

// NO.21-040 | 05/2021

# DISCUSSION PAPER

// SEBASTIAN BLESSE

## Are Your Tax Problems an Opportunity Not to Pay Taxes? Evidence From a Randomized Survey Experiment

# Are Your Tax Problems an Opportunity Not to Pay Taxes? Evidence from a Randomized Survey Experiment \*

SEBASTIAN BLESSE

April 6, 2021

## Abstract

Taxpayers often view tax rules and filing processes as complicated. In this paper I study whether the perceived tax uncertainty among peers leads to a reduction of voluntary tax compliance. I find strong supportive evidence for this hypothesis using a survey experiment for a large representative sample of the German population. Providing randomized information that others are uncertain about how to file their taxable income decreases individual tax morale. This suggests that subjects use negative peer signals as an excuse in order to opt-out of tax compliance. Studying related heterogeneous treatment effects, I find that both older and left-wing subjects are more responsive to tax uncertainty of others. I also show persistent treatment effects among very honest taxpayers in a follow-up survey.

**JEL Classification:** H26, Z13, K42, C9

**Keywords:** Tax Complexity, Taxpayer Uncertainty, Tax Morale, Survey Experiments

---

\***Blesse:** ZEW Mannheim and Goethe University Frankfurt. Email: sebastian.blesse@zew.de. I gratefully acknowledge helpful comments by Alfons Weichenrieder, Friedrich Heinemann, Annika Havlik, Philipp Lergetporer, Justus Nover, Florian Buhlmann and Philipp Dörrenberg. Anna Cresswell and Nicolas Eveneshen provided excellent research assistance. I am thankful for funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project ID 139943784—SFB 884. This paper uses data from the German Internet Panel wave 40 and 41. A study description can be found in Blom et al. (2015). The German Internet Panel (GIP) is funded by the German Research Foundation (DFG) through the Collaborative Research Center 884 "Political Economy of Reforms" (SFB 884).

# 1 Introduction

Many taxpayers regularly face the task of filing their annual tax declaration, which requires processing large amounts of information and abiding ample documentation requirements. This is all the more cumbersome for taxpayers since modern tax systems are typically complex and involve many tax rules, which are often updated on a yearly basis. Therefore, it may not come as a surprise that existing work finds that people frequently have a poor knowledge on existing tax rules and their implications (Stantcheva, 2020).<sup>1</sup> According to a recent representative survey, people perceive tax filing to be very or rather difficult and about 90% support a move towards simpler tax rules (Blesse et al., 2021). Complex tax schedules and filing processes also have important negative implications such as low-take of rates of tax benefits (e.g. Chetty et al. (2013) and Bhargava and Manoli (2015)), large compliance costs on taxpayers (Benzarti, 2020), inattention to new tax incentives (Abeler and Jaeger, 2015) and the fact that taxpayers are confused and misinterpret economic incentives of taxes (Feldman et al., 2016; Aghion et al., 2017).<sup>2</sup>

Against this background, complex tax rules may reduce compliance and enforcement levels (Krause, 2000). Notably, Bellemare et al. (2019) show lab evidence that people use the ambiguity from unintentional filing mistakes in complex tax settings as an excuse for non-compliance. This shows that complex tax rules have a negative direct effect on tax compliance by creating tax payer confusion and opening up wiggle room for individuals to opt-out of compliance. However, even though people also make individual tax compliance decisions in a social context (e.g. Traxler (2010) and Frey and Torgler (2007)), it remains unclear whether there may be indirect negative effects (or spillovers) of observing other taxpayers to be confused. Specifically, taxpayer may use signals about the filing uncertainty of others as an excuse for not paying taxes themselves.<sup>3</sup> In a world where it is the norm rather than the exception to be uncertain about how to file one's taxes, these negative peer signals may decrease the morale costs of tax non-compliance.

This paper addresses this question and provides causal evidence on spillover effects of peer information about tax confusion of others on own tax compliance. Specifically,

---

<sup>1</sup>People seem to systematically underestimate average tax rates (De Bartolome, 1995; Gideon, 2017; Ballard and Gupta, 2018). This may be due to confusion about the concepts of average and marginal tax rates as well as non-linearities in tax schedules. Rees-Jones and Taubinsky (2020) provide experimental evidence that people regularly engage in heuristics to approximate non-linear tax schedules (i.e. "Schmeduling", as discussed by Liebman and Zeckhauser (2004)). Stantcheva (2020) and Aghion et al. (2017) show that people do not correctly understand income tax schedules. Hoopes et al. (2015) report that taxpayers are often not fully informed about taxes but search for relevant information when the tax becomes more salient. Chetty et al. (2009) highlight the role of tax salience for economic behavior.

<sup>2</sup>For a recent review of the economic implications of tax complexity, please see Blesse et al. (2021).

<sup>3</sup>On a related note, Bursztyn et al. (2020) find evidence that people use justifications for adopting socially stigmatized behavior and judge others who use such excuses for stigmatized behavior less harshly. In my specific setting, people may use tax difficulties of other taxpayers as a justification for decreasing pro-social behavior of compliance and to engage in social stigmatized behavior of non-compliance. Thus, signals of taxpayer uncertainty from others should decrease the motivation to pay taxes by oneself.

I show whether information about the perceived uncertainty of others how to file taxes reduces individual tax morale.<sup>4</sup> In order to do this, I designed a survey experiment which asked participants about their tax morale, i.e. their view on the justifiability of tax evasion, and compared the their answers across three experimental groups which differ in the amount of information about taxpayer uncertainty of others. The analysis is inspired by a recent representative survey among the general population in Germany from Blesse et al. (2021) which finds that half of survey respondents find their tax declaration difficult and especially see the chance of forgetting something or reporting falsely as the source of these difficulties (Blesse et al., 2021, p. 56). This observation may be used as a justification for opting-out of tax compliance and judge tax evasion less harshly. The act of tax evasion or even support for evasion likely comes with a social stigma in the eyes of the general population since survey-data shows consistently high tax morale among the German population (see Section 2.1). Thus, signals of taxpayer uncertainty among others should decrease the motivation to pay taxes and improve views on the justification of tax evasion. To put it differently, becoming aware of people not being able to declare their taxes properly may reduce the moral cost of evasion and therefore, people may be more accepting of non-compliance.

The survey experiment was fielded in the German Internet Panel (GIP) which is a large probability-based sample of the general population in Germany.<sup>5</sup> The survey has 3 experimental groups to study the role of taxpayer uncertainty of others for tax morale.

First, the *Uncertainty* condition gives respondents the information about prevalent difficulties of others when filing their annual tax declaration, which is inspired from a previous survey of the author (see Blesse et al. (2021)). Specifically, the information states that in the context of complex tax laws (due to many deduction and exemption possibilities) many citizens are often not sure about whether they report all incomes correctly in their tax declaration. Respondents in this treatment should be more aware of other taxpayers' lack of tax knowledge and their uncertainty on how to file their taxable income properly. This treatment condition thus allows me to test whether people are more lenient towards tax evasion when becoming aware of other taxpayer's uncertainty about how to file their taxes. Individuals may adopt the view that non-compliance may be due to 'honest' mistakes in light of prevalent tax uncertainties among others.

Second, the *Unequal knowledge* treatment augments the social information from

---

<sup>4</sup>I refer to tax morale as all non-pecuniary motives or one's intrinsic motivation to pay taxes in line with Luttmer and Singhal (2014) and Frey and Torgler (2007). Related contributions include Halla (2012), Besley et al. (2019), Alm and Torgler (2006), Richardson (2006), Frey and Torgler (2007), Torgler (2005), Wenzel (2004), Lewis (1982) and Alm et al. (1992). Notably, Dwenger et al. (2016) find tax compliance even in the absence of enforcement, likely due to duty-to-comply preferences.

<sup>5</sup>Other research papers that use GIP data include, among others, Dörrenberg and Peichl (2018), Kerschbamer and Müller (2020), Engelmann et al. (2018), Blesse et al. (2021), Blesse and Heinemann (2020), Müller and Renes (2020), Gsottbauer et al. (2020) as well as Fehr et al. (2020).

the *Uncertainty* group with the contrasting notion that there are also other citizens who possess the relevant tax knowledge or have access to it through tax advisors and make use of that knowledge in order to reduce their tax burden. Accordingly, this treatment condition represents a framing of the uncertainty treatment insofar as it highlights the fact that heterogeneous groups have different opportunities to optimize their after-tax incomes. The augmented statement captures the existence of distinct coexisting tax-subcultures which differ in their levels of tax compliance (Lewis, 1982; Traxler, 2010). Accordingly, in this condition taxpayers should be perceived to differ in their compliance rates due to different abilities to comply with complex taxes, and to influence their after-tax income through tax knowledge. Third and last, a control group does not receive any information. All three experimental groups, however, read the same neutral opening statement about tax evasion in order to become familiarized with the topic.

The results show that informing respondents about the prevalent uncertainties of others when reporting income for tax purposes significantly decreases elicited tax morale. First, this shows that people are not necessarily aware of uncertainties of other taxpayers in the context of filing tax declarations. This is true despite the fact that 50% of respondents elicited in an earlier wave of the GIP that they find tax declarations somewhat or very difficult (see Blesse et al. (2021)). Second, the significant negative effect of the *Uncertainty* treatment shows that the respondents empathize with tax-filing problems of others, making them more lenient towards tax evasion. The *Unequal knowledge* group also significantly reduces elicited tax morale as compared to the control group but one cannot reject the equality of effects between both treatments.

Moreover, I find that the negative information effect of both treatments results entirely from a lower likelihood that respondents chose the answer category representing the highest tax morale. Thus, information about uncertainties regarding tax filing due to complex tax rules especially reduces the prevalence of very honest taxpayers. Related effects are larger for the *Uncertainty* group. I also analyze whether certain groups of participants respond more or less elastically to either of the treatment interventions with respect to their tax morale. Given the longitudinal structure of the GIP survey, I can exploit rich background information on each respondent from previous survey waves. Older respondents show stronger negative responses than younger subjects to the *Uncertainty* information, i.e. that many people are uncertain how to properly declare their taxable income.<sup>6</sup> Regarding the *Unequal knowledge* condition, I find that the result of a lower tax morale is dominantly driven by political ideology. Specifically, left-wing respondents show significantly lower tax morale when being confronted with the augmented treatment condition, as compared to conservative participants. This corresponds to the results of

---

<sup>6</sup>This corresponds with evidence that higher age groups show higher support for simpler taxes (see Blesse et al. (2021)). Aghion et al. (2017) also report that older individuals know less about tax incentives and are also less able to learn, indicating higher cognitive costs in adjusting to complex tax rules.

Alesina et al. (2018) who find that political preferences play a large role for redistributive preferences. In line with their results, I argue that left-wing participants care more about the uncertainty of ‘other’ taxpayers if this may result in unequal opportunities of taxpayers to optimize their after-tax income through differences in tax knowledge at their disposal. Accordingly, leftists are more lenient towards tax evasion, likely as a means of redistribution towards a disadvantaged group of tax filers who either lack tax knowledge or do not have the means to access it through (presumably costly) advice.

I also test the persistence of my results using a follow-up survey in the GIP two months after the main experiment. To the best of my knowledge, this is the first attempt to study within-person changes in tax morale over time. First, the elicited tax morale appears to be relatively stable for individual respondents in the control group over time. Second, the treatment effects do not appear to be persistent regarding their effect on tax morale on average. Nonetheless, the share of very high tax morale is significantly reduced by both types of information even two months after the intervention in the main survey. Thus, peer information about tax uncertainty of others has persistent effects on very honest levels of tax morale and the resulting changes in judgments of evasion grow less strong and more lenient again over time. Persistent treatment effects, however, also suggest that the estimated effects arise due to new information rather than due to an increased salience of taxpayer confusion. Again, I find that political ideology is an important driver of the negative treatment effect of the *Unequal knowledge* condition. The point estimates are virtually identical with the respective effects in the main survey. It appears that left-wing respondents internalized a more lenient view towards tax evasion when being informed that taxpayer uncertainties are common and when possible negative distributional implications for unaware taxpayers are highlighted to the participants.

