American Economic Journal: Economic Policy 2016, 8(3): 203–232
http://dx.doi.org/10.1257/pol.20150083

Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany[†]

By Nadja Dwenger, Henrik Kleven, Imran Rasul, and Johannes Rincke*

We study extrinsic and intrinsic motivations for tax compliance in the context of a local church tax in Germany. This tax system has historically relied on zero deterrence so that any compliance at baseline is intrinsically motivated. Starting from this zero deterrence baseline, we implement a field experiment that incentivized compliance through deterrence or rewards. Using administrative records of taxes paid and true tax liabilities, we use these treatments to document that intrinsically motivated compliance is substantial, that a significant fraction of it may be driven by duty-to-comply preferences, and that there is no crowd-out between extrinsic and intrinsic motivations. (*JEL* C93, D64, H26, H71, K34, Z12)

Is tax compliance driven only by extrinsic motivations related to deterrence and tax policy, or is there also a role for intrinsic motivations such as morals, norms, and duty? The economic theory of tax compliance building on Becker (1968) and Allingham and Sandmo (1972) focuses only on the former and predicts low compliance under low audit probabilities or penalties. This prediction stands in sharp contrast to the empirical observation that tax compliance is high in modern tax systems despite very low audit probabilities and modest penalties. The literature has proposed three ways of resolving this compliance puzzle (e.g., Sandmo 2005; Slemrod 2007; Kleven 2014).

First, modern tax systems make widespread use of third-party information from firms and the financial sector, which creates a divergence between observed audit rates and actual detection probabilities conditional on evading (Kleven, Kreiner,

*Dwenger: Department of Economics, University of Hohenheim, 70593 Stuttgart, Germany (e-mail: nadja.dwenger@uni-hohenheim.de); Kleven: London School of Economics, Houghton Street, London WC2A 2AE, UK (e-mail: h.j.kleven@lse.ac.uk); Rasul: University College London, Drayton House, 30 Gordon Street, London WC1H OAX, UK (e-mail: i.rasul@ucl.ac.uk); Rincke: Department of Economics, University of Erlangen-Nuremberg, Lange Gasse 20, 90403 Nuremberg, Germany (e-mail: johannes.rincke@fau.de). We thank Alan Auerbach, Loukas Balafoutas, Oriana Bandiera, Richard Blundell, Raj Chetty, Michael Devereux, Denvil Duncan, Hilary Hoynes, Larry Katz, Judd Kessler, R.Vijay Krishna, Steve Levitt, Brian McManus, Paul Niehaus, Kathleen Nosal, Ricardo Perez-Truglia, Michael K.Price, Emmanuel Saez, Jonathan Shaw, Monica Singhal, Joel Slemrod, Sarah Smith, Christian Traxler, Shlomo Yitzhaki, anonymous referees, and numerous seminar participants for useful comments. We are grateful for financial support from the Schoeller Foundation and from the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS (grant number RES-544-28-5001). This research project was implemented in a Church District in a metropolitan area of Bavaria, and was approved prior to implementation by the Church District's administrative board.

[†]Go to http://dx.doi.org/10.1257/pol.20150083 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

and Saez 2009; Kleven et al. 2011). Hence, the notion that deterrence is weak is to some extent an illusion. Second, theory assumes that taxpayers have perfect knowledge of deterrence parameters, but in practice there may be misperception. Survey evidence suggests individuals tend to overestimate audit probabilities and penalties associated with tax evasion (Scholz and Pinney 1995; Chetty 2009). Third, individuals may comply due to a wide range of non-pecuniary motivations including moral sentiments, guilt, reciprocity, and social norms (Andreoni, Erard, and Feinstein 1998; Luttmer and Singhal 2014). We label all such motivations under the umbrella term *intrinsic motivations*. The importance of such intrinsic motivations for compliance is the hardest to measure and study empirically, and therefore the least well understood.

We consider a context and natural field experiment that provides new insights on the second and third explanations for the compliance puzzle. In our setting third-party information reporting is not implemented, and our field experiment is designed to reveal extrinsic and intrinsic motivations to comply through the provision of two forms of incentive: (i) the injection of positive deterrence; and (ii) the provision of compliance rewards/recognition.

Our setting is the local church tax in a metropolitan region of Bavaria, Germany. Three features of this setting are important for the empirical analysis. First, it combines taxation with charitable giving: the church tax is compulsory and noncompliance represents a violation of tax law, but the church highlights the good cause and encourages overpayments that are defined as donations. Hence, tax evaders and donors can coexist in this system. Second, the *true* tax base relevant for the church is defined as *reported* taxable income to the government, which we can perfectly observe for each individual by linking church tax records to administrative income tax records. This allows us to compare actual church taxes paid with true taxes owed for each individual, and thus precisely distinguish between evaders, compliers, and donors. This overcomes a key limitation of most tax evasion studies, namely that the outcome of interest is not observed (Slemrod and Weber 2012). Third, even though the church has the legal right to cross-check filed taxes against income tax returns (which would detect evasion with certainty), they have not previously exercised this right. In other words, prior to our field experiment there is zero deterrence in this tax system. Together with the previous point, this implies we can observe compliance in a baseline with zero deterrence, providing a direct measure of intrinsically motivated tax compliance.

To guide the empirical analysis, we set out a conceptual framework that unifies the standard compliance model (Allingham and Sandmo 1972) with the warm-glow model of public goods contributions (Andreoni 1989, 1990). The framework incorporates heterogeneity in intrinsic motivation to allow for the coexistence of evaders, compliers, and donors, as in our empirical setting. We use this to characterize the heterogeneous impacts of compliance incentives on evader and donor types. Our empirical analysis distinguishes throughout between the treatment responses of extrinsically motivated individuals (those who evade in the zero deterrence baseline) and the responses of intrinsically motivated individuals (those who comply or donate in the zero deterrence baseline). Our empirical measure of these motivational types is compelling, because our linked panel data from administrative tax records

and church records allows us to classify individuals into behavioral types using their actual pretreatment compliance behavior.

Our natural field experiment is implemented in collaboration with the Protestant church. We vary the compliance incentives individuals face by manipulating the official tax notification sent to collect the local church tax: 40,000 individual tax payers were randomly assigned to a control group or to 1 of 12 treatments. These treatments varied along three dimensions. The first set of treatments *simplify* the payment of the tax, and aim to correct any *misperceptions* individuals may have on audit probabilities. The second set of treatments vary the *deterrence* parameters individuals face. We do this through the announcement of strictly positive audit probabilities, including both fixed probabilities on all taxpayers and notched probabilities that depend on the tax payment. The third set of treatments offer *compliance rewards* in the form of social recognition, entry into monetary prize draws, or a combination of the two.

Our main findings are as follows. First, a significant fraction of individuals comply in the zero deterrence baseline where compliance should be zero absent intrinsic motivations. Around 20 percent of individuals pay at least the true taxes owed, while the remaining 80 percent of individuals evade taxes and most of them fully evade. Hence, intrinsic motivations can account for a substantial amount of aggregate compliance, but these motives are strongly heterogeneous in the population. The large majority of individuals behave as rational, self-interested taxpayers consistent with the Becker-Allingham-Sandmo framework.

Second, there is sharp bunching at *exact* compliance in the zero deterrence baseline. As there is no extrinsic incentive to locate at exact compliance under zero deterrence, such excess bunching requires either a discontinuity in intrinsic motivation at the point of exact compliance, naturally labeled as a "duty to comply," or the presence of attention or focal point effects of exact compliance. While it is in general difficult to distinguish between these explanations, we exploit our simplification treatment (which makes the point of exact compliance more salient) to shed light on this. We find that the simplification treatment does not increase bunching at exact compliance, suggesting that bunching may be driven more by duty-to-comply preferences.

Third, announcing a zero audit probability (the status quo) has only a small impact on compliance, suggesting there is little misperception on average. Less than 5 percent of baseline compliance can be attributed to misperception of the audit probability, and hence this is not an important confounder in the measurement of baseline intrinsic motivation.

Fourth, tax simplification and deterrence have strong effects on compliance for baseline evaders, but small and mostly insignificant effects for baseline donors. As the enforcement constraint is not binding for baseline donors, deterrence does not directly affect their extrinsic incentives to comply, and hence they should only respond to this treatment if there is crowd-out or crowd-in between extrinsic and intrinsic motivations. Our findings are therefore consistent with the absence of cross-effects between the two types of motivation.

Finally, the provision of compliance rewards has fundamentally different impacts on baseline donors (who increase their donations) and baseline evaders (who increase

their evasion). That is, whether recognition for compliance raises or reduces tax payments hinges on what motivates taxpayers in the first place, with positive effects on the intrinsically motivated and negative effects on the extrinsically motivated. These qualitative patterns arise irrespective of the exact form of the compliance reward, be it in terms of social recognition, entry into monetary prize draws, or a combination of both. This suggests that the behavioral effects are driven by what such compliance rewards signal about the tax institution rather than by the social/private nature of the reward. A natural interpretation is that rewarding taxpayers for contributing to the public good (rather than punishing them for not paying their taxes) signals the *voluntary* aspect of a poorly enforced tax system (and so positively affects the warm glow of donor types) and at the same time downplays the *mandatory* aspect of a legally binding tax system (and so may affect evader types negatively).

This paper contributes to the established literature on tax compliance (surveyed by Andreoni, Erard, and Feinstein 1998; Slemrod and Yitzhaki 2002; Slemrod and Weber 2012), and especially advances an emerging literature using field experiments to study compliance behavior (Blumenthal, Christian, and Slemrod 2001; Slemrod, Blumenthal, and Christian 2001; Kleven et al. 2011; Fellner, Sausgruber, and Traxler 2013; Hallsworth et al. 2014; Del Carpio 2014; Pomeranz 2015). Despite the large amount of work on compliance, there is very little field evidence on the relative importance of, and interaction between, extrinsic and intrinsic motivations to comply with taxes (see Luttmer and Singhal 2014). While we are able to make headway on this question due to the features of our data and setting, we note that these features may also raise issues of external validity. We discuss such issues in the next section.

The paper is organized as follows. Section II provides institutional background, Section III develops our conceptual framework, Section IV describes the experiment and data, Section V presents our empirical results, and Section VI concludes.

I. Institutional Background

The payment of church taxes is a legal obligation for all members of the Catholic and Protestant churches in Germany. There are two separate tiers of church taxes: the federal state and the church district levels. The state church tax is collected by state tax authorities, corresponds to around 9 percent of income tax liabilities, and raises billions of euros annually for both the Protestant and Catholic churches. The local church tax is collected by decentralized church authorities and is much smaller in size. The focus of our study is the local church tax collected by the Protestant church in a major metropolitan area in Bavaria, covering 68 parishes that comprise a Church District.[1]

By default, individuals baptized as Protestants (typically at birth) are church members and therefore liable to pay the local church tax once they turn 18. The vast

[1] The church tax is not unique to Germany: similar institutions exist in Austria, Denmark, Finland, Iceland, and Sweden. The local church tax also exists in other states in Germany (Saxony, Lower Saxony, and Rhineland-Palatinate). The fact that the local church tax represents only around 9 percent of total church revenues is in part due to widespread evasion as we show below.

majority of baptized individuals do *not* attend church as an adult: between 0.8 percent and 8.8 percent of eligible church members regularly attend church services in our sample parishes. Hence, our study is not based on an especially religious sample compared to the general population. We later provide evidence on the representativeness of our taxpayer sample.

