notes: The figure displays the empirical density distribution of contributions made. More than 90% of contributions amounted to 10, 15, 20, 25, 30, 50 or 100 Euro (focal points). The sample consists of all baseline contributors (baseline year 2012, N = 4,817). The bin size is one Euro.

Figure 2: Heterogeneity in Crowding Out
Sample: Baseline Contributors (Intrinsically Motivated)

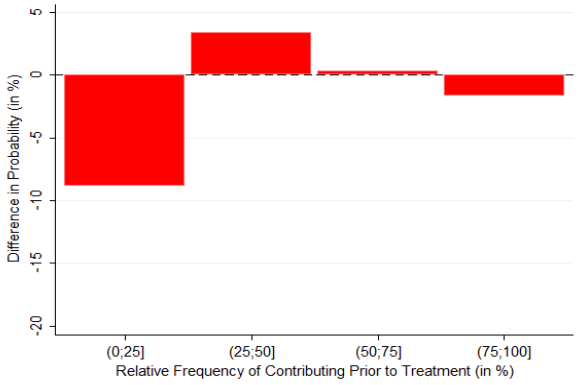
Panel A: Tax Letters, Pooled Effect
(Tax Letters - Donation Letter)



Panel B: Voluntary Tax Letter
(Voluntary Tax Letter - Donation Letter)

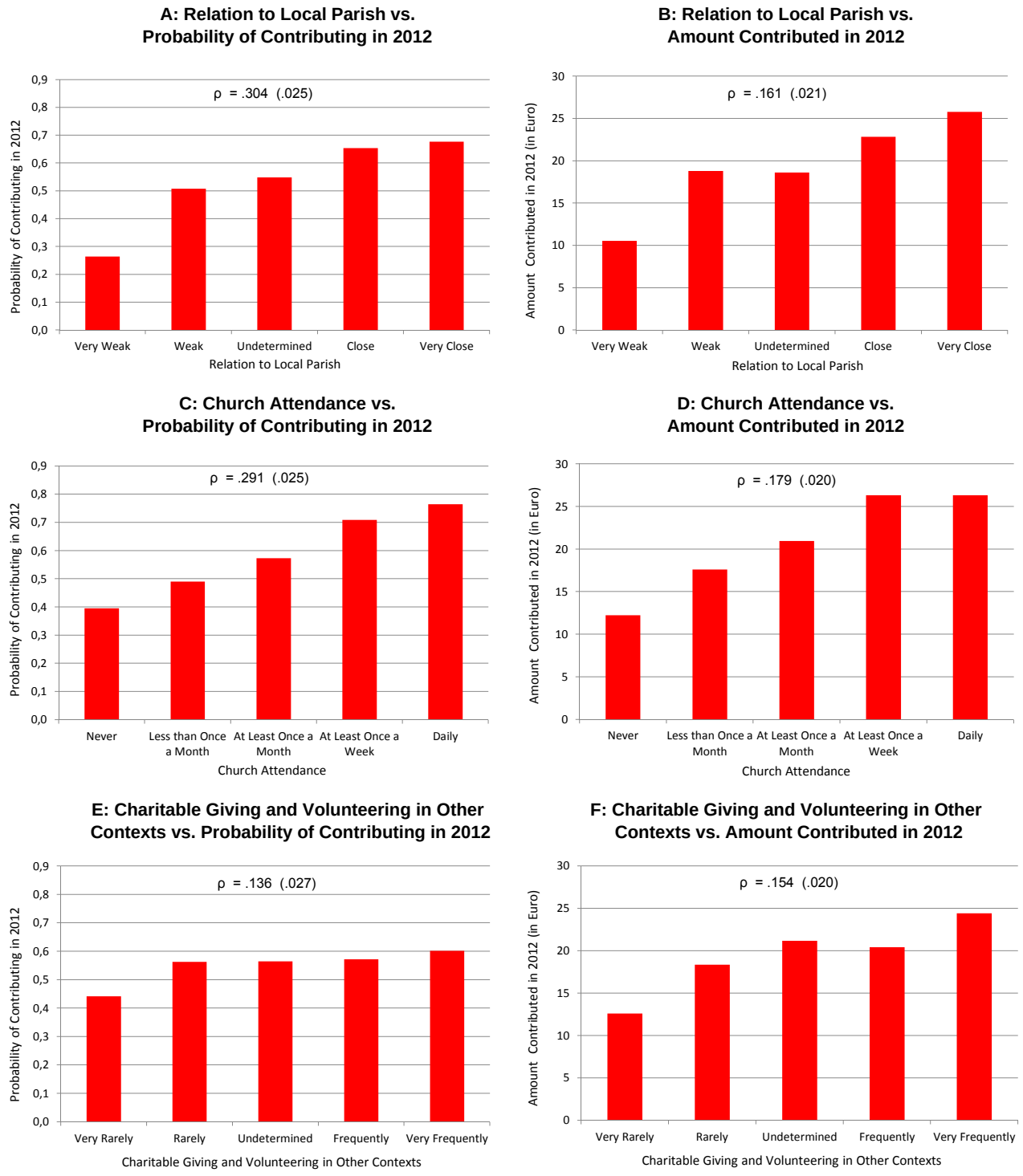


Panel C: Compulsory Tax Letter
(Compulsory Tax Letter - Donation Letter)



Notes: The figures display the difference in probability of contributing in %. Panel A shows the effect of communicating the legal norm by comparing the pooled tax letters (compulsory and voluntary tax letters) to the donation letter. Panel B shows the effect of the voluntary tax letter by comparing the voluntary tax letter to the donation letter. Panel C shows the effect of the compulsory tax letter by comparing the compulsory tax letter to the donation letter. The relative frequency of contributing prior to treatment is measured at the level of the individual as (# of years with strictly positive contribution - # of years solicitation letter was received)/# of years solicitation letter was received; the relative frequency is shown in percent. In all panels the sample is restricted to all individuals who have received at least three solicitation letters prior to treatment and who have contributed between 20 and 100 Euro at least once (N = 2,283).

Figure 3: Baseline Contribution Behavior by Relation to Parish, Charitable Giving, and Church Attendance



Notes: This figure shows the correlations between baseline contribution behavior in 2012 and survey responses on respondents' relation to their local parish (Panels A and B), respondents' church attendance (Panels C and D), and charitable giving and volunteering in other contexts (Panels E and F). In each panel, we distinguish five ordered response categories (Likert scale) on the x-axis. Panels A, C, and E report the average probability of contributing for each response category, while Panels B, D, and F report the average amount contributed for each response category. The sample consists of all survey respondents, excluding those with missing values in either of the following variables: relation to local parish, church attendance, charitable giving in other contexts, and amount contributed in 2012 (N = 2321). In each panel, we also report the polychoric correlation between the two motivational measures considered (ρ), with standard errors in parentheses. The Information on individual contributions in 2012 was pre-coded on questionnaires prior to mail-out.