This study makes several contributions to the literature. First, I add to the general literature on tax compliance and in particular to work regarding the intrinsic motivation of tax compliance (see Luttmer and Singhal (2014) for a recent review). The treatment conditions both speak to papers addressing peer effects and social influences as a channel of tax morale (for a review of the empirical evidence and theoretical work, see Luttmer and Singhal (2014) and Hashimzade et al. (2013), respectively). Cross-country survey evidence supports the hypothesis of ‘conditional cooperation’ where people decide to comply depending on the (perceived) compliance of others (Frey and Torgler, 2007). Several randomized control trials (RCTs), however, provide a mixed picture whether information provision regarding compliance behavior of others has an effect on individual tax compliance. Some studies find positive effects of peer information on compliance behavior, including Hallsworth et al. (2017) who study timely tax payment in the UK. The authors highlight the value of descriptive over injunctive norms (“what others do” vs. “what others think should be done”). Bott et al. (2020) also find improved compliance of peer information among Norwegian taxpayers. Moreover, Del Carpio (2014) shows

that informing taxpayers about average compliance has long-lasting positive effects but information about enforcement levels does not. Notably, empirical support for the relevance of social interactions in tax compliance also comes from studies showing spillovers in tax compliance among peer networks (Drago et al. (2020) for the case of TV license fees; Paetzold and Winner (2016) for reporting a commuter tax allowance). Other RCT studies, however, do not find peer effects on average, such as Slemrod et al. (2001), Fellner et al. (2013), Castro and Scartascini (2015) or Perez-Truglia and Troiano (2018). Also social image concerns may play a role for compliance decisions. Governments can exploit these concerns in order to increase tax compliance through shaming (Perez-Truglia and Troiano, 2018; Dwenger and Treber, 2018). Rewards for compliance, however, can also backfire depending on the intrinsic motivation of taxpayers (Dwenger et al., 2016).

The present paper contributes to this literature by providing novel evidence on the effect of a yet overlooked negative peer information on voluntary tax compliance, i.e. the perceived taxpayer uncertainty of other taxpayers, using a large-scale survey experiment among a representative population sample.<sup>7</sup> My results suggest that once people become aware of tax confusion among others, they significantly reduce their own tax morale.

Second, my results complement recent studies on the effects of tax complexity on tax compliance. These studies find, for instance, that people underreact to new tax rules in more complex tax settings in the lab (Abeler and Jaeger, 2015).<sup>8</sup> Bellemare et al. (2019) show evidence from a lab experiment that the effects of tax complexity on tax compliance are intrinsically linked to distributive fairness. Complex taxes appear to bear justification potential for selfish behavior resulting from the ambiguity of filing mistakes.<sup>9</sup>

---

<sup>7</sup>There are a few other survey experiments on tax morale. Dörrenberg and Peichl (2018) study the role of social norms and reciprocity and Sjoberg et al. (2019) analyze the effect of anti-corruption efforts as well as the opportunity for citizens to voice their public expenditure preferences to the government. Ortega et al. (2016) estimate the effect of government performance on the tax compliance.

<sup>8</sup>Several RCTs also find that simplified communication of tax authorities improves tax compliance (De Neve et al., 2021; Dwenger et al., 2016; Eerola et al., 2019). De Neve et al. (2021) find persistent but diminishing effects of a simplified letter of the tax authorities on compliance. Moreover, simplification appears more cost-effective than traditional enforcement mechanisms. Dwenger et al. (2016) also find differential effects of simplified tax notifications with higher compliance for evaders and null effects for compliers. Other papers also indicate that a lack of awareness and complex welfare eligibility criteria add to low welfare benefit-take up rates (Kleven and Kopczuk, 2011) and simplified information provision to eligible benefit recipients can reduce this problem (Chetty and Saez, 2013; Bhargava and Manoli, 2015). Complex taxes (or perceptions thereof) also have negative effects on survey-based tax compliance (Richardson, 2006; Frey and Torgler, 2007; Torgler, 2005). Eriksen and Fallan (1996) also show that tax knowledge improves tax morale and fairness perceptions of the tax system.

<sup>9</sup>Related to justification for selfish behavior is the behavioral concept of 'moral wiggle room' (Dana et al., 2007). Evidence of pro-social behavior may often come from the fact that people want to appear fair but are, in fact, more selfish and inherently reluctant to share (e.g. Lazear et al. (2012)). In line with this concept, participants in laboratory experiments opt-out of pro-social behavior when uncertainty in the relation of choices and outcomes arise. I find evidence in a large-scale survey experiment on a representative sample of the general population that the intrinsic motivation for the social norm of tax compliance is not robust with respect to information about uncertainties of others from complex tax systems and that people use this information as an opportunity to reduce their tax morale.

However, none of these papers show whether the perceived tax confusion among others (among others, established in Blesse et al. (2021)) itself has an effect on tax compliance. The present paper shows that information about the prevalence of tax filing uncertainty among others offers a morale justification to decrease one's own level of tax morale.

Complex tax systems may thus not only have direct detrimental effects on tax compliance through taxpayer confusion but may also undermine tax morale through negative signals of peer compliance, i.e. once individuals become aware of other confused taxpayers. I argue that these uncertainties may merely serve as an excuse since there are very few opportunities to evade income taxes in Germany for the majority among the general population in the survey on account of mandatory third-party reporting of dependent incomes (see Kleven et al. (2011)) and the general duty to document most deduction possibilities when filing tax declarations (i.e. although they need to be submitted to the tax authorities only after an explicit request). However, the results show that tax confusion of others may still serve as an excuse for the general population for non-compliance.

Third, my findings relate to papers showing that social norms (here, tax morale) are themselves endogenous (e.g. for evidence Bursztyn et al. (2020); for theoretical work, see Benabou and Tirole (2011), Besley et al. (2019) and Acemoglu and Jackson (2017)). I find that information about implications of complex taxes decreases tax morale, implying that the design of tax systems can affect the compliance with the tax system (as hypothesized by Luttmer and Singhal (2014), p. 165). Besley (2020b) also shows that norms like tax morale in theory evolve dynamically over time as part of a social contract between citizens and the state. The social contract entails that if the government fails to sufficiently provide the public good, people then find it legitimate to withdraw their cooperation with the state. The results of my paper suggest that taxpayer uncertainty among others due to complex and ambiguous tax laws provides a possible excuse to opt-out of the social contract with respect to tax compliance in the eyes of the general public.

Fourth and last, I speak to some recent evidence on the persistence of preferences in different domains and the durability of interventions on social norms. While the stability of other economic preferences was a subject of great interest in economics (see Einav et al. (2012) regarding risk preferences; Ubfal (2016) on time-preferences and Helliwell et al. (2016) as well as Bursztyn et al. (2020) on social norms), to the best of my knowledge there is no paper reporting evidence on the within-subject persistence of tax morale.<sup>10</sup>

---

<sup>10</sup>Recently, Besley (2020a) shows strong intergenerational stability of tax morale across birth cohorts. His data, however, do not look at the same individuals over time but at repeated cross-sections.

## 2 Survey details and experimental design

### 2.1 Surveying tax morale

**The main survey.** I field the survey experiment in the *German Internet Panel (GIP)*.<sup>11</sup> The panel is organized as a omnibus survey, where researchers of different fields collect information on attitudes and preferences relevant to economic and political decision making (Blom et al., 2015).<sup>12</sup> The survey ensures representativeness of the general German population aged between 16–75 years based on a probability sample of both the German offline and online population. A random draw of eligible households was recruited with face-to-face interviews and was invited by postal mail to participate in bi-monthly web-based surveys of the GIP. Households which did not have the necessary online access or hardware equipment to participate in the online survey were provided web-access, web-enabled devices and technical assistance (Blom et al., 2015). These efforts constitute a rare and invaluable feature in survey research (Blom et al., 2015, 2016). Blom et al. (2015, 2017) show that the inclusion of previously offline members of the general public in a web-based survey significantly increase survey coverage and representativeness.

The survey-embedded experiment analyzed in this paper was conducted in March 2019 (wave 40) of the GIP. Altogether, 4,890 respondents participated in wave 40. 72 participants dropped out until the survey experiment and the main outcome question was reached. Only 7 respondents chose not to answer the relevant outcome question to my experiment, leaving me with 4,809 persons that participated in the experiment and answered the tax morale question. While the experiment was fielded in wave 40 of the GIP, the panel structure of the survey allows me to gauge detailed information about each respondent from former waves of the GIP as well.<sup>13</sup> The summary statistics discussed in Section 2.4 outline all variables used in the main results of the paper.

**Measuring tax morale.** The measure of tax morale in this paper is based on the following GIP question: *How justifiable do you think it is to evade taxes?* The question is an adapted version of the World Value Survey (WVS) version.<sup>14</sup> Respondents could

---

<sup>11</sup>The GIP is a survey panel which is administered by the University of Mannheim in the framework of the "Collaborative Research Center 884 on Political Economy of Reforms" (SFB 884), which is funded by the German Science Foundation (DFG).

<sup>12</sup>A general description of the GIP and its method of data collection can be found in Blom et al. (2015) and online at <https://www.uni-mannheim.de/en/gip/for-data-users/methodology/>.

<sup>13</sup>The interested reader finds access to the detailed questionnaires via <https://www.uni-mannheim.de/en/gip/for-data-users/questionnaires-and-documentation/>. The survey data are available to the scientific community and every researcher can order access to the GIP data (for free) through the following website: [https://reforms.uni-mannheim.de/io-nas/sowi/reforms/internet\\_panel/Data\\_access/](https://reforms.uni-mannheim.de/io-nas/sowi/reforms/internet_panel/Data_access/).

<sup>14</sup>The WVS question reads: *Please tell me for each of the following statements whether you think it can always be justified, never be justified, or something in between: Cheating on tax if you have the*

answer on a 6 point scale with response categories ranging from absolutely justifiable, justifiable, rather justifiable, rather not justifiable, not justifiable, and absolutely not justifiable.<sup>15</sup> Higher values imply higher tax morale.

The question measures the general and unframed view of a respondent towards the norm (and thus, his or her intrinsic motivation) to comply with taxes. Only about 10.06% of all respondents stated that tax evasion is absolutely justifiable, justifiable, or rather justifiable to them (i.e. 1.14, 2.52, and 6.4% of all answers, respectively), with 89.94% of answers depicting that evading taxes is either rather not justifiable (19.67%), not justifiable (35.95%) or absolutely not justifiable (34.31%). Comparably high levels of tax morale for Germany were also found in Dörrenberg and Peichl (2018) in an earlier wave of the GIP. The numbers are also well in line with similar questions from the WVS and EVS (see Möhlmann (2014), Torgler (2002) and Feld and Torgler (2007) for tax morale in Germany and related West-East differences). In the WVS wave 7 (2017-2020) Germany also ranks 3rd with respect to stated tax morale levels in a sample of 48 countries (World Value Survey, 2020). Hence, Germany is a country with a relatively high tax culture, where evading taxes comes with a social stigma.

Survey-based measures of voluntary tax compliance are not free of criticism. For instance, respondents may elevate their self-reported morality in surveys and answer not truthfully (Andreoni et al., 1998). In line with this argumentation, Elffers et al. (1987) link tax audit and survey data and find that self-reported tax compliance can be substantially overstated. Other research, however, finds that survey-based levels of tax morale and (perceived) tax evasion are strong correlates (Richardson, 2006; Frey and Torgler, 2007; Torgler et al., 2008). Cummings et al. (2009) replicate the findings from artefactual field experiments on tax compliance with self-reported survey measures of tax morale. Several papers show also that tax morale is negatively correlated with the size of the shadow economy (e.g. Alm and Torgler (2006) and Torgler (2005)). Notably, Halla (2012) finds a direct causal link between tax morale and compliance behavior. According to Andreoni et al. (1998, p. 837) survey data is especially useful in order to test taxpayer motivations and behavior. This is exactly what the present paper does by using survey-based tax morale with experimental interventions to test changes in voluntary compliance.