We now describe three institutional features that are central to our study.

Tax Base and Tax Schedule.—The local church tax is a progressive income tax as shown in Figure A1 in the online Appendix. The schedule is a step function with an exemption level of €8,005 in annual income followed by 6 tax brackets in which the tax liability varies from €5 in the lowest bracket to €100 in the highest bracket. The tax base is a broad income measure (wages, business income, capital income, pensions, etc.) with no deductions. Importantly, the income components included in the church tax base are also taxable under the personal income tax and must be reported separately to state tax authorities. By defining the *true* taxable income for the church tax as *reported* taxable income for the personal income tax, the Church District is essentially leveraging on the far larger administrative capacity of the state tax authority. Reported taxable income may of course be subject to misreporting due to personal income tax evasion, but it is still defined as true income for the purposes of the church tax.

Tax Collection and Enforcement.—The Church District mails a tax notification (shown in the online Appendix) to all resident church members in May each year to collect the local church tax. A bank transfer form prefilled with the church's bank account information and the individual's local church tax number is attached to the notice. Church members are asked to self-assess their income and taxes owed according to the tax schedule, and to transfer the appropriate amount to the church's bank account by September. Although the church has the legal right to cross-check self-assessed income against information from personal income tax returns held by the state tax authorities (which would detect church tax evasion with certainty), they have never exercised this right in the past. In other words, prior to our field experiment, there was zero deterrence in this tax system and hence any compliance would have to be driven by some form of intrinsic motivation.[2]

Mandatory Taxes and Voluntary Donations.—It is possible for individuals to overpay their local church tax liability. Unlike conventional taxes, overpayments are encouraged and not refunded to individuals. As funds raised mostly remain within the parish, we can think of such overpayments as charitable donations to the local public good of parish services. This feature allows for the coexistence of tax evaders (who pay less than their legal obligation) and donors (who pay more than their legal obligation). We identify whether an individual is extrinsically or

[2] Individuals who do not pay their taxes before the September deadline receive a reminder in October requesting the transfer of the appropriate amount by the end of the calender year. If the payment has still not been made by the end of the year, no further action is taken by the Church tax authorities.

intrinsically motivated based on her actual past evasion/donation behavior under the zero enforcement regime.[3]

While these institutional features are useful for our empirical design, there is a trade-off with external validity: the features that make this setting well-suited to study motives for tax compliance are also features that distinguish our setting from other tax systems. Four potential threats to external validity are worth discussing. First, the tax is very small and this may affect compliance behavior, especially if inattention or other optimization frictions are important. We directly explore the potential role of attention/salience effects in one of our experimental treatments.

Second, the fact that the local church tax relies on zero enforcement may signal to taxpayers that, even though the tax is a legal obligation, church authorities do not consider it an important civic obligation. If so, this would undermine intrinsic motivation and imply that our finding of substantial intrinsically motivated compliance is *downward* biased relative to other tax settings.

Third, if the local church tax funds a service that taxpayers value more than the public expenditures funded by other taxes, this could raise intrinsic motivations to comply relative to other contexts. To address this point, we note that our estimates do not differ much across church parishes with varying levels of participation in religious services. Moreover, to reiterate, participation rates in church activities are uniformly low and the vast majority of those liable for the tax are not regular churchgoers as adults.

Fourth, contrary to other tax systems, in our context it is possible that individuals contribute through direct donations to the church instead of via the church tax system. This would lead us to underestimate intrinsic motivation in the baseline (as some tax evaders could be contributing directly) and potentially overestimate the effect of incentives on revenues (as some of the effect could reflect substitution between direct contributions and tax payments). However, these potential biases are unlikely to be important in our setting: (i) private individual donations to the church are very small in Germany (as in many other European countries), accounting for less than 4 percent of total revenues for the parishes in our sample; and (ii) at the parish level, there is little correlation between changes in private donations between 2011 and 2012 (the year of the field experiment) and the estimated aggregate change in tax revenues caused by our treatments.

II. A Warm-Glow Model of Tax Compliance

To guide the empirical analysis, we present a conceptual framework that unifies the standard deterrence model (Becker 1968; Allingham and Sandmo 1972) with the warm-glow model of public goods donations (Andreoni 1989, 1990). Our

[3] Besides encouraging overpayments (donations), the social pressures to comply with church taxes are not very different from those related to standard personal income taxes: whether an individual makes a payment to the local church tax remains private information, and individual or aggregate information on compliance is not communicated within or across parishes. Charitable giving is tax deductible in Germany, and this also applies to overpayments of the local church tax. Hence, there is no incentive to give to the church separately from the local church tax.

framework embodies both extrinsic motivations (deterrence) and intrinsic motivations (warm-glow) to comply with taxes.

We consider taxpayers with true income $\overline{z}$ facing a tax schedule $T(\overline{z})$ under truthful reporting. They decide on reported income z and tax payment $T(z)$ facing a probability of audit and penalty for evasion. Denoting consumption by c, utility is given by $u\big(c, T(z), s\big)$, where the inclusion of taxes paid $T(z)$ as an explicit argument captures the warm glow of giving, or intrinsic motivation, and s is a preference parameter capturing the strength of such intrinsic motivation. We assume that the marginal rate of substitution between intrinsic and extrinsic benefits u'_T/u'_c is increasing in s and equal to zero for $s = 0$. We allow for heterogeneity in intrinsic motivation, captured by a cdf $F(s)$. The Allingham-Sandmo model corresponds to the special case where all individuals have $s = 0$.[4]

Agents choose reported income z to maximize expected utility, which can be written as

$$(1)\qquad (1-p)\cdot u\big(\overline{z} - T(z), T(z), s\big) + p\cdot u\Big(\overline{z} - T(z) - I\{z<\overline{z}\}[1+\theta]\big[T(\overline{z}) - T(z)\big], T(z), s\Big),$$

where p is the audit probability, θ is the penalty rate on tax evasion, and $I\{z<\overline{z}\}$ is an indicator for evading taxes. This specification naturally assumes that warm glow depends on the *voluntary* tax payment $T(z)$ in both the audited and unaudited states. That is, an evader does not obtain warm glow from being forced to pay additional taxes $T(\overline{z}) - T(z)$ due to an audit.

Consistent with our empirical setting, the model allows for taxpayers to fall in three different categories: Those who underpay taxes $T(z) < T(\overline{z})$ (evaders), those who pay exactly the right amount $T(z) = T(\overline{z})$ (compliers), and those who overpay taxes $T(z) > T(\overline{z})$ (donors). Changes in extrinsic or intrinsic incentives create movements across these three compliance categories (extensive margin) and reporting responses within the evasion and donor categories (intensive margin).

Consider first the intensive margin choice of z, which is governed by

$$(2)\qquad (1-p)u'_{c_N} + p\big(1 - I\{z<\overline{z}\}[1+\theta]\big)u'_{c_A} = E\big[u'_T\big],$$

where u'_{c_N} and u'_{c_A} denote marginal utilities of consumption in the non-audited and audited states, respectively, while $E\big[u'_T\big]$ is the expected marginal utility of tax payments due to intrinsic motivation. This condition highlights the trade-off between the extrinsic (consumption) costs and the intrinsic (warm glow) benefits of paying taxes.[5] In online Appendix A.1 we formally characterize intensive margin responses

[4] Allingham and Sandmo (1972) did consider a case with social stigma from being caught evading, but the stigma idea is conceptually different from the warm-glow idea analyzed here.

[5] In the Allingham-Sandmo model, where $s = 0$, we have $E\big[u'_T\big] = 0$ and $I\{z<\overline{z}\} = 1$, in which case (2) simplifies to the standard condition $u'_{c_A}/u'_{c_N} = (1-p)/(p\theta)$.

to changes in deterrence and the strength of warm glow. We show that, under a natural assumption on preferences, deterrence increases reported income for evaders while it does not affect reported income for donors. The differential deterrence response between evaders and donors follows from the fact that enforcement is not a binding constraint for donors.

Consider now the extensive margin choice between being an evader, complier, or donor. The model predicts bunching at the point of exact compliance $z = \overline{z}$ due to the fact that evaders and donors are treated asymmetrically: In the event of an audit, evaders have to pay the unpaid tax topped up by the penalty rate θ, whereas donors are not reimbursed for the overpaid tax nor rewarded at rate θ. This asymmetry creates a *kink* in the consumption possibility set at $z = \overline{z}$ and produces excess bunching at this point. Formally, assuming smooth preferences, there exists cutoffs $\overline{s}_1, \overline{s}_2$, such that a fraction $F(\overline{s}_1)$ of the population are evaders ($z < \overline{z}$), a fraction $F(\overline{s}_2) - F(\overline{s}_1)$ are compliers ($z = \overline{z}$), and a fraction $1 - F(\overline{s}_2)$ are donors ($z > \overline{z}$). The cutoffs are given by

$$(3)\quad \frac{u_T'\big(\overline{z} - T(\overline{z}), T(\overline{z}), \overline{s}_1\big)}{u_c'\big(\overline{z} - T(\overline{z}), T(\overline{z}), \overline{s}_1\big)} = 1 - p\big[1 + \theta\big] \quad \text{and} \quad \frac{u_T'\big(\overline{z} - T(\overline{z}), T(\overline{z}), \overline{s}_2\big)}{u_c'\big(\overline{z} - T(\overline{z}), T(\overline{z}), \overline{s}_2\big)} = 1,$$

implying $\overline{s}_1 < \overline{s}_2$ and therefore excess bunching at $z = \overline{z}$ for any positive deterrence incentive, $p\big[1 + \theta\big] > 0$. In online Appendix A.1, we characterize extensive margin responses to changes in deterrence and the strength of warm glow. We show that stronger deterrence reduces the fraction of evaders, increases the fraction of compliers (bunching), and does not affect the fraction of donors. Stronger warm glow, on the other hand, reduces the fraction of evaders and increases the fraction of donors, leaving the effect on the fraction of compliers indeterminate.

Our empirical setting starts from a baseline of zero deterrence in which the tax authority never audits ($p = 0$). In this case, equation (2) shows that reported income z satisfies $u_T'/u_{c_N}' = 1$ for each taxpayer, so that compliance is driven solely by intrinsic motivation. Furthermore, from equation (3), we have $\overline{s}_1 = \overline{s}_2$ and therefore zero excess bunching at exact compliance. Empirically, however, we find strong bunching at exact compliance even in the zero deterrence baseline. There are two potential reasons for this that can easily be incorporated in the model. The first possibility is that intrinsic motivation (warm-glow preferences) feature a discontinuity at $z = \overline{z}$. This would be the case if taxpayers are discretely more motivated to be law-abiding than to be marginal evaders, naturally labeled *duty-to-comply* preferences. This could be accounted for by allowing for a discrete jump in warm glow (a notch) at exact compliance. The second possibility is that exact compliance is a focal point and that bunching is therefore driven by attention or salience effects. As we precisely measure compliance in the zero deterrence baseline, we are able to estimate the amount of such intrinsically motivated bunching and to use our experiment to explore if it is driven by duty-to-comply or attention.

III. Design, Data, and Empirical Method

A. *The Natural Field Experiment*

The Protestant church mails out a tax notification for the local church tax in May of each year. Our field experiment manipulated the content of notifications sent out in 2012. Mail-out recipients were randomly assigned either to a control group or one of three groups of treatment. The first group of treatments *simplify* the details of the tax and correct any *misperception* individuals might have on audit probabilities. The second group of treatments manipulate *deterrence* parameters through the suggestion of strictly positive audit probabilities or an audit probability notch. The third group of treatments offers *compliance rewards/recognition*.[6]

The online Appendix shows the format and content of the mail-out letter for the control group (T1). The same mail-out design had been used in earlier years. This standard notification comprises a cover page (with the remittance slip at the foot of the first page) and an information leaflet about church activities. The standard mail-out clearly states on the front page that, "the local church tax forms part of the general church tax," and that the "letter serves as a tax certificate." On the second page it makes precise that the tax is "a compulsory contribution" and explicitly lists the legal foundations for the tax. However, in other regards, the standard mail-out appears poorly designed: important details, such as the payment deadline and tax schedule, are only mentioned on the second page. We now describe how the mail-out design varied in each treatment group. Table A1, at the end of this paper, overviews all the treatments and provides the exact wording used in each.[7]

Treatment Group 1: Tax Simplification and Misperception.—The tax simplification treatment (T2) makes two changes to the tax notification design: (i) it is significantly shorter and makes salient the legal obligation to pay; (ii) payment deadlines and the tax schedule are presented on the cover page. All other design aspects remained unchanged relative to the control group. We might reasonably expect tax simplification to impact baseline evaders because some noncompliance might be driven by them being misinformed/inattentive toward the local church tax.

All subsequent treatments then add one paragraph on the cover page of this simplified mail-out (as shown in the online Appendix). While it is well known among taxpayers that enforcement is lax in this setting, the misperception treatment (T3) aims to correct for any remaining misperception by making explicit that there is zero enforcement of the tax. This is communicated by explicitly stating that $p = 0$. We assigned twice as many individuals to this treatment than to any other treatment

[6] Following procedures from earlier tax years, a reminder was sent to non-payers in October 2012. The reminder letter is the same for all and makes no mention of the original treatment assignment. The reminder sets a final payment deadline of December 31, 2012.

[7] Cagala, Glogowsky, and Rincke (2015) present evidence from a small-scale survey among a random sample of those liable for the Catholic church tax in Bavaria: almost 90 percent of those receiving a tax notification confirmed they had read it. Hence, while our analysis focuses throughout on intent-to-treat effects, the corresponding average treatment effects should only be slightly scaled up.

to ensure we had statistical power to detect changes in tax payments arising from potential misperception. The natural comparison is with T2.

As with the simplification treatment, we might expect responses to the misperception treatment to vary across taxpayer types: some baseline compliers might have been paying the tax because they previously perceived p to be larger than zero. By making explicit that $p = 0$, the treatment intends to fully eliminate extrinsic motivation for compliance, so that tax payments can only be driven by some form of intrinsic motivation. T3 therefore allows us to cleanly estimate the importance of such intrinsically motivated compliance.

Treatment Group 2: Deterrence.—The second group of treatments inject deterrence into the tax system. They do so by informing taxpayers the audit probability p is unconditionally set to some strictly positive value, namely $p = 0.1$, 0.2, or 0.5. These p-treatments are denoted T4, T5, and T6, respectively, and make clear that the church district has the legal right to delegate tax enforcement to the church tax authorities, to whom a tax filer's income is known. The natural comparison group for these $p > 0$ treatments is the $p = 0$ treatment, so that we pin down the precise comparative static impacts of deterrence through Δp.

These p-treatments were truthfully implemented in that income self-assessment was verified, but in practice no monetary penalty followed if the individual was caught misreporting. Like previous tax enforcement field experiments, we do not observe individual beliefs about penalties. These beliefs are particularly difficult to gauge in our context, because the zero-audit policy of the church implies that taxpayers have never had to face penalties. However, the conceptual framework makes precise that any behavioral response to $p > 0$ must reflect a positive expected penalty.

A final deterrence treatment (T7) introduces an audit probability notch: individuals face an audit probability of 0.5 if they pay less than or equal to €10, and face a zero audit probability otherwise. There are two natural comparison groups to this notch treatment: the T3 misperception treatment that sets $p = 0$, and the T6 treatment that sets $p = 0.5$ for all payments.

Treatment Group 3: Compliance Rewards.—The final group of treatments are designed to reveal motivations for compliance through the provision of rewards/recognition. These treatments differ in the exact form in which the reward for compliance is provided. The first offers a potential reward in the form of *social recognition* (T8), through a small probability of an individual's timely compliance being publicly announced in a local newspaper. The next two treatments offer entry into *monetary* prize draws as a reward for complying, a purely *private* form of recognition that is unannounced to others. There are two randomly assigned reward values (€250, €1,000), denoted Treatments T9 and T10. The final form of reward combines social and private recognition for compliance, so taxpayers have the opportunity to be named in a local newspaper *and* to be entered in the higher valued prize draw. This treatment is denoted T11.[8]

[8] Rewards were offered for payments of at least €5 (not the true payment owed) to prevent individuals inferring any change in likelihood of being audited. The winners of all rewards were drawn by lot, before local church

For all these compliance rewards, the probability of winning the reward is close to zero: for the social recognition treatment this follows from the fact that many individuals pay some church tax and are therefore potentially eligible for the newspaper acknowledgment; for the monetary reward treatment the notification makes explicit that the probability of winning is $1/1{,}000$. As such, these compliance rewards have essentially no impact on the (expected) extrinsic incentives individuals face to comply, and so they should change compliance only if they impact intrinsic motivation. In particular, individuals may respond to the offer of such rewards if they affect perceptions about the nature of the tax institution. Indeed, a natural interpretation of such treatments is that rewarding taxpayers for contributing to the public good (rather than punishing them for not paying their taxes) signals the *voluntary* aspect of a poorly enforced tax system, and at the same time downplays the *mandatory* aspect of a legally binding tax system. If so, compliance rewards may have heterogenous impacts across baseline types, with donor types being encouraged to respond positively and evader types being more negatively impacted.

Finally, we also implemented treatments that provide information on social norms over compliance, or that use moral suasion. The literature has considered very similar treatments (Blumenthal, Christian, and Slemrod 2001; Fellner, Sausgruber, and Traxler 2013) and so we do not focus on them. In the online Appendix, we discuss fully the weak effects of such cheap talk letters, very much replicating findings in the literature.

B. *Data Sources*

Our analysis links panel data from two administrative data sources: church district records containing *actual* church taxes paid by each individual $T(z)$, and state income tax records containing *true* church taxes owed $T(\overline{z})$ as implied by reported taxable income to the federal state. Church taxes due in year t depend on reported taxable income to the federal state in year $t-1$. The church district's payment records cover 2008–2012, which we have linked with the state's income tax records for 2007–2011 using information on names, date of birth, and zip code. The linked sample consists of 39,782 individuals that are included in the field experiment.[9]

Table A2 in the online Appendix presents evidence on the representativeness of our sample relative to other subgroups of tax filers in 2007, the last year for which

officials in December 2012, and immediately notified about their prize. Winners of the social reward had to provide consent for their name to be published. The advertisement thanking church members for their local church tax payment was published in early 2013 (after the final payment deadline of December 31, 2012 to avoid any impact on outstanding payments). Monetary prizes were paid in private in January 2013.

[9] Our administrative tax records allow us to observe tax compliance behavior across the income distribution. As the lower portion of Figure A1 highlights: 29 percent of our sample have an income below €24,999 (falling into the first two payment bins), while 13 percent of the sample have an income above €70,000 and lies in the highest payment bin. There are two restrictions on the data linkage. First, administrative records are available only for those that file a tax declaration. In the area our study is based in, 60 percent of Protestants file a tax declaration. Second, the tax base for the local church tax is *individual* taxable income. This raises an issue among joint filers: in the administrative records, individual shares of taxable household income are available only for joint filers who belong to different religious denominations. Hence, we exclude married couples in which both spouses are Protestants (thus ameliorating concerns over within household treatment spillovers). Given the advice of the church, we also excluded individuals 75 years old or older from the field experiment.

nationwide personal income tax statistics are available.[10] There are only minor differences in gender, age, children, taxable incomes, and income sources between our sample (columns 1a, 1b) and: (i) the general population in the same metro area (columns 2a, 2b); and (ii) non-church members in the same metro area (columns 3a, 3b). These similarities are not altogether surprising: those liable for the church tax are individuals baptized, typically at birth, into the church; as adults, the vast majority of them do not attend church regularly and hence our sample is not skewed toward overrepresenting religious individuals.

The other sampling concern relates to attrition from our linked panel. Individuals can attrit for multiple reasons: falling below the tax exemption threshold, relocating outside the Church District, not filing a tax return, or opting-out of the Protestant church. This last cause is of most concern for the interpretation of our results. However, rates of attrition are relatively low: less than 3 percent of individuals attrit each year for *any* reason, and 87 percent of individuals are observed in *all* years 2008–2012. In the online Appendix, we provide evidence on the correlates of attrition, and summarize those findings as showing: (i) attrition is *uncorrelated* to treatment assignment, and (ii) there is no differential attrition across treatments by past compliance behavior. Our working sample is based on those 89 percent of individuals (35,603) for whom we observe taxable income for up to four years pretreatment (2008–2011).

Individuals were randomly assigned to treatment within strata.[11] Online Appendix Table A4 presents evidence on sample characteristics and balance across treatments. Around 51 percent of our taxpayer sample are men, the average age is 45, 42 percent are married, half have at least one child, and average taxable income is €43,000. Column 10 shows an F-test of the significance of the covariate set from being assigned to that specific group relative to the T1 control group (in brackets) and the T2 Tax Simplification (in braces). The evidence shows the samples are well balanced across treatments.[12]

C. *Identifying Evaders, Compliers, and Donors*

As we observe both actual tax payments $T(z)$ and true taxes owed $T(\overline{z})$, we can precisely measure compliance at the individual level and therefore estimate

[10] In Germany, individuals are obliged to file a tax return if they receive business income or income from self-employment: around 38 percent of the population files a tax return. Single filers comprise unmarried individuals and married couples who choose to file two separate tax returns. The vast majority of married couples are joint filers and benefit from the associated reduction in the progressivity of the personal income tax. One parent of each underage child is entitled to child allowances. Tax raising communities in Germany refer to religious communities that collect taxes within the scope of the personal income tax. The Protestant and Catholic churches are by far the largest tax raising communities and cover around 60 percent of the population.