Moreover, the tax morale question represents a hypothetical and abstract concept

---

*chance.* The question is measured on a 10 point scale with one (1) meaning 'never justifiable' and ten (10) meaning 'always justifiable'. The European Value Survey (EVS) uses an identical formulation of the question at hand. The tax morale question in the GIP does not include the qualification for taxpayers to evaluate tax evasion if they had the opportunity to cheat. This is because the present analysis focuses on changes in the justifiability of 'honest' mistakes in tax filing due to taxpayer uncertainty from complex taxes (see the experimental interventions in Section 2.2). Hence, the tax morale question is more general and abstracts from the possibility of seeing a chance to evade taxes and intentionally use it.

<sup>15</sup>The original question in German reads: *Fuer wie vertretbar halten Sie es Steuern zu hinterziehen?*. Response categories in original wording read: *fuer sehr vertretbar, fuer vertretbar, fuer eher vertretbar, fuer eher nicht vertretbar, fuer nicht vertretbar* and *fuer ueberhaupt nicht vertretbar*.

of tax compliance and does not ask directly about tax evasion behavior of respondents, which should increase honesty in responses. Similar to the WVS and EVS, the GIP survey is a comprehensive questionnaire where participants are not likely to get suspicious about individual tax compliance questions and may answer more honestly. Based on these considerations, I argue that the tax morale question in the GIP suits the purpose of measuring changes in voluntary tax compliance across the following experimental groups. Please note that the question explicitly asks the respondents about the justifiability of tax evasion rather than the acceptance of the legal counterpart of tax avoidance.

## 2.2 Experimental interventions

Before eliciting their tax morale, survey participants were allocated to three experimental groups in a between-subject design: the control group, the *Uncertainty* as well as the *Unequal knowledge* group. The experiment uses an augmented treatment structure which gradually adds information. This means that respondents in the *Uncertainty* group receive the same information as the control group plus extra information, and those in the *Unequal knowledge* group receive the same information as in the *Uncertainty* group as well as extra information. Figure A.1, A.2 and A.3 of the Appendix show screenshots of the respective experimental conditions.<sup>16</sup> Please note that while it was not possible for me to survey individual-level prior beliefs for the respective information treatments, treatment effects may not only occur due to information but also due to a higher salience or a priming of the information. However, given the fact that Section 3.4 provides evidence on the persistence of my treatment interventions, I argue that my treatments indeed convey new information that respondents then adapt in their stated decision-making.

Control group participants only read a short opener which states that tax evasion is frequently discussed by the media. All experimental groups are preceded by that same statement in order to introduce the topic of tax evasion and indicate its relevance to the respondents. Just like survey items in other omnibus questionnaires, most questions in the GIP are preceded by a short and neutral statement to familiarize respondents with a new set of questions. 1,603 participants are being allocated to the control group.

Members of the *Uncertainty* group receive the following statement after the short introductory opener about the relevance of tax evasion in the public debate: *Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure whether they report all incomes correctly in their tax declaration.* Subjects are thus informed about the tax confusion of others and their uncertainty about declaring their taxable income.<sup>17</sup> Since taxpayer

---

<sup>16</sup>See Dörrenberg and Peichl (2018) for a similar treatment structure.

<sup>17</sup>The treatment reflects foremost and primarily a tax morale channel which Luttmer and Singhal (2014) categorize as peer effects and social influences. According to their review of the literature, these

uncertainty cannot be easily measured, I do not rely on information on actual taxpayer behavior but on information about taxpayer perceptions regarding tax filing (see below).<sup>18</sup> The treatment represents a negative shock to intrinsic motivation (see also Besley et al. (2019)) through information on peer behavior (here, perceptions). To be precise, the information acts as a negative shock for those who were not aware of this fact and assumed that all or most taxpayers perceive tax filing as unproblematic. Specifically, respondents are informed about the perceived difficulty among others in the German population with respect to filing income taxes. The treatment is motivated by the fact that half of the respondents in a prior GIP survey stated that they think filing their annual income tax declaration is difficult or very difficult if they have filed their taxes themselves (see Blesse et al. (2021)). Since I did not want to restrict the statement to those who file their taxes themselves (about 62.8% of all respondents in the GIP, see Blesse et al. (2021)) as compared to those who do not, i.e. if they do not have to file a tax return (for instance, retirees) or if someone else files their return on their behalf (e.g. a spouse or a tax adviser), I refer only to "many citizens" as the relevant peer group and do not provide exact figures on taxpayer uncertainty. The treatment then states that many citizens are often not sure, whether they report all incomes correctly in their tax declaration.<sup>19</sup> Among others, this perception is well in line with recent papers reporting that complexity in tax filing due to many tax rules leads to cognitive costs and poor decision quality (Aghion et al., 2017; Benzarti, 2020; Abeler and Jaeger, 2015) as well as creating ambiguity in compliance behavior (Bellemare et al., 2019).

The treatment information in the *Uncertainty* condition effectively reduces the moral costs of evasion in the context of complex taxes since it signals that uncertain taxpayers are not necessarily compliant. Non-compliance to complex tax rules is likely to be unintended since people would not be aware of the consequences of their actions

---

channels, however, are not mutually exclusive and often go hand in hand. In fact, the treatment links intrinsic motivation for tax compliance to government policy, e.g. the provision of transparent tax laws, and can thus also be attributed to reciprocal motives of tax morale. If tax morale is (at least) in part derived from benefits of taxation and the legitimacy or fairness of government policy, then tax morale could be undermined (e.g. Besley (2020b), Besley et al. (2019) and Luttmer and Singhal (2014)). In this specific case, people may perceive complex taxes as unfair (as shown by Blesse et al. (2021)) and may opt-out of the social contract of complying with tax laws in response to unfair complexity. Early lab evidence of Alm et al. (1992) further indicates that uncertainty in fiscal parameters of the economic decision to evade taxes in the presence of public goods, i.e. in the presence of reciprocity, decrease tax compliance significantly, highlighting a complementary role of uncertainty and reciprocity.

<sup>18</sup>Also, other papers resort to randomly giving survey-participants information about the beliefs, preferences and actions of others (Bursztyn et al., 2020.?; Coibion et al., 2018). Just as in the present paper, the relevant information from these papers is derived from survey-based perceptions.

<sup>19</sup>Supporting evidence for this implication of uncertainty comes from another GIP survey from March 2020, which asked about exactly what makes the income tax system difficult and where 40.16, 41.69 and 45.48% of all respondents reported that the scope of documentation requirements, the fear of forgetting something important when filing one's tax return or the fear of filing something wrong in the tax return is making income tax declarations difficult (multiple answers were possible, see Blesse et al. (2021)).

(or would not know any better).<sup>20</sup> The treatment represents a descriptive social norm which likely leads respondents to think evasion is more justifiable if taxpayer uncertainty is common and resulting non-compliance is likely due to 'honest' mistakes of taxpayers. Essentially, the condition also tests whether respondents are conditionally cooperative such that their intrinsic motivation to comply depends on the compliance behavior of others (e.g. Frey and Torgler (2007) and Traxler (2010)). Altogether, I hypothesize that the tax morale for this treatment is lower (as compared to the control group) because it informs respondents that taxpayer uncertainty is common among the general population. 1,604 respondents receive the *Uncertainty* information.

Participants of the augmented treatment group additionally receive a further statement after reading the opener. The respective *Unequal knowledge* group receives the following message: *Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration. Citizens who know tax laws well or use a tax attorney can possibly take advantage of the complexities of the tax law and reduce their tax burden.* While the first treatment only gives information regarding the uncertainty of ordinary citizens with respect to filing taxes, the second treatment highlights that society is comprised of heterogeneous tax-payer groups by stating that other taxpayers have the relevant tax knowledge (or access to it through tax attorneys) and can actively influence their after-tax income. The *Unequal knowledge* condition thus captures the existence of distinct coexisting tax-subcultures which differ in their levels of tax compliance (Lewis, 1982; Traxler, 2010). In the present context, these groups differ in compliance due to different abilities to comply with complex taxes and to influence their after-tax income through tax knowledge.<sup>21</sup> Knowledgeable taxpayers are also communicated as to make active use of tax knowledge in order to optimize their tax burden,

---

<sup>20</sup>Please note that the German tax law distinguishes between tax evasion and careless tax deficiency (so called *leichtfertige Steuerverkürzung*). Both are based on withholding information to the tax authorities which leads to a reduction of tax revenues but tax deficiency is based on light-mindedness and a lack of knowledge while tax evasion is done on purpose (Fromm, 2019). Tax deficiency is thus associated with lower levels of punishment than tax evasion and is classified as a petty offense (unlike the criminal act of evasion). The law differentiates between these two acts based on the personal (ability to process) tax knowledge. While the outcome question of the experiment, however, clearly refers to the act of evasion, the treatment may lead to more empathy towards uncertain taxpayers which are presumably less able to comply with the law and their non-compliance to be the likely result of 'honest' mistakes in reporting. This logic is also in line with practiced tax law, although taxpayers are obliged to search for relevant information to a reasonable extent. Legal information in this footnote are based on Fromm (2019).

<sup>21</sup>One may argue that everyone could potentially have access to relevant tax knowledge through active websearch and consulting tax authorities (Hoopes et al., 2015), tax advisers (Chetty and Saez, 2013), preparation software or their peer or personal networks (Alstadsæter et al., 2019; Chetty et al., 2013). Acquiring information is, however, costly. According to Hoopes et al. (2015) taxpayers are rationally attentive and search for tax-related information only when it seems useful to them. Other evidence indicates that learning about taxes and how to benefit from them takes time, is far from complete, and is especially hard for individuals with lower income and education, as well as the elderly (Aghion et al., 2017). Hence, even though people learn about taxes in various ways (if need be), complex taxes remain an important source of uncertainty for many taxpayers when making economic decisions.

i.e. they exert avoidance activities. Evasion motives are explicitly not referred to in order to avoid possible confounding effects due to the fact that people could associate the *Unequal knowledge* group with such intentions. The condition also qualifies the first statement further since many (but not all) citizens are subject to this uncertainty with respect to complex taxes. The treatment is again motivated by earlier work on tax filing perceptions as well as the finding that a majority believes that complex taxes (which are based on abundant deduction possibilities) are contributing to larger perceived income inequalities (see Blesse et al. (2021)) and that the rich engage in more tax evasion and avoidance opportunities (Stantcheva, 2020).<sup>22</sup> Essentially, the *Unequal knowledge* treatment augments the *Uncertainty* condition in such a way that states many citizens may be confused about filing their taxable income but other and more knowledgeable groups exist which have better opportunities to optimize and improve their after-tax incomes.<sup>23</sup>

Conditional cooperation may be different with respect to these two groups of taxpayers who differ in their opportunities to optimize their after-tax income through complex tax rules (Traxler, 2010). The augmented treatment design enables me to test whether people only care for the fact that taxpayer uncertainty exists or whether they also care for the fact that there are others who know how to report their taxes. For the empirical analysis, different outcomes are possible depending on what respondents perceive as the social reference group (Traxler, 2010) when forming their beliefs on tax morale under heterogeneous knowledge levels of the general population: First, if tax morale is significantly higher for the *Unequal knowledge* than for the *Uncertainty* treatment, then respondents likely view that uncertainty is less prevalent among the general population than under the first treatment condition. Second, if tax morale is significantly lower for the *Unequal knowledge* than for the *Uncertainty* condition, then respondents appear to empathize relatively more with uncertain taxpayers and view it as a legitimate reason not to comply with taxes, likely due to the fact that respondents perceive filing errors as unintentional and because information acquisition or tax advise are perceived as costly. Third, if the estimated coefficients of both treatments are different from zero but the difference between both comparison groups is statistically indistinguishable, we cannot

---

<sup>22</sup>This perception of the general population corresponds with papers which find that, predominantly, the very rich evade (Alstadsæter et al. (2019)), that more educated (and richer) individuals know more about the tax code, learn faster and can adjust accordingly to tax incentives (Aghion et al., 2017) and that tax professionals can help their clients to reduce tax liabilities (Andreoni et al., 1998; Slemrod, 1989) and are an important source of tax-relevant information for their clients (Chetty and Saez, 2013; Chetty et al., 2013).