[11] Two randomization strata were used: (i) the individual's church tax bracket in 2011; (ii) the number of pretreatment years the individual is observed for in the administrative records.

[12] The other key identifying assumption is that there are no spillovers across treatments. Four points bolster the credibility of our design on this point: (i) on within-household spillovers, we reiterate that our sample matched to administrative tax records only covers households in which one spouse is Protestant; (ii) individuals in the church hierarchy were excluded from the field experiment, including administrative staff, priests, and a few historically generous donors; (iii) there was no media coverage of the field experiment; (iv) we set up a telephone inquiry line for individuals to call in case they had any comments/queries after receiving their tax notification: this received 162 calls in total (corresponding to 0.34 percent of treated individuals), with queries mostly relating to the tax base.

compliance responses to the different experimental treatments. What is more, our panel data allows us to measure pretreatment compliance behavior from 2008–2011, a period with zero tax enforcement and therefore no extrinsic incentive to comply. This allows us to identify *baseline compliance types* under zero enforcement: baseline evaders are those who underpay pretreatment, baseline compliers are those who pay exactly the right amount, and baseline donors are those who overpay. These categories then proxy for motivational types, with baseline compliers/donors being *intrinsically motivated* and baseline evaders being *extrinsically motivated*. The ability to distinguish between these different types enables us to study heterogeneous treatment effects with respect to motivation, thereby speaking to the interaction between extrinsic and intrinsic incentives, and compare with our conceptual framework, which predicts that those effects should be strongly heterogeneous.

While information on past behavior can be combined in many ways to define baseline types, we use a simple approach based on individual behavior in 2011, the year immediately preceding our field experiment.[13] Using one year of data to identify baseline types is reliable in our setting because of a high degree of persistence in individual behavior across years. To see this, note that for the balanced panel of individuals in our control group that are observed for *all* years 2008–2011 that evaders in 2011 had on average evaded for 2.79 out of the previous three years, while compliers/donors in 2011 had on average complied/donated for 2.09 out of the previous three years. Table A5 documents the high degree of persistence in individual behavior over time using a multinomial logit model. To summarize, we find: (i) the best predictor of current compliance type is lagged type, for example, those who evaded in 2010 are 87 times more likely to evade in 2011 relative to complying; and (ii) most other covariates have no predictive power on being an evader or a donor relative to a complier.[14]

D. *Empirical Method*

We first consider extensive margin responses to the different treatments, estimating a linear probability model,

$$\text{Prob}\,(i \text{ evades}) = \alpha + \beta I(T_i = j) + \pi E_{i,pre} + \lambda_s + u_i, \tag{4}$$

where $I(T_i = j)$ is an indicator equal to one if individual i is assigned to treatment j, $E_{i,pre}$ is the number of times individual i has evaded in the pretreatment years, λ_s are dummies for randomization strata, and u_i is an error term. The coefficient of interest β measures the percentage point impact of treatment on the probability of

[13] Columns 11–13 in Table A4 show the samples across treatments to be balanced *within* each of these baseline types.

[14] Older individuals are significantly more likely to donate. Those with wage income or liable for trade tax (a proxy for being an entrepreneur) are significantly more likely to evade, all else equal. However, the marginal impacts of these covariates are far smaller than the impact of the individual's own past compliance. If there are high transaction costs of compliance, individuals might periodically pay large amounts so to, on average over time, pay the total payment owed. To check for this we examined whether those that donate in any given tax year are significantly less likely to make a payment the following year: we find no evidence for this pattern of payments.

evasion.[15] We estimate an analogous specification for the probability of donating as a function of treatment (conditioning on the number of times the individual donated in pretreatment years, $D_{i,pre}$).[16]

We also consider total responses that combine the extensive and intensive margins. Here, we estimate the OLS specification

$$y_i = \delta + \gamma I(T_i = j) + \theta \overline{y}_{i,pre} + \lambda_s + \varepsilon_i, \tag{5}$$

where y_i is the tax payment of individual i post-treatment, $\overline{y}_{i,pre}$ is the average tax payment pretreatment, and $I(T_i = j)$ and λ_s are as defined above. In addition to tax payments, we also consider a coarser compliance outcome that gives us more statistical power: a dummy variable equal to one if the individual *increases* the tax payment over its pretreatment level.

IV. Empirical Results

A. *Compliance in the Zero Deterrence Baseline*

We begin by exploiting an important feature of our setting: that we can accurately measure tax compliance in a legally binding tax system with a zero deterrence baseline. If such zero deterrence is well understood (as we largely confirm below), there should be zero compliance absent intrinsic motivations to pay taxes. Table 1 documents compliance in the baseline using data from the T1 Control group. Column 1 shows the full sample, while columns 2 and 3 split the sample into evaders (the extrinsically motivated) and compliers/donors (the intrinsically motivated). Three points are of note.

First, a significant fraction of individuals comply in the zero deterrence baseline: 20.9 percent of individuals make a payment greater than or equal to their true tax liability, while the remaining 79.1 percent make a payment smaller than their true tax liability. Second, among the evaders, 91.9 percent of them are *full* evaders and pay zero tax, while the remaining 8.1 percent are *partial* evaders and pay some tax. Third, among those that make at least the correct payment, 55.5 percent are exact compliers and 44.5 percent are donors.

These findings have implications for the compliance puzzle debate. The fact that almost 80 percent of individuals evade and 73 percent fully evade in the zero deterrence baseline implies that the Becker-Allingham-Sandmo framework is 70–80 percent correct in our setting. At the same time, there coexists a substantial fraction of individuals whose compliance is driven by some form of intrinsic motivation not captured by the standard model: about 20 percent comply or overpay and about 27 percent pay at least something even though the tax system is completely

[15] If we leave out pretreatment compliance $E_{i,pre}$ and strata fixed effects λ_s, the specification corresponds to a simple comparison of means across treatment groups. We consider this unconditional specification in the online Appendix, and show that it gives very similar results as (4).

[16] All the extensive margin results reported are also robust to estimating a multinomial logit model for choice type k (evader, complier, donor), conditioning on treatment assignment, the number of times individual i has been of type k (evader, complier, donor) in the pretreatment years, and dummies for the randomization strata.

Table 1—Compliance under Zero Deterrence

Control group, means	Full sample (1)	Evaders (extrinsically motivated) (2)	Compliers/donors (intrinsically motivated) (3)
Number of individuals	2,532	2,004	528
Percentage of all individuals	100%	79.1%	20.9%
Full evaders	72.7%	91.9%	—
Partial evaders	6.4%	8.08%	—
Compliers	11.6%	—	55.5%
Donors	9.3%	—	44.5%
Payment amount	€10.32	€1.87	€42.40

Notes: The sample of individuals are all those assigned to the T1 control group in 2012 (2,532 individuals). The column headings refer to behavior in 2012, the year of the field experiment. Evaders are defined as those who pay strictly less than their legal tax liability, compliers are those who pay exactly their legal tax liability, and donors are those who pay strictly more than their legal tax liability.

unenforced. Hence, both sides of the compliance puzzle debate may feel justified: while the Becker-Allingham-Sandmo model is a good approximation for the majority of taxpayers, it does leave out a nontrivial element of intrinsically motivated tax compliance.[17]

Duty-to-Comply.—As the conceptual framework makes clear, individuals have no extrinsic incentive to bunch at exact compliance in the zero deterrence baseline. Such bunching requires either a discontinuity in intrinsic motivation at exact compliance, naturally labeled as a "duty-to-comply," or that the point of exact compliance represents a focal point for intrinsically motivated taxpayers.

Figure 1, panel A presents evidence on such bunching by showing, for those that make a positive payment, the histogram of differences between taxes paid and taxes owed in the T1 Control group. This shows large and sharp bunching at $T(z) = T(\overline{z})$ despite no extrinsic incentive to locate there. We use the bunching methodology developed by Saez (2010) and Chetty et al. (2011) to quantify the amount of excess bunching at exact compliance: the bunching estimate shown in the figure, $b = 7.2$, implies that over 7 times as many taxpayers are observed at exact compliance than would be otherwise expected given smooth preferences as inferred from other parts of the distribution of $T(z) - T(\overline{z})$. The strong tendency of intrinsically motivated taxpayers to comply exactly with the letter of the law can also be gauged from Table 1. This table shows that, among the 20 percent of individuals who feature some form of intrinsic motivation to comply, more than half of them locate at the point of exact compliance.

[17] We note that the compliance/donation rate of 20.9 percent to the local church tax is far higher than those typically observed in large-scale field experiments on charitable giving, where response rates typically vary between 2 percent and 5 percent for fundraising campaigns, despite those campaigns often being targeted to those with affinity toward the charitable cause (Karlan and List 2007; Huck, Rasul, and Shephard 2015). This suggests the local church tax is not viewed merely as a form of charitable donation, and that the *legal* obligation to pay has significant bite. This is reaffirmed if we recall that the vast majority of baptized individuals do not participate in church activities (with attendance rates less than 5 percent in the average parish). Hence, intrinsic motivation does not appear entirely due to behaviors confined to the religious.

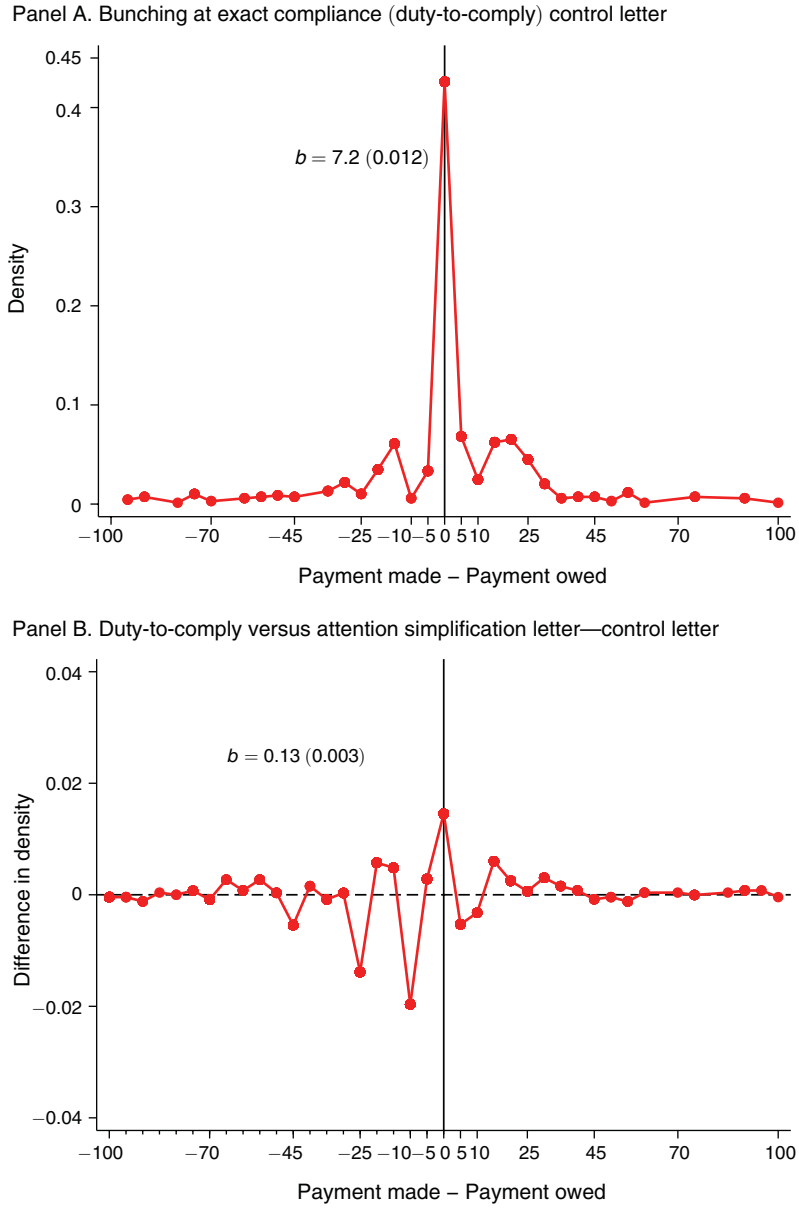

FIGURE 1. COMPLIANCE DISTRIBUTION UNDER ZERO DETERRENCE

Notes: Panel A displays the raw distributions of the difference between payment made and payment owed for the control letter. In panel A, the sample consists of compliers and donors with strictly positive payments. Panel B plots differences in the densities of tax payments made–tax payments owed between different treatment groups. The objective is to see if bunching at exact compliance (demarcated by the vertical line at zero) responds to the tax simplification treatment. Hence, panel B shows the difference between the T2 simplification group and the T1 control group. In panel B, the sample consists of all individuals. The bin size in both panels is 5 euro.

It is conceptually difficult to distinguish between duty-to-comply and attention/focal point explanations for the observed bunching at exact compliance, and to some extent this can be viewed as a matter of labeling rather than substance. Nevertheless, in the next section we attempt to make progress on the distinction between the two explanations by considering how bunching changes in response to our simplification treatment, which makes the point of exact compliance more salient.

B. *Compliance Responses to Treatment*

Table 2 presents our core results on how tax compliance is causally affected by tax simplification (panel A), misperception (panel B), deterrence (panel C), and compliance rewards (panel D). For each panel we show both extensive margin and total responses in three samples: the full sample, baseline evaders (extrinsically motivated), and baseline donors (intrinsically motivated). The full sample results are presented in Part I of Table 2, while the results for baseline evaders and baseline donors are shown in Part II. The extensive response estimates are based on the linear probability model (4) for the probability of evading and an analogous specification for the probability of donating. The total response estimates are based on the specification in (5): the outcomes we consider are the total tax payment and the probability of increased payment. All treatment effects are reported as percentages of the average outcomes in the relevant comparison group, and at the foot of each panel we show the average outcome in the comparison group.

Tax Simplification.—Panel A of Table 2 presents the results of the T2 Tax Simplification treatment. Pooling all taxpayers, columns 1–4 show that simplification (i.e., making salient the legal obligation to pay and making deadlines and the tax schedule more prominent) significantly reduces the probability of evasion by 2.45 percent, and causes individuals to significantly increase tax payments by 9.73 percent. The remaining columns in panel A show the effects on both margins to be driven by baseline evaders (columns 5–8). In this subsample of taxpayers, simplification of the tax notification significantly reduces the probability of evasion by 2.66 percent, and increases payment amounts by 43.4 percent. The fact that extrinsically motivated individuals are not more likely to donate (column 6) highlights that the primary response to the simplification treatment is largely driven by such individuals changing their behavior from being full evaders to being exact compliers.

On baseline donors, columns 9–12 show tax simplification has no significant impact on either margin of behavior. These null impacts suggest their tax compliance is not driven by them being confused. All the findings are robust to unconditionally estimated treatment effects as documented in the online Appendix (Table A6).[18]

Taken together, the results of the tax simplification treatment imply that a considerable degree of tax evasion may be due to the complexity of tax notifications. Our

[18] We further note that all these findings are additionally robust to: (i) controlling only for randomization strata; (ii) excluding controls for pretreatment behaviors; (iii) additionally controlling for the full set of individual controls shown in the balancing Table A4; and (iv) restricting the sample to the balanced panel of individuals observed in all tax years 2007–2010.

Table 2—Treatment Effects on Compliance—Part I

Control group, means	Full sample			
	Probability of evading (1)	Probability of donating (2)	Payment amount (3)	Probability of payment increase (4)
Panel A. Tax simplification				
Simplification versus control				
Effect of tax simplification	−2.45	−0.438	9.73	33.61
	(0.971)	(6.90)	(3.73)	(10.25)
Average outcome in comparison group	79.29%	9.24%	€10.29	7.89%
Observations	5,076	5,076	5,076	5,076
Panel B. Misperception				
Zero audit probability versus simplification				
Effect of correcting misperception	0.942	−7.23	−0.766	−10.60
	(0.889)	(5.65)	(3.05)	(6.75)
Average outcome in comparison group	77.30%	9.75%	€11.65	10.92%
Observations	7,641	7,641	7,641	7,641
Panel C. Deterrence				
Positive audit probability versus zero audit probability				
Effect of deterrence	−3.13	13.71	10.45	26.93
	(0.660)	(4.59)	(2.37)	(5.84)
Average outcome in comparison group	78.04%	8.93%	€11.63	9.42%
Observations	12,692	12,692	12,692	12,692
Panel D. Compliance rewards				
Reward versus simplification				
Effect of compliance rewards	0.259	−0.040	1.24	−9.48
	(0.821)	(5.23)	(2.86)	(6.21)
Average outcome in comparison group	77.30%	9.75%	€11.65	10.92%
Observations	12,632	12,632	12,632	12,632

(*continued*)

results contribute to a nascent empirical literature examining the real world importance of salience/information costs for taxes and benefits (Chetty, Looney, and Kroft 2009; Finkelstein 2009; Chetty and Saez 2013; Bhargava and Manoli 2015). Although not part of our framework, these findings can be couched in the notion that the complexity of a decision making environment drives status quo bias (Kahneman, Knetsch, and Thaler 1991) or that subjects can only take a small number of tax rules into account (Eliaz and Spiegler 2011). Either interpretation is consistent with the documented responses to simplification and the high degree of persistence in behavior over pretreatment years shown in Table A5 for example.[19]

[19] Boyer, Dwenger, and Rincke (2014) present evidence from a natural field experiment related to the equivalent Catholic Church tax in Bavaria. Their experiment is designed to make salient that the local church tax is legally binding. Their paper and field experiment were developed entirely subsequent to our analysis and the methods they use to measure intrinsic motivation are based on those presented in this paper. They find such manipulations of tax notifications significantly increase compliance among those identified to be extrinsically motivated, and actually reduce compliance of those identified to be intrinsically motivated. The first of these results closely mirrors our finding on tax simplification: some noncompliance is likely driven by misunderstanding of or inattention toward the

Table 2—Treatment Effects on Compliance—Part II (*continued*)

	Baseline evaders (extrinsically motivated)				Baseline donors (intrinsically motivated)			
	Probability of evading (5)	Probability of donating (6)	Payment amount (7)	Probability of payment increase (8)	Probability of evading (9)	Probability of donating (10)	Payment amount (11)	Probability of payment increase (12)
Panel A. Tax simplification								
Simplification versus control								
Effect of tax simplification	−2.66	6.58	43.40	64.82	−5.25	−4.04	−6.65	−37.29
	(0.747)	(22.86)	(10.60)	(13.69)	(19.67)	(6.97)	(4.85)	(19.38)
Average outcome in comparison group	94.98%	1.91%	€3.13	6.12%	17.32%	62.34%	€39.94	15.58%
Observations	4,007	4,007	4,007	4,007	476	476	476	476
Panel B. Misperception								
Zero audit probability versus simplification								
Effect of correcting misperception	1.53	−8.89	−9.83	−11.03	−16.75	1.52	8.79	32.37
	(0.715)	(17.47)	(6.75)	(7.55)	(17.63)	(5.78)	(4.78)	(28.02)
Average outcome in comparison group	92.35%	2.18%	€4.84	10.53%	15.92%	61.63%	€40.16	8.57%
Observations	6,049	6,049	6,049	6,049	723	723	723	723
Panel C. Deterrence								
Positive audit probability versus zero audit probability								
Effect of deterrence	−3.12	36.89	33.67	29.81	− 0.093	7.07	2.10	30.85
	(0.536)	(15.22)	(6.28)	(6.64)	(15.48)	(4.22)	(3.25)	(19.16)
Average outcome in comparison group	93.80%	1.93%	€4.05	9.00%	12.55%	61.72%	€45.08	10.67%
Observations	9,979	9,979	9,979	9,979	1,261	1,261	1,261	1,261
Panel D. Compliance rewards								
Reward versus simplification								
Effect of compliance rewards	1.27	5.24	−5.46	−15.58	− 11.64	2.02	4.87	48.34
	(0.664)	(16.17)	(6.33)	(6.90)	(15.11)	(4.95)	(3.83)	(25.27)
Average outcome in comparison group	92.35%	2.18%	€4.84	10.53%	15.92%	61.63%	€40.16	8.57%
Observations	9,909	9,909	9,909	9,909	1,247	1,247	1,247	1,247

Notes: Robust standard errors are in parentheses. Estimations account for pretreatment payment behavior, dummies for randomization strata, and parish fixed effects. We split the sample into evaders, donors, and compliers (results not shown) based on pretreatment compliance.

Finally, we use the tax simplification treatment to probe further our finding of excess bunching at exact compliance in the zero deterrence baseline. As stated earlier, it is in general difficult to distinguish between duty-to-comply and focal point/attention explanations for such excess bunching. One way to make headway on the distinction is to exploit our simplification treatment, which makes the point of exact compliance more salient. If attention is the main reason for bunching at exact compliance, one would expect bunching to increase in response to the simplification treatment. In Figure 1, panel B we therefore show the difference in the densities of $T(z) - T(\overline{z})$ between the tax simplification and control groups. The graph clearly shows that the simplification treatment does not increase bunching at exact

local church tax. The second result links to our later study of compliance rewards, that highlight intrinsic motivations can be impacted by how the tax institution is viewed.

compliance, pushing further the interpretation toward such bunching being driven by duty-to-comply preferences.[20]

Misperception.—Our ability to measure intrinsic motivations at baseline hinges on taxpayers being aware that there is zero deterrence. We now test this assertion using the T3 misperception treatment where we make explicit that $p = 0$. On all other dimensions this treatment is identical to the T2 tax simplification letter, which is therefore the natural comparison group.