<sup>23</sup>In addition to fairness considerations due to reciprocity (see footnote 17 above), the unequal opportunities to influence after-tax incomes presented here may also change fairness perceptions with respect to the allocation of tax payments across taxpayers (see Hashimzade et al. (2013), p. 972). Complex taxes may thus also reduce tax morale through a more negative fairness perception of the tax system. Blesse et al. (2021) finds that informing individuals about higher scope for avoidance and evasion opportunities in complex tax systems undermines the perception of the existing tax rules to be fair. Compliance decisions for complex taxes are also directly linked to distributional motives (Bellemare et al., 2019).

reject the notion that respondents only care for the disadvantaged and unaware taxpayers when stating their tax morale. Please note that unlike to the *Uncertainty* treatment, I do not have a clear a priori expectation about the effects of the second condition.

The respective treatment condition is comprised of 1,602 respondents in total.

### 2.3 Follow-up survey

In order to test the persistence of the experimental conditions of my main experiment in wave 40 of the GIP, I use a follow-up question two months after the main survey in the subsequent wave of the GIP. Follow-up surveys are also typically an effective way to alleviate concerns about experimenter demand bias (e.g. Alesina et al. (2018); Haaland and Roth (2019, 2020)). Altogether, 4,824 respondents participated in wave 41 of the GIP in May 2019. 4,760 people answered the respective question on tax morale.

The tax morale question itself is slightly different, which should make it harder to find persistent effects. The wording of the tax morale question in wave 41 reads: *How justifiable do you think it is to evade taxes if a good opportunity to do so presents itself?* with the same response categories as in wave 40 of the GIP. The question now added a qualification with respect to the justification of tax evasion if a good opportunity to do so presents itself. The question is thus closer to the WVS version and is equal to the question Dörrenberg and Peichl (2018) use in an earlier wave of the GIP.<sup>24</sup> The question was preceded by the same opener statement as in wave 40 to gauge the interest of respondents and to familiarize them with the context of tax evasion.

The distribution of responses for members of the control group appear to be largely stable over time and in between questions (see Section 3.4). Note that any concerns about consistent answering between subsequent survey waves of the GIP are very negligible since both waves are fielded with a comparably large time gap of two months when compared to other follow-up surveys present in the literature which are typically fielded only a few weeks after the main intervention. Consistent survey answers in wave 40 and 41 are further alleviated by the fact that the GIP is a omnibus survey module on a repeated basis for the same set of individuals. This means that in both survey waves, tax morale questions were placed in between questions of other topics which should further disguise the connection between the question in the main survey and the follow-up question. While the tax morale question in wave 40 was placed in between questions about the debt brake and fiscal equalization in Germany as well as feedback questions at the end of the survey,

---

<sup>24</sup>The choice not to ask respondents exactly the same question in a follow-up experiment is inspired by recent efforts made in obfuscated follow-up surveys (pioneered by Haaland and Roth (2019) and Haaland and Roth (2020)) which aim to alleviate concerns of experimenter demand effects in survey experiments. Effectively, obfuscated follow-up surveys measure persistence of treatment effects from information experiments using a short follow-up survey after the main information intervention while taking steps to disguise the connection between main survey and follow-up (Haaland et al., 2020).

the tax morale question in wave 41 was embedded in a question about redistribution preferences as well as public support for EU policy reforms. The respondents are also not aware of the identity of the involved researchers when answering questions.

## 2.4 Summary Statistics

Table 1 provides detailed summary statistics of the demographic structure of the GIP survey data.<sup>25</sup> The table reflects the respective sample information regarding household size, age structure, marital and retirement status as well as employment, education, household income and political orientation. Not all information were necessarily surveyed in wave 40 where the survey experiment was fielded but may come from earlier survey waves. Specifically, I matched information on both political preferences derived from a variable with respect to a person's self-placement on a 11 point left-right scale as well as elicited net household incomes from wave 37 to the current survey. Moreover, some attributes of respondents that do not regularly change over time (e.g. gender or education status) are typically updated only about once a year and are also linked to the current survey wave.

The table indicates that the most frequent household size covers two household members (43%) and 17% of participants live in single households. The sample pool reaches a balanced gender composition with 48% females and covers all age cohorts with a somewhat higher share of older participants (24% being between 49 and 58 years old and 31% being at least 58 years old) corresponding with an ageing German society. 58% of all participants are married. The sample comprises 46% full-time workers and 19% retirees. Participants are also well distributed across education levels. About 31% of all respondents are highly educated, i.e. have a tertiary degree. Individuals are split in 5 net household income groups ranging from poor (less than 1500 Euro) to rich (more than 4500 Euro). 30%, 27% and 14% of respondents have net household incomes of less 1500 Euro,  $1500 \geq x < 2500$  Euro or  $2500 \geq x < 3500$  Euro. Only 6% and 5% have either between 3500 and 4500 Euro or are 'rich' with more than 4500 Euro. 13% of all participants do not declare their household income to the interviewers and income information of 6% of individuals cannot be matched to the current survey. Furthermore, the sample appears to have a relatively balanced distribution with respect to political ideology, with 40% of individuals being labeled as conservatives and 44% being categorized as left-wing. 11% of respondents do not declare a political ideology (non-partisans).<sup>26</sup>

---

<sup>25</sup>Similar to Blesse et al. (2021), I argue that since the GIP is based on a probability-based representative sample of the general German population, there is no need to compare demographic background information from survey respondents to official population statistics or other representative data.

<sup>26</sup>Left-right preferences are measured on a 11-point scale from right to left, where 'conservative' is  $\leq 5$  on this scale, see Blesse et al. (2021) for a similar coding). I cannot match political preferences for 5% of respondents to wave 40 of the GIP.

## 2.5 Balancedness of experimental groups

Table A.1 of the Appendix tests the balancedness of the experimental groups with respect to the demographic information summarized earlier in Section 2.4. Specifically, I run three OLS regressions for each covariate which can be formulated as follows:  $y_i = \beta Covariate_i + \epsilon_i$ . While  $Covariate_i$  represents the respective covariate listed above, the dependent variables depict dummy variables for the respective treatment status for each participant  $i$ , i.e. being either among the control group, the *Uncertainty* or the *Unequal Knowledge* group. Testing balancedness of covariates across treatment groups in survey-embedded experiments using this procedure is standard in the literature (for similar approaches, see for example Alesina et al. (2018)). Given 9 covariates and 3 dummies representing my experimental conditions, I end up with 27 separate regressions in total. These regressions yield 66 estimated coefficients and I find that only one is significant at the 10% significance level and 5 and 2 are significant at the 5 and 1% significance level, respectively. This is by and large in line with the estimates being significant within their margin of error. All in all, randomization by computer appears to have worked properly, which allows me to derive causal inference using the implemented survey experiment. Please note that I still control for covariates in my main specifications in order to reduce standard errors and increase the precision of the estimates of interest.

## 3 Results

This section outlines the results of the experimental interventions on tax morale.

### 3.1 Main results

**Graphical evidence.** Figure 1 shows the average response for the tax morale question and a 95% confidence interval. Average responses on tax morale on a 6-point Likert scale equal 5.08 in the control group as well as 4.78 and 4.83 in the *Uncertainty* and *Unequal Knowledge* group, respectively. P-values from two sample t-tests indicate that responses differ significantly with  $p < 0.01$  between control group and both treatment groups. The respective p-value of 0.1381 from this test does, however, not reject the hypothesis of mean equality of the *Uncertainty* and *Unequal Knowledge* group.

**Regression model.** Now I present the results of the following OLS regression

$$Morale_i = \beta_1 Treat_i + \beta_2 Covariate_i + \epsilon_i,$$

where I measure tax morale  $Morale_i$  of respondent  $i$  on a 6-point Likert scale (ranging from 1 evasion 'absolutely justifiable' to 6 'absolutely not justifiable').  $Treat_i$  represent

dummy variables for either the *Uncertainty* and *Unequal Knowledge* group. The respective coefficients of these variables are the treatment effects of interest relative to the omitted control group. In my preferred specification, I also control for a full set of covariates ( $Covariate_i$ ) which comprise demographic information such as household size, age groups, marital status, gender, employment status, retirement and education categories as well as disposable household income and political preferences, i.e. left-wing ideology.

Reported standard errors are robust to heteroscedasticity. I use OLS specifications in the main regressions in order to ease interpretation of the treatment effects of interest. Alternative model specifications, including Linear Probability Models (LPM) as well as Ordered Probit Models are tested as robustness checks in Section 3.2.

**Main effects.** Table 2 shows the regression results for the effects of the experimental interventions on tax morale. Column (1) shows the coefficients of the treatment dummies relative to the control group without controlling for any covariates. The other columns add further covariates in order to improve statistical precision. Specifically, column (2) controls for demographic information including gender, age, marital status, household size, employment status, retirement status, and education, while column (3) adds household income information and column (4) additionally accounts for political preferences.<sup>27</sup>

The coefficient of the *Uncertainty* treatment is negative in all specifications and statistically significant at the 1% level. This suggests that providing information about the prevalence of uncertainty in filing taxes among citizens in the context of complex tax systems does indeed have a negative effect on tax morale, making people more accepting of tax evasion. Uncertainty here likely results in 'honest' mistakes of taxpayers in their income reporting rather than evading on purpose. The effects are very stable across specifications. In my preferred specification of Column (4) which controls for a full set of controls, the *Uncertainty* treatment reduces tax morale by about 6.8% (-0.300/4.358) when evaluated at the conditional control group average, relative to the control group which did not receive further information. Therefore, the results imply that individual tax morale is elastic to salient uncertainties among peers from complex tax systems. Hence, respondents use information about taxpayer uncertainty as an excuse for a lower motivation to comply with taxes and judge evasion less harshly. The effect relative to the control group amounts to 0.278 of a standard deviation. The effect size may be not very large in economic terms but needs to be interpreted against the background of a very strong support for tax morale among respondents.

The *Unequal knowledge* treatment also shows consistently negative treatment effects across all columns and is statistically significant at the 1% level when compared to the control condition. The treatment essentially augments the uncertainty treatment with

---

<sup>27</sup>For a discussion of the effects of individual covariates on tax morale, please see the paragraph below.

a contrasting statement that certain people (i.e. those with tax knowledge or those with access to tax advise) can indeed make proper use of the tax system in their favor. The respective coefficients are thus somewhat weaker as compared to the *Uncertainty* treatment. When evaluated at the control group average, the treatment effect ranges from a -5.5 to -5.6% change to the control group. Again, point estimates are very stable across specifications. Joint F-tests do not show that both treatment groups are significantly different at conventional levels. In my preferred specification of Column (4), the p-value of the respective F-test is 0.143. Also in economic terms the difference of effect sizes is small, representing only about 0.1 standard deviations. Hence, highlighting the presence of other knowledgeable taxpayers which can have optimize their tax burden does not affect the likelihood of people using the fact that other taxpayers are confused when filing taxes as an excuse for a lower motivation to pay taxes. Taken together, these findings suggest that information about taxpayer confusion from complex taxes among others offers an opportunity for taxpayers to make excuses and reduce their own tax morale. Highlighting the fact that other taxpayers may be less affected by this uncertainty and can utilize complex tax rules does not affect voluntary compliance further.