Panel B of Table 2 shows the results. Columns 1–4 show that averaging across all taxpayers, there are no significant effects of trying to correct for misperceptions on either the extensive or total response margins. However, breaking down the impacts across taxpayer types, the remaining columns show that correcting misperception does have a small but statistically significant effect on the behavior of baseline evaders: they become significantly more likely to evade when they are explicitly told there is zero deterrence, and their tax payments fall (although this effect is not statistically significant).[21]

These findings confirm that compliance in the zero deterrence baseline is virtually unaffected by misperception and is therefore largely intrinsically motivated. That there is little misperception at baseline is not very surprising: the complete absence of enforcement in this established tax system is unlikely to go unnoticed, especially since this has been the status quo for a long time. Of course, while these findings help rule out misperception as a confounder in our setting, they do not imply that misperception is a nontrivial issue in other enforcement settings. In systems with nonzero deterrence, given that deterrence strategies are typically confidential, there remains scope for misperception among taxpayers (Scholz and Pinney 1995; Chetty 2009; Del Carpio 2014).

Deterrence.—In the standard Allingham-Sandmo framework, tax compliance is driven by extrinsic incentives due to audit probabilities (p) and penalties (θ). Panel C of Table 2 documents the compliance impact of higher audit probabilities by pooling together the treatments that inject strictly positive audit probabilities $p = 0.1, 0.2, 0.5$ (treatments T4–T6) into the zero enforcement baseline. To make the variation completely unambiguous and increase power, we compare all these positive p-treatments to the T3 misperception treatment in which $p = 0$. This eliminates noise from idiosyncratic variation in perception.

Considering first the full sample of taxpayers, columns 1–4 show that increased deterrence causes significant reductions in the probability of evasion, increases in the probability of donating, and increased tax payments. Considering heterogeneous treatment responses in the remaining columns, we see that the deterrence effects on both margins are nearly entirely driven by their impacts on baseline evaders (the extrinsically motivated). These results are largely consistent with our

[20] While duty motives have been much discussed in the literature (Scholz and Pinney 1995; Andreoni, Erard, and Feinstein 1998), we are among the first to provide nonparametric evidence of such effects.

[21] These responses among baseline evaders underpin the credibility of our experimental design: the fact that they are willing to evade more when told that the tax system is not enforced suggests that the notification letters were viewed as authentic by those taxpayers.

conceptual model, which predicts positive deterrence effects on the extrinsically motivated and zero deterrence effects on the intrinsically motivated for whom enforcement is not a binding constraint (see Propositions 1 and 2 in the online Appendix A.1).

Two further points are of note. First, the magnitude of each impact is quantitatively similar to those documented in panel A on Tax Simplification. Second, the weak response to these deterrence treatments among the intrinsically motivated speak to the literature examining the potential crowd-out of intrinsic motivations from the provision of extrinsic incentives. For example, if intrinsically motivated taxpayers believe that under $p > 0$, other individuals pay taxes only because of deterrence, this could erode their own intrinsic or social motivation to comply by changing perceptions about other taxpayers' true motives. Our results suggest that no such extrinsic-intrinsic crowd-out exists in this setting; if anything we observe a slight crowd-in of intrinsic motivations.

In Table 3, panel A we break down the pooled impact into the separate impacts of each of the uniform audit probability treatments (again, we split the table into Part I for the full sample and Part II for baseline evaders and baseline donors). This reveals the additional insight that the deterrence effects are quite similar across treatments T4–T6. This lack of gradient could be an artifact of how individuals perceive audit-threat letters like T4–T6: they may respond to the general message of stronger deterrence rather than the specific probability provided. Audit probabilities communicated through such letters are likely to be perceived differently than audit probabilities inferred from actual audit experiences over time. This is of course a generic issue for all tax enforcement experiments, not just ours. We next analyze a different kind of audit-threat letter than what has been considered in the previous literature—namely the audit notch treatment T7—which works very powerfully and suggests that there *is* a gradient.

Notched Audit Probabilities.—In the notched audit probability treatment T7, the tax notification letter announces $p = 0.5$ for payments less than or equal to €10 and $p = 0$ for payments above €10. Such a notch provides a strong incentive for individuals who would otherwise pay less than or equal to €10 to pay just above €10, thereby creating a hole in the payment distribution below the cutoff and excess bunching in the payment distribution just above the cutoff. The theory of notches and how to use them to estimate behavioral responses has been developed by Kleven and Waseem (2013). Here we build on their methodology by taking advantage of the fact that the notch is randomized.

The top panels of Figure 2 illustrate conceptually how individuals should respond to notches by comparing (hypothetical) density distributions of payments for individuals in the audit notch treatment group (solid red line in panel A) and the control group (dashed black line in panel A). The density for the audit notch group features missing mass at and below the cutoff along with excess bunching just above, whereas the density for the control group is smooth around the cutoff as they do not face the notch. Panel B shows the difference in densities between the treatment and control groups: this difference will be zero above the bunch due to random assignment.

Table 3—Individual Treatment Effects on Compliance—Part I

	Full sample			
	Probability of evading (1)	Probability of donating (2)	Payment amount (3)	Probability of payment increase (4)
Panel A. Deterrence				
Positive audit probability versus zero audit probability				
Deterrence, pooled effect	−2.45	−0.438	9.73	33.61
	(0.971)	(6.90)	(3.73)	(10.25)
Deterrence, individual effects				
Audit probability = 0.1	−3.29	5.38	9.52	29.76
	(0.898)	(6.08)	(3.20)	(8.05)
Audit probability = 0.2	−3.11	17.61	11.48	26.81
	(0.923)	(6.44)	(3.37)	(8.11)
Audit probability = 0.5	−2.99	18.27	10.38	24.17
	(0.912)	(6.31)	(3.30)	(8.01)
Average outcome in comparison group	78.04%	8.93%	€11.63	9.42%
Observations	12,692	12,692	12,692	12,692
Panel B. Compliance rewards				
Reward versus simplification				
Compliance rewards, pooled effect	0.259	−0.040	1.24	−9.48
	(0.821)	(5.23)	(2.86)	(6.21)
Compliance rewards, individual effects				
Social reward	0.185	2.97	0.245	−11.60
	(1.03)	(6.68)	(3.51)	(7.71)
Small private reward	0.450	−4.59	−1.15	−10.88
	(1.03)	(6.74)	(3.56)	(7.74)
Large private reward	1.02	−3.30	2.12	−15.30
	(1.00)	(6.60)	(3.98)	(7.63)
Social and private reward combined	−0.618	4.75	3.74	−0.15
	(1.04)	(6.57)	(3.73)	(7.89)
Average outcome in comparison group	77.30%	9.75%	€11.65	10.92%
Observations	12,632	12,632	12,632	12,632

(*continued*)

The bottom panels of Figure 2 show empirical density differences between the audit notch treatment group and different comparison groups. The comparison group in panel C is the T2 tax simplification treatment, while the comparison group in panel D is the T3 misperception treatment. As the raw distributions are lumpy because most individuals pay in one of the statutory tax bins (0, 5, 10, 25, 45, 70, 100), we show the distributions in €5 bins, and average densities within statutory tax bins. The qualitative findings are similar for the two comparison groups and consistent with the conceptual model: there is a large hole in the bins below €10 and large excess bunching just above €10. The amount of excess bunching between €10–€25 (scaled by the average density in the comparison group below the notch) is shown by the estimate b, with bootstrapped standard errors as in Chetty et al. (2011) and Kleven and Waseem (2013). When comparing to the tax simplification treatment in panel C, we have $b = 0.42$: the excess mass above the notch is 42 percent of the average density in the comparison group below the notch. When

TABLE 3—INDIVIDUAL TREATMENT EFFECTS ON COMPLIANCE—PART II (*continued*)

	Baseline evaders (extrinsically motivated)				Baseline donors (intrinsically motivated)			
	Probability of evading (5)	Probability of donating (6)	Payment amount (7)	Probability of payment increase (8)	Probability of evading (9)	Probability of donating (10)	Payment amount (11)	Probability of payment increase (12)
Panel A. Deterrence								
Positive audit probability versus zero audit probability								
Deterrence, pooled effect	−2.66	6.58	43.40	64.82	−5.25	−4.04	−6.65	−37.29
	(0.747)	(22.86)	(10.60)	(13.69)	(19.67)	(6.97)	(4.85)	(19.38)
Deterrence, individual effects								
Audit probability = 0.1	−3.09	14.43	31.69	34.91	15.80	−2.07	2.91	41.68
	(0.741)	(19.43)	(8.73)	(9.19)	(21.44)	(5.76)	(4.38)	(26.31)
Audit probability = 0.2	−3.60	44.22	42.19	29.86	7.45	10.92	−0.544	22.62
	(0.773)	(21.67)	(8.89)	(9.17)	(19.63)	(5.39)	(3.94)	(23.67)
Audit probability = 0.5	−2.69	52.86	27.48	24.55	−25.90	12.08	4.41	29.25
	(0.749)	(22.28)	(9.23)	(9.10)	(21.44)	(5.66)	(4.11)	(26.60)
Average outcome in comparison group	93.80%	1.93%	€4.05	9.00%	12.55%	61.72%	€45.08	10.67%
Observations	9,979	9,979	9,979	9,979	1,261	1,261	1,261	1,261
Panel B. Compliance rewards								
Reward versus simplification								
Compliance rewards, pooled effect	1.27	5.24	−5.46	−15.58	−11.64	2.02	4.87	48.34
	(0.664)	(16.17)	(6.33)	(6.90)	(15.11)	(4.95)	(3.83)	(25.27)
Compliance rewards, individual effects								
Social reward	1.02	17.93	−6.38	−16.84	−11.96	3.17	3.50	40.87
	(0.824)	(21.07)	(7.99)	(8.57)	(19.62)	(6.35)	(4.66)	(34.04)
Small private reward	1.22	2.66	−10.10	−17.50	−11.95	−4.56	5.15	56.00
	(0.825)	(20.60)	(7.87)	(8.50)	(18.48)	(6.32)	(4.57)	(32.76)
Large private reward	2.09	−7.38	−10.57	−21.24	−4.55	2.72	3.16	35.00
	(0.794)	(19.69)	(7.59)	(8.34)	(19.25)	(6.32)	(5.08)	(32.38)
Social and private reward combined	0.777	7.66	4.93	−6.95	−18.58	7.38	7.89	62.16
	(0.841)	(20.47)	(8.34)	(8.75)	(20.02)	(6.37)	(5.24)	(34.60)
Average outcome in comparison group	92.35%	2.18%	€4.84	10.53%	15.92%	61.63%	€40.16	8.57%
Observations	9,909	9,909	9,909	9,909	1,247	1,247	1,247	1,247

Notes: Robust standard errors are in parentheses. See note in Table 2 for details on control variables and definition of types (baseline evaders versus baseline donors).

comparing to the zero audit probability treatment in panel D, the effects are even stronger: the excess mass above the notch is 62 percent of the average density in the comparison group. These bunching estimates are highly significant, much more so than the uniform audit probability treatments considered above (and in the previous literature). That is, randomizing a notched audit probability vastly increases power compared to conventional randomizations of uniform audit probabilities.

Online Appendix Table A7 digs deeper by comparing both the notched audit probability treatment (with $p = 0.5$ below a cutoff) and the T6 uniform audit probability treatment (with $p = 0.5$ everywhere) to the T3 misperception treatment (with $p = 0$). To begin with, Column 1a considers the *total* average treatment effect of the notched and uniform audit probabilities. The effects are roughly similar in size (slightly larger for the notch) and highly significant for both treatments. However, the audit notch estimate obtained this way is attenuated because it does not account for the fact that individuals initially above the cutoff (where p remains zero) are untreated. Hence, column 1b uses the bunching estimate in Figure 2D to obtain the

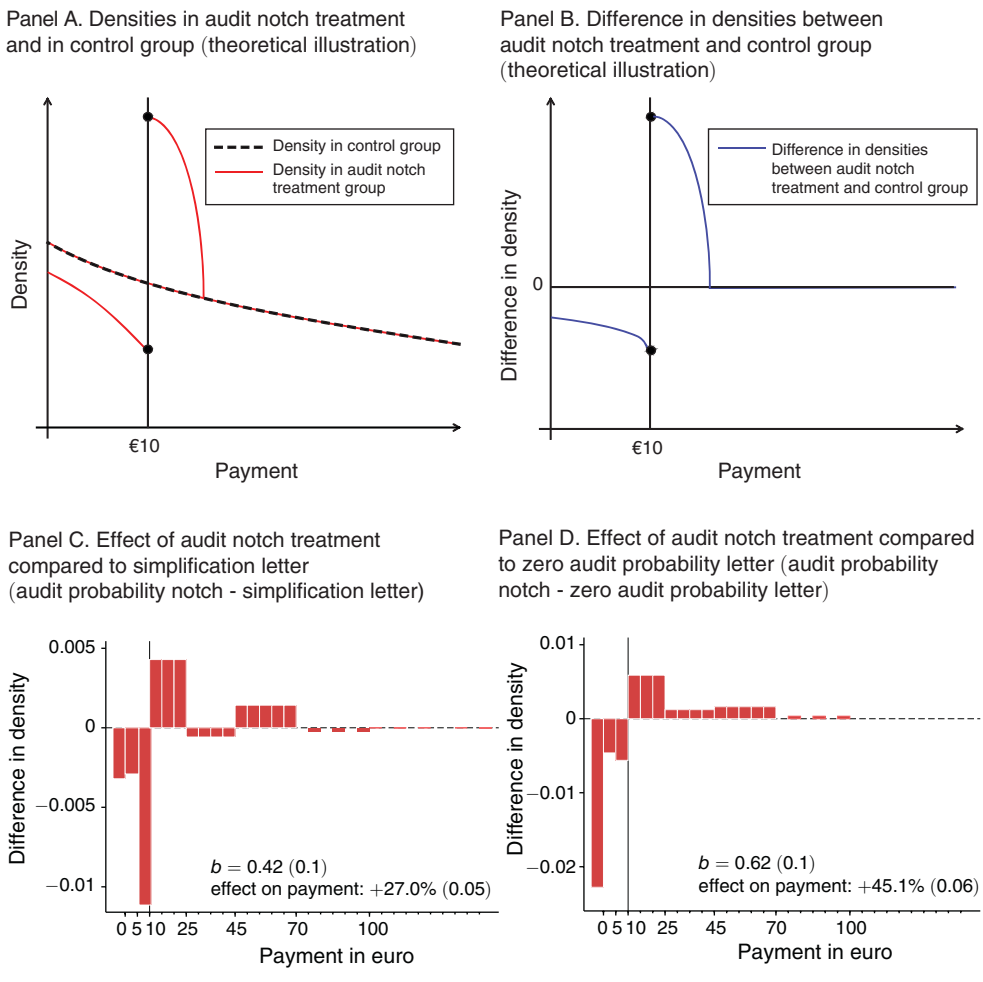

FIGURE 2. EFFECT OF AUDIT PROBABILITY NOTCH ON COMPLIANCE

Notes: Panel A provides a theoretical illustration of the distribution of payments made for the audit probability notch treatment (compared to the distribution of payments in the control group). Similarly, panel B illustrates the difference in densities between the audit probability notch treatment and the control group. Panels C and D display the difference in the empirical density distributions of payments made. The density distribution of the audit probability notch letter group is compared to the density distribution of the simplification letter group in panel C and to the density distribution of the zero audit probability letter group in panel D. In both lower panels, the dashed horizontal line denotes zero difference in density distributions between the compared letter groups. The vertical line denotes the threshold at which the audit probability dips from 50 percent (payments below) to 0 percent (payments above). Bunching b is the excess mass just above the threshold (scaled by the average counterfactual density below the notch). In both panels, the sample consists of baseline evaders, who paid less than the amount owed prior to treatment (baseline year 2011). The sample is limited to those with payments weakly smaller than 150 euro. The bin size is 5 euro. We account for differences in the size of tax brackets below and above the threshold by averaging densities within tax brackets.

correct *local* average treatment effect on tax payments. The estimated audit notch impact of 45 percent constitutes the correct comparison with the uniform audit probability impact of 29 percent, and so the notched audit-threat letter induces a much stronger response than the uniform audit-threat letter.

Compliance Rewards.—We complete our analysis by using the rewards/recognition treatments to probe motivations for tax compliance. We first pool together all such rewards treatments (T8–T11), and then later consider the individual impacts of each type of reward (social recognition, monetary prize draws, and a combination of the two). The comparison group is the T2 tax simplification treatment. As discussed earlier, the probability of actually winning each reward is very close to zero. We therefore view the salient feature of these rewards as being what they signal about the institution of the local church tax system. In particular, the offer of rewards for compliance (in contrast to punishment for noncompliance) highlights the voluntary aspect of an unenforced tax system, which may have very different effects across different compliance types.

Panel D of Table 2 presents our findings. For the full sample, we show in columns 1–4 that the offer of compliance rewards has no significant impact on either the extensive or total response margins of tax compliance. However, the remaining columns show that pooling taxpayers masks the considerable heterogeneity in compliance responses to rewards across taxpayer types. Among baseline evaders (the extrinsically motivated) the offer of rewards/recognition for compliance causes them to: (i) significantly increase their probability of evading by 1.27 percent; (ii) significantly reduces the likelihood they increase payments by 16 percent. Among baseline donors (the intrinsically motivated) the offer of rewards/recognition: (i) does not significantly impact their probability of donating; and (ii) significantly increases the likelihood they increase the size of their donation. This is remarkable given the considerable levels of donation/overpayment among this type of taxpayer at baseline.[22]

Two further points are of note. First, the sharply heterogeneous effects of rewards across taxpayer types again highlight the importance of being able to cleanly classify individuals as extrinsically or intrinsically motivated for the study of tax compliance. Pooling all taxpayers leads to the (incomplete) conclusion that the provision of rewards does not impact tax compliance. Second, by highlighting the voluntary aspect of an unenforced tax system, the reward treatments induce qualitatively similar responses among baseline evaders as the misperception treatment that made explicit $p = 0$ and, thus, also emphasized that tax payments are effectively voluntary. Baseline donors, on the other hand, respond as if these rewards positively shock their warm glow and thus crowds-in their intrinsic motivations.

In Table 3, we report the separate impacts of each form of reward. Recall that these rewards are of three types: (i) T8: provides individuals with a *purely social* reward through the possibility of their name being publicly announced in a local newspaper; (ii) T9–T10: provide individuals a *purely private* reward through their entry into small/high valued monetary prize draws; and (iii) T11: combined social and private rewards so taxpayers have the opportunity to be recognized in a local newspaper *and* be entered in the high-valued monetary prize draw. Hence, the differences between these treatments are whether the reward takes the form of social

[22] The spirit of these results match findings from other contexts in which very low-value rewards motivate prosocial behavior. For example, Stutzer, Goette, and Zehnder (2011) find that offering lottery tickets increases blood donations; Chetty, Saez, and Sándor (2014) find that offering a $100 gift card to journal referees significantly reduces the time taken to send reports.

or private recognition and the value of the private reward. All other dimensions are held constant across rewards: the number of individuals named in the social recognition component of T8 and T11 remains the same; and in T9–T11 the identity of monetary prize winners and their prize value remains private information.

Panel B of Table 3 documents a very uniform pattern of impacts across the different forms of reward, although we sometimes lose statistical significance when focusing on individual reward treatments. Across all three samples, the sign of the treatment effect is almost always the same for each of the individual rewards and for the pooled effect. For example, when considering the probability of increasing tax payments as our outcome, all four reward letters have a negative effect on the extrinsically motivated and a positive effect on the intrinsically motivated.

Our findings thus highlight that the offer of rewards can significantly impact tax compliance: the first-order impact will depend on taxpayers' underlying motivations; the form in which rewards are offered are less consequential in our setting.[23] The heterogeneous treatment responses across taxpayer types reveals a subtle trade-off for a social planner. The net benefit of offering such rewards depends both on the magnitude of responses for extrinsically and intrinsically motivated taxpayers, and on the underlying distribution of those types in the population. This is a timely insight given the use of rewards or recognition for tax compliance is becoming more prevalent, especially in developing countries. Moreover, over half of US states have utilized "name and shame" programs revealing top debtors (Luttmer and Singhal 2014).

These results also shed more light on the potential crowd-out/in between extrinsic and intrinsic motivations. We earlier documented that the manipulation of deterrence parameters (extrinsic incentives) had little impact on the intrinsically motivated, consistent with the absence of strong cross-effects between forms of motivation. Our results on the provision of compliance rewards are consistent with this insight: the qualitative similarity of responses to social and monetary rewards, as well as their interaction, suggests that intrinsically motivated tax compliance is not crowded-out by the provision of monetary rewards.[24]

V. Conclusion

This paper contributes to the large literature on tax compliance, and specifically to an emerging literature on intrinsic motivations for compliance (Luttmer and Singhal 2014). We provide novel insights on the relative importance of extrinsic and intrinsic motivations for tax compliance in a large representative sample of German taxpayers. We shed light on each motivation and their interaction using

[23] As such, there is little value added in discussing further the interpretation of the different types of reward. It remains an open question for future research whether social and private recognition can have different effects on tax compliance in other settings, say because *social* rewards leverage against intrinsically motivated individuals contributing to the tax because they have social image concerns or a desire to signal to others their type or conspicuous generosity (Bénabou and Tirole 2003, 2006; Ellingsen and Johannesson 2011).