**Discussing the anatomy of tax morale.** Table 3 shows the ‘anatomy’ of elicited tax morale using OLS regressions of tax morale on different sets of covariates via a broad set of observable respondent characteristics. The control variables are similar to those being used in the baseline regressions studying the effect of both treatment interventions on tax morale (see Table 2). However, Table 3 illustrates the influence of individual background characteristics on tax morale. Thus, the results of Table 3 show which (socio-demographic) group is more or less likely to find tax evasion justifiable. Please note that the respective coefficients of the included covariates show only (conditional) correlations but do not reflect causal relationships. Detecting heterogeneous responses to the tax morale question allows me, however, to get an impression of how well drivers of tax morale among the general population in the GIP correspond to existing evidence.

Column (1)–(3) always control for the allocation of respondents to both treatment interventions and gradually add covariates. Specifically, Column (1) considers demographic features (gender, age, marital status, household size, employment status, retirement status, and education), column (2) adds net household income categories and column (3) also includes political preferences, i.e. left-right political ideology. The results suggest that important determinants of elicited tax morale are age and gender. Specifically, we find that older respondents have a statistically higher tax morale which is independent of other factors that go along with seniority such as marital status, employment, being retired or having higher income. Moreover, women have a higher tax morale. Both age and gender effects are in line with the evidence from developed and developing countries (e.g. Frey and Torgler (2007) and Alm and Torgler (2006); for a brief review of

this literature, see Dörrenberg and Peichl (2013)) and also fit to findings from a related paper on tax morale using GIP data (Dörrenberg and Peichl, 2018). Corroborating with most existing papers (again, see the brief review in Dörrenberg and Peichl (2013)), being married correlates significantly with higher tax morale, albeit only at the 10% level (for example, see also Alm and Torgler (2006)). Unlike Dörrenberg and Peichl (2018), however, we do find robust evidence that stated tax morale increases statistically with a higher individual education status. According to Table 3, education is a significant and positive determinant of tax morale. While existing evidence is also ambiguous with respect to the effect of income on tax morale,<sup>28</sup> I find statistically significant effects for the highest declared income category suggesting that higher incomes have a somewhat higher tax morale. Interestingly, I also find statistically significant and positive effects on tax morale when respondents do not declare their income in the survey. Furthermore, the results suggest that respondents with a left-wing political ideology have a higher tax morale compared to rather conservative individuals. This finding matches evidence from Dörrenberg and Peichl (2018) using an earlier wave of the GIP. Finally, I do not find significant effects of household size, employment status, and retirement on tax morale.

### 3.2 Robustness checks and extensions

**Sensitivity checks.** In the following, I present evidence that speaks to the robustness of the main results. Firstly, I test whether a binary measure of tax morale leads to different treatment effects. In order to do this, I code tax morale as a dummy variable being 1 if responses are equal to 4, 5 or 6 for the question: "How justifiable do you think it is to evade taxes?" which is measured on a 6-point scale, and 0 if otherwise. Table A.2 of the Appendix estimates the respective Linear Probability (LPM) models on both treatment dummies and using a binary measure of tax morale. The effects remain qualitatively very similar to the main specification with both treatment effects being consistently significant across specifications at the 1% level. Again, the *Uncertainty* treatment results in larger point estimates than the *Unequal knowledge* group, although both groups are repeatedly not significantly different from one another according to joint F-tests.

While the main specifications (and the LPM models) use OLS in order to ease the interpretation of the treatment effects, the tax morale question is measured on a 6-point Likert scale, which may speak for alternative model specifications that directly account for the discrete and ordinal nature of the outcome variable. Hence and secondly, I use ordered probit models to account for the ordered nature of the tax morale variable. Table A.3 of the Appendix shows the respective results. The treatment effects are again very similar to the baseline results using ordinary least squares. Accounting for the ordinary nature

---

<sup>28</sup>Some papers find positive (e.g. Lago-Peñas and Lago-Peñas (2010)), some insignificant (e.g. Konrad and Qari (2012)) and others find negative effects (e.g. Alm and Torgler (2006)).

of the outcome variable, we can, however, reject that both treatment groups are equal at the 10% level. Therefore, it seems that the negative treatment effects of the *Uncertainty* condition are statistically somewhat stronger than with the second augmented treatment group. However, just like in the main results the respective differences between both experimental groups are typically not meaningful in an economic sense.<sup>29</sup>

**Do changes in tax morale vary across compliance levels?** Moreover, I show how the experimental conditions change the distribution of response categories differently across experimental groups. Preliminary evidence can be taken from a test of difference in distributions.<sup>30</sup> To test the shift in response distributions econometrically, I do the following: For each of the six response categories of the tax morale question ranging from (1) the respondent finds tax evasion absolutely justifiable to (6) the respondent states that he/she deems tax evasion absolutely not justifiable, I create an individual dummy variable. Then, I use each of these six binary tax morale variables as alternative dependent variables in LPM specifications. Table 4 shows the respective results for all response categories and both treatment conditions.<sup>31</sup> All specifications account for a full set of covariates. This exercise helps to understand whether the negative information effect of tax confusion among peers on individual tax morale results from a decline in several answer categories or only from the highest levels of motivation (see Dwenger et al. (2016) and Dwenger and Treber (2018) for the role of different motivational types for the effectiveness of rewards or shaming on tax compliance, respectively).

While most answers can be found in the highest response category of tax evasion being absolutely not justifiable (41.42% of answers in the control group), both treatment conditions show large negative treatment effects in this category which are statistically significant at the 1% level. While respondents in the *Uncertainty* group chose this category about 12.5 percentage points less often on average, subjects in the augmented treatment chose it about 8.1 percentage points less often. According to the table, the entire negative average treatment effects emerge from a reduction of answers in the highest

---

<sup>29</sup>In addition to potential concerns about the implications of misspecification to my main estimates, I also check whether the inclusion of covariates in the main results leads to a sample selection bias since not all covariates are necessarily observed for all respondents. Table A.4 of the Appendix shows the regression results when using the reduced sample of 4,635 respondents for which all covariates are available, which we use as control variables across all specifications. My main results prevail when gradually adding control variables and keeping the number of observations fixed throughout all specifications.

<sup>30</sup>While I rejected the hypothesis of mean equality between both treatment groups in my baseline estimates, the distributions appear to differ between groups. P-values from pairwise Kolmogorov-Smirnov tests indicate that responses differ significantly with  $p < 0.01$  between control group and both treatment groups, respectively. The respective p-value of 0.075 from that test rejects the equality of the distributions in the *Uncertainty* and *Unequal knowledge* groups at the 10% level.

<sup>31</sup>Figure A.4 of the Appendix illustrates the treatment effects across individual answer categories to the tax morale question graphically without conditioning on additional covariates in a regression framework. The results are, however, very similar to Table 4 in the main text.

category of tax morale, i.e. from the most "honest taxpayers". The difference in effects between both treatment groups, however, is significant at the 1% level. Other answer categories do not show significant negative treatment effects. However, the treatment groups shift response patterns somewhat differently. While both groups show significantly more answers in response category 4 (tax evasion somewhat not justifiable), i.e. they alleviate concerns about evasion by making it look more legitimate through the prevalence of uncertainty in filing taxes from complex tax declarations, the augmented *Unequal knowledge* treatment significantly shifts preferences from category 6 even towards response category 2 (evasion being justifiable). The *Uncertainty* group also significantly shifts respondents towards category 3 (evasion being somewhat justifiable) if compared to the control group. Altogether, this set of findings speaks to the fact that peer information about taxpayer confusion especially reduces the share of highly motivated individuals in our between subject design. The *Unequal knowledge* condition has a significantly lower effect on very high levels of tax morale than the *Uncertainty* group and thus alleviates the extent to which participants perceive tax confusion to be present among others.

### 3.3 Heterogeneous treatment effects

In the following, I analyze whether certain groups of participants respond more or less elastically to either of the treatment interventions with respect to their tax morale. In order to do this, I interact the treatment dummies with different characteristics of the sample population. Given the longitudinal structure of the survey data, I can exploit rich background information on each respondent from previous survey waves. I implement an explorative analysis as to whether specific individual characteristics (listed below) can explain my baseline findings.<sup>32</sup> All related results need thus to be interpreted as suggestive and not as causal. Specifically, I tested the following heterogeneities:

**Demographic structure.** Demographic information comprise marital status, sex, being full time employed, religion, being rich (i.e. in the highest income category), having dependent work, being self employed (all drawn from wave 40) and education (wave 39).

**Individual perceptions of income tax system.** Other tests explicitly check for heterogeneous treatment effects by individual tax perceptions. Specifically, I test sub-group

---

<sup>32</sup>Given the vast amount of detailed background characteristics for each respondent (e.g. a 20 to 25 minute survey was sent to all participants every second month since the start of the GIP in 2012 until my experiment was fielded in March 2019), which all could be used to identify predictors of treatment effects using machine learning (ML) algorithms (Chernozhukov et al., 2018; Wager and Athey, 2018), I abstain from data driven selection methods for my sub-group analysis due to power and attrition problems which would complicate ML methods. Instead, I choose to run the sub-group analysis of treatment effects in an exploratory fashion in order to document interesting results following the argument in Duflo et al. (2020). At the same time, I interpret these results as suggestive and not as causal in nature.

effects of the treatment indicators with respect to the perceived difficulty in filing one's own income tax declaration, the personal use of itemization in annual income tax declarations, the preference for a more simple tax system as well as the perceived implication of deductions on income inequalities (wave 36; for more detail, see Blesse et al. (2021)) as well as prior levels of tax morale drawn from wave 39. I also use data for individual attitudes on whether tax evasion should be prosecuted and whether taxpayers should have more influence on public spending (both retrieved from wave 39).<sup>33</sup>

**Attitudes towards redistribution and government.** I also check whether redistributive preferences (wave 35) and government trust (wave 40) affect treatment effects.

Generally, I find that the treatment effects for both experimental conditions appear to be very stable across most sub-groups. The results suggest that only political ideology and older age groups (both from wave 40) appear to have significant and robust effects on the experimental interventions and thus, contribute to my treatment effects.

Table 5 depicts the respective results from the interaction models. I only report interaction effects which are significant at conventional levels. Note that I also do not find that other variables explain the treatment effects in a consistent manner. Column (1)-(2) report interaction effects of both treatments with measures of political ideology. Column (3)-(4) reports interaction effects of both treatments with age categories. For each dimension of heterogeneity, I first report results without control variables (column (1) & (3)) and then add a full-set of covariates with information on demographics, income and political preferences used in earlier regressions (column (2) & (4)).

Older respondents ( $48 < \text{age} \leq 58$  and  $\text{age} > 58$ ) respond significantly more negative regarding their tax morale when being confronted with the prevalence of uncertainty in filing income taxes. The negative effect of the *Uncertainty* group in Table 5 increases with age but only becomes significant for the two oldest age groups at conventional levels.<sup>34</sup> This corresponds with evidence from Blesse et al. (2021) that preferences towards more simple tax systems are more prevalent among older citizens. These findings also correspond to results from Aghion et al. (2017) that older individuals know less about tax incentives and are also less able to learn, indicating cognitive costs in adjustment

---

<sup>33</sup>I also test whether it matters if respondents declare themselves or whether they let others report their tax return (such as spouses or tax attorneys). The respective information comes from wave 36 of the GIP. While I cannot exactly distinguish the third party that reports one's income, this information may be valuable by itself in order to answer whether the exposure to self-reporting makes one more sensitive to information about tax complexity in tax reporting. While I do not find that own reporting is driving the respective treatment effects in my experiment, this may be due to the fact that I can only use about 2,000 observations of my main estimation sample (including controls) from wave 40.

<sup>34</sup>I also find a significant negative effect for respondents who are retired but the effect is arguably nested in the oldest age group. Since the interaction effect for  $\text{age} > 58$  prevails even when controlling for a full set of controls (including retirement status), I argue that the age effect drives the main results with respect to the *Uncertainty* group independent of the retirement status of the respective participant.

to complex tax rules. My results suggest that this group is more lenient towards tax evasion when being informed about the uncertainty of many taxpayers when filing taxes. Older respondents also have a higher tax morale on average (see Table 3). I argue that higher cognitive costs among old individuals (see Aghion et al. (2017)) make them more susceptible for using the uncertainty of others as an excuse for lower tax morale.

Table 5 also shows that the negative effect of the augmented treatment group can be explained by political ideology. Participants with a left-leaning political ideology have a significantly more negative shift in their tax morale in the *Unequal knowledge* treatment as compared to conservatives. Thus, left-wing ideology does not only correlate with higher tax morale (for the respective results, see Table 3), it also makes people also more receptive for distributional concerns regarding tax morale in the context of tax complexity. This is in line with Alesina et al. (2018) who find a large role of political preferences in a survey experiment on redistributive preferences. In their study, left-wing respondents demand more redistribution after receiving pessimist information about inequality of opportunity. Corresponding with their finding, I find that left-wing participants care more about the uncertainty of ‘other’ taxpayers when they are framed as inequitable. Specifically, they reduce their tax morale in response to the information that some taxpayers cannot optimize their after-tax income due their lack of tax knowledge, while others with that respective knowledge or sufficient access to it can. Accordingly, subjects with left-wing ideology appear to find tax evasion more justifiable, likely as a means of redistribution towards the ‘unlucky’ mass of confused tax filers. My results are robust to interaction models which measure political preferences using individual German parties where I find that the respective effect is driven by the socialist party (*Die Linke*).

### 3.4 Follow-up

Now I present the empirical results on the persistence of the experimental conditions using a follow-up survey two months after the main experiment (see 2.3 for details).

**Main results.** In order to draw conclusions on the persistence of the experimental conditions regarding tax morale, I match the respective treatment and control group variation from the tax morale experiment in wave 40 of the GIP to the subsequent follow-up survey wave (wave 41) where a similar but slightly changed tax morale question was asked to the same respondents (see Section 2.1 and 2.3 for details on the exact wording). Given the panel structure of the GIP, I can exploit how responses to the tax morale survey item have changed for the same individual two months after the main experiment.

First, it may be interesting to know for the reader how persistent tax morale preferences are over time. Given that most papers on the issue of survey-measures of tax compliance use repeated cross-sections of survey data rather than longitudinal surveys

(e.g. WVS or EVS), not much is known about how much individuals change their opinion on the acceptability of tax evasion over time. To the best of my knowledge, I provide first evidence on within-subject stability regarding tax morale. Figure 2 shows the changes in individual response patterns between survey waves for control group members. Specifically, it depicts the difference of tax morale responses from wave 41 to wave 40. Notably, the majority of participants with about 55.3% do not change their opinion with respect to tax morale. Only 20.6% of respondents show a decrease and approximately 24.09% of participants state an increase in tax morale between subsequent survey waves. Altogether, tax morale preferences appear to be rather stable over time for a given individual.

Second, I study how the experimental interventions of the main experiment change tax morale over time using the follow-up question. In order to do that, I restrict the sample of the follow-up survey wave to respondents who were also part of the experiment in the main survey. This leaves me with a high rate of 91.47% ( $N=4,405$ ) which underwent the experiment in wave 40 and answered the tax morale question in the follow-up survey. Accordingly, I argue that attrition should not be a main concern when interpreting the follow-up results.<sup>35</sup> Covariates in the follow-up are also well balanced with respect to the relevant treatment interventions from wave 40 (see Table A.6 of the Appendix).

In order to estimate treatment effects, I simply regress tax morale survey responses (now from wave 41) on the treatment dummies of the main experiment in wave 40 using simple OLS, omitting the control group dummy. Table 6 depicts the results of the respective estimations. While column (1) includes only the treatment dummies in the regression but abstains from including covariates, columns (2) to (4) gradually add information on demographic characteristics, income and political ideology of the individual respondents.

While the *Uncertainty* treatment does not appear to have strong lasting effects on stated tax morale, its augmented version, the *Unequal knowledge* condition, which further specifies who can properly use tax deductions to their advantage, indeed shows a significant negative effect on tax morale. The effect is statistically significant at the 5% level and appears to be somewhat weaker than in the main experiment (i.e. 27% of the estimate in the main survey). The effect is also significantly different from the effect of the *Uncertainty* condition as illustrated by the rejection of a joint F-test with a p-value of 0.086. However, the step-wise inclusion of covariates in order to increase precision of estimation in columns (2) to (4) renders the treatment effect of the augmented experimental group insignificant at conventional levels (i.e. with a p-value of 0.120).

As seen before in Section 3.2, average treatment effects of peer information may, however, mask important heterogeneities and especially very honest taxpayers may be deterred by the tax confusion of others. Hence, I again estimate the respective treatment effects of treatment conditions on individual tax morale answer categories in order to in-

---

<sup>35</sup>Table A.5 of the Appendix also shows that neither of the treatment conditions of wave 40 significantly affect the likelihood to answer the follow-up question. The calculations are based on 4,438 observations.

vestigate whether different types of tax morale react differently even two months after the intervention. Table 7 shows the respective results. It seems that also in the follow-up only very honest tax morale categories are negatively affected by peer information about filing uncertainty. Both treatment conditions are significantly different from the control condition although the *Unequal knowledge* effect is somewhat larger and statistically significant at the 1% level as compared to the 10% significance level of the *Uncertainty* condition. Both effects are not significantly different from each other. While the negative effect among very honest tax responses is shifted to harsher judgements on tax evasion in the main survey (see Table 4), we now see a statistically significant shift of answers towards category 5 (i.e. from evasion being absolutely not justifiable towards evasion being not justifiable only). Altogether, it appears that peer information about tax uncertainty of others has persistent effects on very honest levels of tax morale but that resulting changes in judgments grow less harsh over time and effects become somewhat less strong. The *Uncertainty* and *Unequal knowledge* effect amount to 24.6% (-0.031/-0.126) and 64.2% (-0.052/-0.081) among very honest tax morale levels in the follow-up when compared to the main survey (and the baseline results in Table 2), respectively. Persistent treatment effects also suggest that new information rather than the mere salience of the conveyed arguments is the *main* source in the observed changes in voluntary tax compliance.<sup>36</sup>

**Heterogeneous effects on persistence.** As for the main experiment, I only report heterogeneous effects of the experimental interventions in the follow-up survey when significant effects at conventional levels could be found. Table 8 depicts the heterogeneous effects for the follow-up question on tax morale. While Column (1) does not condition on covariates, Column (2) includes a full-set of covariates covering information on demographics, income as well as political preferences. In line with the null results on the average treatment effects of the *Uncertainty* condition on the elicited tax morale in the follow-up, I do not find significant sub-group effects for this treatment. There is no particular sub-group of the sample where lasting effects from the main survey carry over to the follow-up questionnaire. However, for the augmented *Unequal knowledge* condition, very similar effects emerge in wave 41 for left-wing respondents when compared to the baseline effects from wave 40. Respondents with left-wing political ideology can also explain the negative baseline effects for that experimental intervention in the follow-up survey. Negative effects of the augmented treatment on the elicited tax morale for left-wing subjects (as compared to conservatives) are statistically significant at the 1% level without conditioning on additional covariates. The results are very similar when including a full-set

---

<sup>36</sup>Different treatment effects in the follow-up for both treatment conditions may also be due to a slightly different question wording which now asks how justifiable tax evasion is when an opportunity presents itself (rather than the justifiability in a general sense as in wave 40). Respondents may have the feeling that these opportunities can be mostly associated with people that have tax knowledge (or access to it), partly explaining the negative effect of the *Unequal knowledge* condition.

of controls but are somewhat less precise and significant at the 5% level.

Persistent effects of the *Unequal knowledge* treatment for this sub-group suggest that respondents with a left-wing ideology identify themselves with and empathize with disadvantaged uncertain taxpayers and thus internalize the perception of prevalent uncertainties among peers in their own views on tax morale (for the relevance of identification within groups for the norm of tax compliance, see also Wenzel (2004)).

The results for the follow-up are remarkable for several reasons. First, as argued above, tax morale is rather stable over time (even despite slightly different question wordings). Second, the main experiment only gave short statements about certain aspects of the public perception of the tax system but the overall extent of the intervention was small and subtle. Third, the follow-up estimates should be interpreted as being conservative because the follow-up was fielded two months after the main experiment, which is a substantial amount of time as compared to other follow-up surveys in the literature. Typically, follow-ups are implemented shortly after the main intervention (e.g. one week, for instance, in Alesina et al. (2018); Haaland and Roth (2019, 2020)).

## 4 Conclusion

Taxes are complicated and people also frequently perceive them as such, which may result in uncertainty among many taxpayers how to file taxes themselves but may also create room for negative spillovers on tax compliance. The present paper addresses the latter and answers the question whether perceived taxpayer confusion among others serves as an excuse for tax non-compliance using a novel survey experiment among Germans.

I find supportive evidence for the hypothesis that the intrinsic motivation of tax compliance is undermined when people are informed about the prevalent uncertainty among other taxpayers. Given that a large majority among the German population is in favor of a simpler tax system and about half perceive their tax returns as difficult (see Blesse et al. (2021)), this finding speaks to the fact that people on average do not extrapolate their own difficulties with respect to tax filing onto others but people do empathize with the uncertainty of others in filing taxes once they are aware of it. The treatment effect is largely not different from the *Unequal knowledge* treatment, which augments the uncertainty information with a statement about other taxpayers possessing the tax knowledge (or having access to it through tax advice) to reduce their tax burden.

Negative effects of peer information are driven by a reduction among very honest taxpayers. Older respondents also show more empathy with taxes being evaded when they become aware of the tax uncertainty of others. Similarly, left-wing subjects show a higher tax morale in the control group (as compared to conservatives) but have significantly lower tax morale when receiving the augmented treatment. This effect remains virtually

identical in a follow-up survey two months after the main experiment. It appears that left-wing respondents internalized a more lenient view on tax evasion over time when being informed that many citizens are uncertain on how to file their taxes and, unlike those possessing knowledge about complex taxes, cannot use their tax declaration to reduce their tax burden. I also find significant negative effects of both treatment conditions among very high tax morale levels in the follow-up but it appears that resulting changes from peer information in judgments on tax evasion grow somewhat less strong over time.

Hence, by randomly providing respondents with information about tax uncertainty among others, I find evidence of a potentially long-lasting erosion of tax morale among very high levels of tax morale as well as for left-wing subjects in a high-morale context such as the German income tax. My findings can likely be applied to other developed countries with high levels of tax morale and complicated tax systems (such as the US).

Ultimately, my results raise concerns that complex tax systems may not only have direct detrimental effects on tax compliance through taxpayer confusion but may also undermine tax morale through negative signals of peer compliance, i.e. once one becomes aware of other confused taxpayers. In turn, these negative spillover effects may lead to lower effective fiscal capacities in typically high morale contexts in the long-run (Besley, 2020b) if treatment effects are persistent or if individuals receive regular signals about the negative perceptions of other taxpayers. Taxpayers can (at least to some extent) become aware of related perceptions of others. First, the political debate recurrently involves reform proposals which aim at simplifying the tax schedule (see Blesse et al. (2021)) and which are often motivated by the fact that many taxpayers are confused by the tax schedule. Second, recent evidence of Chetty et al. (2013) indicates that tax knowledge can diffuse locally in networks such as among welfare recipients who need to file similar tax forms or through the presence of tax advisers. One can argue that not only tax knowledge but also related uncertainties can spread among peers although only to a limited extent since information likely diffuses rather within than between groups.

Against this background, policy makers not only have to account for negative direct effects of taxpayer confusion such as low take-up rates of tax benefits (Chetty et al., 2013; Bhargava and Manoli, 2015) or suboptimal behavioral adjustments (e.g. Abeler and Jaeger (2015) and Feldman et al. (2016)) but also do need to consider potential related negative spillover effects on voluntary tax compliance when designing tax rules.