[24] Gneezy, Meier, and Rey-Biel (2011) review the field evidence on extrinsic-intrinsic crowd-out. Studies that find no such cross-effects (in a variety of non-tax contexts) include Dal Bó, Finan, and Rossi (2013); Ashraf, Bandiera, and Jack (2014); and Chetty, Saez, and Sándor (2014).

experimental manipulations of deterrence, tax simplification, misperception, and rewards/recognition.

We make headway on these questions by exploiting unique aspects of our data and setting. Our data allows us to precisely measure tax compliance in contrast to many earlier studies (Slemrod and Weber 2012), and to cleanly identify extrinsically and intrinsically motivated taxpayers based on their pretreatment compliance behavior in a zero deterrence baseline. Furthermore, the tax system studied is one in which overpayments are encouraged, thus creating the coexistence of evaders and donors and allowing us to integrate the study of tax compliance with the study of charitable giving. While these topics have largely been studied separately, they naturally belong together as any imperfectly enforced tax system involves an element of voluntary giving.

We conclude by highlighting two directions for future research. First, our finding that 20 percent of individuals pay at least true taxes owed in a baseline with no pecuniary incentive to comply suggests a need for more research that identify the key intrinsic or social motivations to comply and study how these respond to policy.[25] We have provided evidence that duty-to-comply motives may be one important mechanism in the context of taxation, but we have also shown that other forms of intrinsic motives play a role—and are affected by policies that provide recognition—for a subset of taxpayers who are willing to pay taxes above and beyond the letter of the law.

Second, while we find significant effects of deterrence and reward incentives on compliance behavior, the effect of these marginal incentives are relatively modest compared to the baseline evasion rate of 80 percent. When pooling the effect of all of our incentive treatments, we find that collectively they reduce the aggregate evasion rate by only about 4 pp. In contrast, previous work has shown that third-party information reporting and tax withholding is able to reduce evasion to almost zero (Kleven et al. 2011). Hence, while incentives on the margin do matter, this paper along with the recent literature show that it is not possible to make a tax system fully successful without information and tax collection systems that make compliance more or less automatic. The next generation of compliance studies should therefore provide more direct comparisons between the impact of marginal incentives—be they economic or social in nature—and the impact of mechanisms related to informational and administrative procedures. The longer-term aim would be to unify separate strands of the recent economics literature, which have identified the importance of institutional/administrative features for individual behavior in contexts as diverse as pro-social behavior, benefits take-up, savings, and voting.

[25] It is instructive to compare the levels of intrinsic motivation we document to those in DellaVigna, List, and Malmendier (2012). They combine a natural field experiment and a structural model to estimate the share of potential donors to a charitable cause that is intrinsically motivated. Despite their very different setting, they report a quantitatively similar share of individuals who are intrinsically motivated to give (25 percent) as we find in our zero deterrence baseline.

Appendix

Table A1—Summary of Treatments

Treatment	Description	Wording of additional paragraph relative to tax simplification treatment (T2)
Control (T1)	Tax notice unchanged	
Treatment group 1: Tax simplification and misperception		
Tax simplification (T2)	Tax notice simplified	
Misperception (T3)	Zero audit probability	Please note that, according to Article 9 para. 4 of the Church Levy Collection Act, the Evangelical-Lutheran congregation can delegate the collection of the local church tax to the church tax authority. The church tax authority can officially assess your income. However, the Evangelical-Lutheran congregation does not make use of this option. There is no verification of church members' own income assessment.
Treatment group 2: Deterrence		
Audit probability 10% (T4)	Audit probability of 10%	Please note (. . .) your income. In order to ensure a fair tax collection, we consider it necessary to verify the church members' self-assessment for every tenth [fifth, second] church member. In other words, the self-assessment of 10% [20%, 50%] of church members will be verified.
Audit probability 20% (T5)	Audit probability of 20%	
Audit probability 50% (T6)	Audit probability of 50%	
Audit probability notch (T7)	Audit probability of 50% for payments $\leq$ €10 and zero audit probability for payments $>$ €10	Please note (. . .) your income. While there will be no verification of church members' self-assessment for payments above €10, there may be a verification of payments at €10 or lower. In order to ensure a fair tax collection, we consider it necessary to verify the church members' self-assessment for every second church member paying €10 or less. In other words, the self-assessment of 50% of church members paying €10 or less will be verified.
Treatment group 3: Compliance rewards		
Social reward (T8)	Lottery with individual's timely compliance being announced in local newspapers	Among all individuals paying a local church tax of at least €5 no later than September 30, 2012, we will randomly draw 100 church members. If you belong to the church members drawn by lot we will contact you and ask for consent before publishing your name in a newspaper advertisement. With this advertisement, published in the [names of three local newspapers], we are going to thank the allotted church members by name for funding our work. Funds for financing the advertisement have been kindly found to this end.
Small private reward (T9)	Lottery, prize draw €250	All individuals paying a local church tax of at least €5 no later than September 30, 2012 are going to take part in a lottery. From every 1,000 local church taxpayers one will be drawn to win a prize of €250 [€1,000]. The prize has been kindly funded to this end.
Large private reward (T10)	Lottery, prize draw €1,000	
Social and private reward combined (T11)	Lottery combining social reward and large private reward	Among all individuals (. . .) for funding our work. In addition, out of the 100 church members mentioned above, we will randomly draw 15 members who will each win a prize of €1,000. Funds for the advertisement and the prizes have been kindly found to this end.

Notes: Treatments T3 to T11 are based on the simplified notice T2 and add one paragraph on the cover page relative to T2. We implemented two additional treatments on social norms and moral suasion (T12 and T13). See the online Appendix for further details.

REFERENCES

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

Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy* 97 (6): 1447–58.

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *Economic Journal* 100 (401): 464–77.

Andreoni, James, Brian Erard, and Jonathan Feinstein. 1998. "Tax Compliance." *Journal of Economic Literature* 36 (2): 818–60.

Ashraf, Nava, Oriana Bandiera, and B. Kelsey Jack. 2014. "No margin, no mission? A field experiment on incentives for public service delivery." *Journal of Public Economics* 120: 1–17.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169–217.

Bénabou, Roland, and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70 (3): 489–520.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96 (5): 1652–78.

Bhargava, Saurabh, and Day Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review* 105 (11): 3489–3529.

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

Boyer, Pierre C., Nadja Dwenger, and Johannes Rincke. 2014. "Do Taxes Crowd Out Intrinsic Motivation? Field-Experimental Evidence from Germany." Unpublished.

Cagala, Tobias, Ulrich Glogowsky, and Johannes Rincke. 2015. "Who to Target in Fundraising? A Field Experiment on Gift Exchange." Unpublished.

Chetty, Raj. 2009. "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance." *American Economic Journal: Economic Policy* 1 (2): 31–52.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics* 126 (2): 749–804.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review* 99 (4): 1145–77.

Chetty, Raj, and Emmanuel Saez. 2013. "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients." *American Economic Journal: Applied Economics* 5 (1): 1–31.

Chetty, Raj, Emmanuel Saez, and László Sándor. 2014. "What Policies Increase Prosocial Behavior? An Experiment with Referees at the Journal of Public Economics." *Journal of Economic Perspectives* 28 (3): 169–88.

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics* 128 (3): 1169–1218.

Del Carpio, Lucia. 2014. "Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru." Unpublished.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier. 2012. "Testing for Altruism and Social Pressure in Charitable Giving." *Quarterly Journal of Economics* 127 (1): 1–56.

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 2016. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany: Dataset." *American Economic Journal: Economic Policy*. http://dx.doi.org/10.1257/pol20150083.

Eliaz, Kfir, and Ran Spiegler. 2011. "Consideration Sets and Competitive Marketing." *Review of Economic Studies* 78 (1): 235–62.

Ellingsen, Tore, and Magnus Johannesson. 2011. "Conspicuous Generosity." *Journal of Public Economics* 95 (9): 1131–43.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler. 2013. "Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information." *Journal of the European Economic Association* 11 (3): 634–60.

Finkelstein, Amy. 2009. "E-Z Tax: Tax Salience and Tax Rates." *Quarterly Journal of Economics* 124 (3): 969–1010.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives* 25 (4): 191–209.

Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev. 2014. "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." National Bureau of Economic Research (NBER) Working Paper 20007.

Huck, Steffen, Imran Rasul, and Andrew Shephard. 2015. "Comparing Charitable Fundraising Schemes: Evidence from a Natural Field Experiment and a Structural Model." *American Economic Journal: Economic Policy* 7 (2): 326–69.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives* 5 (1): 193–206.

Karlan, Dean S., and John A. List. 2007. "Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment." *American Economic Review* 97 (5): 1774–93.

Kleven, Henrik Jacobsen. 2014. "How Can Scandinavians Tax So Much?" *Journal of Economic Perspectives* 28 (4): 77–98.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søeren Pedersen, and Emmanuel Saez. 2011. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." *Econometrica* 79 (3): 651–92.

Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez. 2009. "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries." National Bureau of Economic Research (NBER) Working Paper 15218.

Kleven, Henrik J., and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *Quarterly Journal of Economics* 128 (4): 669–723.

Luttmer, Erzo F. P., and Monica Singhal. 2014. "Tax Morale." *Journal of Economic Perspectives* 28 (4): 149–68.

Pomeranz, Dina. 2015. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax." *American Economic Review* 105 (8): 2539–69.

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3): 180–212.

Sandmo, Agnar. 2005. "The Theory of Tax Evasion: A Retrospective View." *National Tax Journal* 58 (4): 643–63.

Scholz, John T., and Neil Pinney. 1995. "Duty, Fear, and Tax Compliance: The Heuristic Basis of Citizenship Behavior." *American Journal of Political Science* 39 (2): 490–512.

Slemrod, Joel. 2007. "Cheating Ourselves: The Economics of Tax Evasion." *Journal of Economic Perspectives* 21 (1): 25–48.

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

Slemrod, Joel, and Caroline Weber. 2012. "Evidence of the invisible: Toward a credibility revolution in the empirical analysis of tax evasion and the informal economy." *International Tax and Public Finance* 19 (1): 25–53.

Slemrod, Joel, and Shlomo Yitzhaki. 2002. "Tax Avoidance, Evasion, and Administration." In *Handbook of Public Economics*, Vol. 3, edited by Alan J. Auerbach and Martin Feldstein, 1423–70. Amsterdam: North-Holland.

Stutzer, Alois, Lorenz Goette, and Michael Zehnder. 2011. "Active Decisions and Prosocial Behaviour: A Field Experiment on Blood Donation." *Economic Journal* 121 (556): F476–93.