## References

- Abeler, J. and S. Jaeger (2015). Complex tax incentives. *American Economic Journal: Economic Policy* 7(3), 1–28.
- Acemoglu, D. and M. O. Jackson (2017). Social norms and the enforcement of laws. *Journal of the European Economic Association* 15(2), 245–295.
- Aghion, P., U. Akcigit, M. Lequien, and S. Stantcheva (2017). Tax simplicity and heterogeneous learning. NBER Working Paper No. 24049.
- Alesina, A., S. Stantcheva, and E. Teso (2018). Intergenerational mobility and support for redistribution. *American Economic Review* 108(2), 521 – 554.
- Alm, J., B. Jackson, and M. McKee (1992). Institutional uncertainty and taxpayer compliance. *American Economic Review* 82(4), 1018–1026.
- Alm, J., G. H. McClelland, and W. D. Schulze (1992). Why do people pay taxes? *Journal of Public Economics* 48(1), 21–38.
- Alm, J. and B. Torgler (2006). Culture differences and tax morale in the United States and in Europe. *Journal of Economic Psychology* 27(2), 224–246.
- Alstadsæter, A., N. Johannessen, and G. Zucman (2019). Tax evasion and inequality. *American Economic Review* 109(6), 2073–2103.
- Alstadsæter, A., W. Kopczuk, and K. Telle (2019). Social networks and tax avoidance: Evidence from a well-defined norwegian tax shelter. *International Tax and Public Finance* 26(6), 1291–1328.
- Andreoni, J., B. Erard, and J. Feinstein (1998). Tax compliance. *Journal of economic literature* 36(2), 818–860.
- Ballard, C. L. and S. Gupta (2018). Perceptions and realities of average tax rates in the federal income tax: Evidence from Michigan. *National Tax Journal* 71(2), 263–294.
- Bellemare, C., M. Deversi, and F. Englmaier (2019). Complexity and distributive fairness interact in affecting compliance behavior. CESifo Working Paper No. 7899.
- Benabou, R. and J. Tirole (2011). Laws and norms. NBER Working Paper No. 17579.
- Benzarti, Y. (2020). How taxing is tax filing? using revealed preferences to estimate compliance costs. *American Economic Journal: Economic Policy forthcoming*.
- Besley, T. (2020a). Reply to: Comments on a state capacity, reciprocity, and the social contract. *Econometrica* 88(4), 1359–1362.
- Besley, T. (2020b). State capacity, reciprocity and the social contract. *Econometrica* 88(4), 1307–1335.

- Besley, T., A. Jensen, and T. Persson (2019). Norms, enforcement, and tax evasion. NBER Working Paper No. 25575.
- Bhargava, S. and D. 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.
- Blesse, S., F. Buhlmann, and P. Dörrenberg (2021). Attitudes towards tax simplification: Evidence from a representative survey and experiments. [https://www.bwl.uni-mannheim.de/media/Lehrstuehle/bwl/Doerrenberg/tax\\_complexity\\_20210108.pdf](https://www.bwl.uni-mannheim.de/media/Lehrstuehle/bwl/Doerrenberg/tax_complexity_20210108.pdf). Online; accessed 18 January 2021.
- Blesse, S. and F. Heinemann (2020). Citizensâ trade-offs in state merger decisions: Evidence from a randomized survey experiment. *Journal of Economic Behavior & Organization* 180, 438–471.
- Blom, A. G., M. Bosnjak, A. Cornilleau, A.-S. Cousteaux, M. Das, S. Douhou, and U. Krieger (2016). A comparison of four probability-based online and mixed-mode panels in Europe. *Social Science Computer Review* 34(1), 8–25.
- Blom, A. G., C. Gathmann, and U. Krieger (2015). Setting up an online panel representative of the general population: The German Internet Panel. *Field methods* 27(4), 391–408.
- Blom, A. G., J. M. Herzing, C. Cornesse, J. W. Sakshaug, U. Krieger, and D. Bossert (2017). Does the recruitment of offline households increase the sample representativeness of probability-based online panels? Evidence from the German Internet Panel. *Social Science Computer Review* 35(4), 498–520.
- Bott, K. M., A. W. Cappelen, E. Ø. Sørensen, and B. Tungodden (2020). Youâve got mail: A randomized field experiment on tax evasion. *Management Science* 66(7), 2801–2819.
- Bursztyn, L., G. Egorov, and S. Fiorin (2020). From extreme to mainstream: The erosion of social norms. *American Economic Review* 110(11), 3522–48.
- Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020). Misperceived social norms: Women working outside the home in saudi arabia. *American economic review* 110(10), 2997–3029.
- Bursztyn, L., I. K. Haaland, A. Rao, and C. P. Roth (2020). I have nothing against them, but... NBER Working Paper No. 27288.
- Castro, L. and C. Scartascini (2015). Tax compliance and enforcement in the pam-pas evidence from a field experiment. *Journal of Economic Behavior & Organization* 116, 65–82.

- Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernandez-Val (2018). Generic machine learning inference on heterogenous treatment effects in randomized experiments. NBER Working Paper No. 24678.
- Chetty, R., J. N. Friedman, and E. Saez (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *American Economic Review* 103(7), 2683–2721.
- Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *American Economic Review* 99(4), 1145–77.
- Chetty, R. and E. Saez (2013). Teaching the tax code: Earnings responses to an experiment with EITC recipients. *American Economic Journal: Applied Economics* 5(1), 1–31.
- Coibion, O., Y. Gorodnichenko, S. Kumar, and J. Ryngaert (2018). Do you know that i know that you know...? Higher-order beliefs in survey data. NBER Working Paper No. 24987.
- Cummings, R. G., J. Martinez-Vazquez, M. McKee, and B. Torgler (2009). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior & Organization* 70(3), 447–457.
- Dana, J., R. A. Weber, and J. X. Kuang (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory* 33(1), 67–80.
- De Bartolome, C. (1995). Which tax rate do people use: Average or marginal? *Journal of Public Economics* 56(1), 79–96.
- De Neve, J.-E., C. Imbert, J. Spinnewijn, T. Tsankova, and M. Luts (2021). How to improve tax compliance? evidence from population-wide experiments in belgium. *Journal of Political Economy* 129(5), 000–000.
- Del Carpio, L. (2014). Are the neighbors cheating? Evidence from a social norm experiment on property taxes in Peru. mimeo.
- Dörrenberg, P. and A. Peichl (2013). Progressive taxation and tax morale. *Public Choice* 155(3-4), 293–316.
- Dörrenberg, P. and A. Peichl (2018). Tax morale and the role of social norms and reciprocity. Evidence from a randomized survey experiment. *CESifo Working Paper No. 7149*.
- Drago, F., F. Mengel, and C. Traxler (2020). Compliance behavior in networks: Evidence from a field experiment. *American Economic Journal: Applied Economics* 12(2), 96–133.

- Duflo, E., A. Banerjee, A. Finkelstein, L. F. Katz, B. A. Olken, and A. Sautmann (2020). In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in economics. NBER Working Paper No. 26993.
- Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2016). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. *American Economic Journal: Economic Policy* 8(3), 203–32.
- Dwenger, N. and L. Treber (2018). Shaming for tax enforcement: Evidence from a new policy. *CEPR Discussion Paper No. 13194*.
- Eerola, E., T. Kosonen, K. Kotakorpi, T. Lyytikäinen, and J. Tuimala (2019). Tax compliance in the rental housing market: Evidence from a field experiment. *VATT Institute for Economic Research Working Papers* 122.
- Einav, L., A. Finkelstein, I. Pascu, and M. R. Cullen (2012). How general are risk preferences? Choices under uncertainty in different domains. *American Economic Review* 102(6), 2606–38.
- Elffers, H., R. H. Weigel, and D. J. Hessing (1987). The consequences of different strategies for measuring tax evasion behavior. *Journal of Economic Psychology* 8(3), 311–337.
- Engelmann, D., E. Janeba, L. Mechtenberg, and N. Wehrhoefer (2018). Preferences over taxation of high-income individuals: Evidence from a survey experiment. Technical report.
- Eriksen, K. and L. Fallan (1996). Tax knowledge and attitudes towards taxation; A report on a quasi-experiment. *Journal of Economic Psychology* 17(3), 387–402.
- Fehr, D., D. Muller, and M. Preuss (2020). Social mobility perceptions and inequality acceptance. *University of Innsbruck Working Papers in Economics and Statistics No. 02*.
- Feld, L. P. and B. Torgler (2007). Tax morale after the reunification of Germany: Results from a quasi-natural experiment. *CESifo Working Paper No. 1921*.
- Feldman, N. E., P. Katuščák, and L. Kawano (2016). Taxpayer confusion: Evidence from the child tax credit. *American Economic Review* 106(3), 807–35.
- Fellner, G., R. Sausgruber, and C. Traxler (2013). Testing enforcement strategies in the field: Threat, moral appeal and social information. *Journal of the European Economic Association* 11(3), 634–660.
- Frey, B. S. and B. Torgler (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics* 35(1), 136–159.
- Fromm, F. (2019). Leichtfertige Steuerverkürzung: 6 Hinweise vom Fachanwalt. <https://www.frankfromm.de/leichtfertige-steuerverkuerzung/>.

- Gideon, M. (2017). Do individuals perceive income tax rates correctly? *Public Finance Review* 45(1), 97–117.
- Gsottbauer, E., D. Müller, S. Müller, S. T. Trautmann, and G. Zudenkova (2020). Social class and (un)ethical behavior: Causal versus correlational evidence. *University of Innsbruck Working Papers in Economics and Statistics No. 10*.
- Haaland, I. and C. Roth (2019). Beliefs about racial discrimination and support for pro-black policies. CESifo Working Paper No. 7828.
- Haaland, I. and C. Roth (2020). Labor market concerns and support for immigration. *Journal of Public Economics* 191, 104256.
- Haaland, I., C. Roth, and J. Wohlfart (2020). Designing information provision experiments. CAGE Working Paper No. 484.
- Halla, M. (2012). Tax morale and compliance behavior: First evidence on a causal link. *B.E. Journal of Economic Analysis & Policy* 12(1).
- Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics* 148(April), 14–31.
- Hashimzade, N., G. D. Myles, and B. Tran-Nam (2013). Applications of behavioural economics to tax evasion. *Journal of Economic Surveys* 27(5), 941–977.
- Helliwell, J. F., S. Wang, and J. Xu (2016). How durable are social norms? Immigrant trust and generosity in 132 countries. *Social Indicators Research* 128(1), 201–219.
- Hoopes, J. L., D. H. Reck, and J. Slemrod (2015). Taxpayer search for information: Implications for rational attention. *American Economic Journal: Economic Policy* 7(3), 177–208.
- Kerschbamer, R. and D. Müller (2020). Social preferences and political attitudes: An online experiment on a large heterogeneous sample. *Journal of Public Economics* 182(February), 104076.
- Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica* 79(3), 651–692.
- Kleven, H. J. and W. Kopczuk (2011). Transfer program complexity and the take-up of social benefits. *American Economic Journal: Economic Policy* 3(1), 54–90.
- Konrad, K. A. and S. Qari (2012). The last refuge of a scoundrel? Patriotism and tax compliance. *Economica* 79(315), 516–533.
- Krause, K. (2000). Tax complexity: Problem or opportunity? *Public Finance Review* 28(5), 395–414.

- Lago-Peñas, I. and S. Lago-Peñas (2010). The determinants of tax morale in comparative perspective: Evidence from European countries. *European Journal of Political economy* 26(4), 441–453.
- Lazear, E. P., U. Malmendier, and R. A. Weber (2012). Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics* 4(1), 136–63.
- Lewis, A. (1982). *The psychology of taxation*. Martin Roberson: Oxford.
- Liebman, J. B. and R. J. Zeckhauser (2004). Schmeduling. mimeo.
- Luttmer, E. F. and M. Singhal (2014). Tax morale. *Journal of Economic Perspectives* 28(4), 149–68.
- Möhlmann, A. (2014). Persistence or convergence? The East–West tax-morale gap in Germany. *FinanzArchiv/Public Finance Analysis* 70(1), 3–30.
- Müller, D. and S. Renes (2020). Fairness views and political preferences: evidence from a large and heterogeneous sample. *Social Choice and Welfare*, 1–33.
- Ortega, D., L. Ronconi, and P. Sanguinetti (2016). Reciprocity and willingness to pay taxes: Evidence from a survey experiment in Latin America. *Economía* 16(2), 55–87.
- Paetzold, J. and H. Winner (2016). Taking the high road? Compliance with commuter tax allowances and the role of evasion spillovers. *Journal of Public Economics* 143(November), 1 – 14.
- Perez-Truglia, R. and U. Troiano (2018). Shaming tax delinquents. *Journal of Public Economics* 167(November), 120–137.
- Rees-Jones, A. and D. Taubinsky (2020). Measuring "schmeduling". *The Review of Economic Studies* 87(5), 2399–2438.
- Richardson, G. (2006). Determinants of tax evasion: A cross-country investigation. *Journal of international Accounting, Auditing and taxation* 15(2), 150–169.
- Sjoberg, F. M., J. Mellon, T. C. Peixoto, J. Z. Hemker, and L. L. Tsai (2019). Voice and punishment: A global survey experiment on tax morale. *World Bank Policy Research Working Paper No. 8855*.
- Slemrod, J. (1989). The return to tax simplification: An econometric analysis. *Public Finance Quarterly* 17(1), 3–27.
- Slemrod, J., M. Blumenthal, and C. Christian (2001). Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *National Tax Journal* 54(1), 125 – 138.

- Stantcheva, S. (2020). Understanding tax policy: How do people reason? Technical report, National Bureau of Economic Research.
- Torgler, B. (2002). Does culture matter? Tax morale in an East–West–German comparison. *FinanzArchiv/Public Finance Analysis* 59(4), 504–528.
- Torgler, B. (2005). Tax morale in Latin America. *Public Choice* 122(1-2), 133–157.
- Torgler, B., I. C. Demir, A. Macintyre, and M. Schaffner (2008). Causes and consequences of tax morale: An empirical investigation. *Economic Analysis and Policy* 38(2), 313–339.
- Traxler, C. (2010). Social norms and conditional cooperative taxpayers. *European Journal of Political Economy* 26(1), 89–103.
- Ubfal, D. (2016). How general are time preferences? Eliciting good-specific discount rates. *Journal of Development Economics* 118(C), 150–170.
- Wager, S. and S. Athey (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association* 113(523), 1228–1242.
- Wenzel, M. (2004). An analysis of norm processes in tax compliance. *Journal of Economic Psychology* 25(2), 213–228.
- World Value Survey (2020). World Values Survey Wave 7 (2017-2020): WVS Results By Country 2017-2020 v1.4.

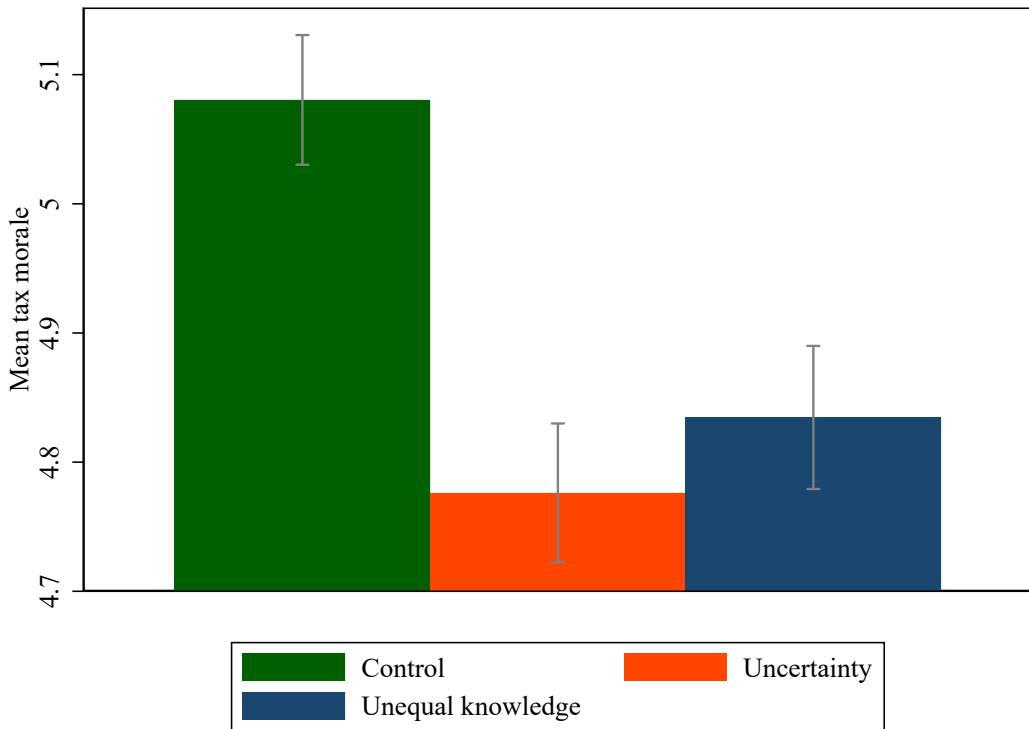
## 5 Main figures and tables

Table 1: Summary statistics

	N	Mean	Std.Dev.	Min	Max
<b><i>Experimental intervention</i></b>					
Control	4890	0.33	0.47	0	1
Uncertainty	4890	0.33	0.47	0	1
Unequal knowledge	4890	0.33	0.47	0	1
<b><i>Demographics</i></b>					
Single households	4876	0.17	0.38	0	1
2	4876	0.43	0.50	0	1
3	4876	0.18	0.39	0	1
4	4876	0.16	0.36	0	1
5+	4876	0.05	0.23	0	1
Age <=28	4887	0.14	0.35	0	1
Age 29-38	4887	0.15	0.36	0	1
Age 39-48	4887	0.16	0.36	0	1
Age 49-58	4887	0.24	0.43	0	1
Age >= 59	4887	0.31	0.46	0	1
Married	4890	0.58	0.49	0	1
Female	4888	0.48	0.50	0	1
Full time-employed	4866	0.46	0.50	0	1
Retired	4866	0.19	0.40	0	1
Low education	4746	0.04	0.20	0	1
Low-med education	4746	0.42	0.49	0	1
High-med education	4746	0.24	0.42	0	1
High education	4746	0.31	0.46	0	1
<b><i>Personal net income</i></b>					
Poor	4890	0.30	0.46	0	1
2	4890	0.27	0.44	0	1
3	4890	0.14	0.35	0	1
4	4890	0.06	0.23	0	1
Rich	4890	0.05	0.21	0	1
No income stated	4890	0.13	0.34	0	1
Not merged	4890	0.06	0.23	0	1
<b><i>Political orientation</i></b>					
Conservatives	4890	0.40	0.49	0	1
Left-wing	4890	0.44	0.50	0	1
Non partisans	4890	0.11	0.31	0	1
Not merged	4890	0.05	0.23	0	1

Notes: The table depicts the summary statistics for all treatment group dummies and all covariates used. I define variables as follows, Control, *Uncertainty* and *Unequal Knowledge* group realizations represent the respective allocations of respondents to either group; household size comprises single households and household with 2, 3, 4 and 5+ members; age categories are  $\leq 28$ , 29-38, 39-48, 49-55 and  $\geq 59$ ; Married equals 1 if respondent is married, 0 otherwise; Female equals 1 if respondent is female, 0 otherwise; Full time equals 1 if respondent is full time employed, 0 otherwise; Retired equals 1 if respondent is retired, 0 otherwise; education categories comprise low (secondary schooling, no job training), low to medium education (upper secondary schooling or finished job training), high to medium education (upper secondary schooling and finished job training) and high education (tertiary education); household income variables define net monthly household incomes on a 5-point scale from poor, i.e. 1 ( $\leq 1500$  Euro), 2 ( $1500 \leq x < 2500$  Euro), 3 ( $2500 \leq x < 3500$  Euro), 4 ( $3500 \leq x < 4500$  Euro) to 5 being rich ( $\geq 4500$  Euro) as well as a dummy for no answers (No income stated) and a dummy for those observations which had not been in wave 37 of the GIP where the income question was asked; conservatives equals 1 if  $> 5$  on a 11-scale left-right placement variable, for  $\leq 5$  left-wing equals 1. Non partisans did not report a score for the left-right placement variable. I include a dummy for observations which had not been in the GIP wave where the political preference question was asked. Data comes from the German Internet Panel (GIP) wave 40, except for political preferences as well as household incomes (wave 37).

Figure 1: Tax morale by experimental group



Notes: Average tax morale by experimental group with 95% confidence bars. The outcome variable is the survey question about tax morale preferences as described in Section 2.1. Treatment groups as described in Section 2.2. Total number of observations is 4,809 with even distribution across experimental groups. Data come from GIP wave 40 (Blom et al., 2015).

Table 2: Effect of experimental intervention on tax morale, Baseline results

	(1)	(2)	(3)	(4)
Uncertainty	-0.304*** (0.038)	-0.300*** (0.038)	-0.299*** (0.038)	-0.300*** (0.038)
Unequal knowledge	-0.246*** (0.038)	-0.240*** (0.038)	-0.242*** (0.038)	-0.242*** (0.038)
Constant	5.080*** (0.026)	4.494*** (0.102)	4.420*** (0.109)	4.358*** (0.112)
test (p-val T1 vs. T2)	0.138	0.132	0.153	0.143
N	4809	4635	4635	4635
r2	0.015	0.037	0.039	0.043
Demographics	No	Yes	Yes	Yes
Household Income	No	No	Yes	Yes
Political Preference	No	No	No	Yes

Notes: The table presents the effects of the randomized treatment interventions on tax morale preferences. This is estimated by OLS regressions of tax morale preferences on treatment dummies. Tax morale is measured on a 6 point scale based on the question, “How justifiable do you think it is to evade taxes?”. Answer categories range from absolutely justifiable (1) to absolutely not justifiable (6). The experimental groups are, Control group, *Uncertainty* and *Unequal Knowledge* group. Control is omitted, implying that the effects are relative to the Control Group. All participants receive the following information, “There are frequent reports on tax evasion in the media.” Participants in the *Uncertainty* group receive the following information, “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration.” Participants in the *Unequal Knowledge* group receive the following information, “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration. Citizens who know tax laws well or use a tax attorney can possibly take advantage of the complexities of the tax law and reduce their tax burden.” The line p-val T1 vs. T2 presents the p-values from t-tests which compare if the regression coefficient for the *Uncertainty* group is different from the regression coefficient for the *Unequal Knowledge* group. Columns (1)-(4) differ in the included sets of covariates. (1): no covariates, (2): gender, age, marital status, household size, employment status, retirement status, and education, (3): (2) plus net household income, (4): (3) plus political preferences. The scale of the outcome variable is 1 (absolutely justifiable) to 6 (absolutely not justifiable). Robust standard errors are in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 3: Anatomy of tax morale

	(1)	(2)	(3)
Household size	0.015 (0.018)	0.005 (0.019)	0.005 (0.018)
Age	0.073*** (0.016)	0.072*** (0.016)	0.072*** (0.016)
Married	0.079* (0.040)	0.070* (0.041)	0.076* (0.041)
Female	0.110*** (0.035)	0.104*** (0.035)	0.097*** (0.035)
Full time-employed	-0.033 (0.040)	-0.050 (0.041)	-0.041 (0.041)
Retired	0.058 (0.052)	0.061 (0.052)	0.061 (0.052)
Education	0.072*** (0.017)	0.066*** (0.018)	0.062*** (0.018)
Income cat 2		0.121* (0.072)	0.121* (0.072)
3		0.115 (0.074)	0.116 (0.074)
4		0.114 (0.077)	0.113 (0.077)
Rich		0.173** (0.078)	0.177** (0.077)
No income stated		0.275*** (0.105)	0.288*** (0.105)
Not merged		0.207*** (0.073)	0.225*** (0.076)
Left-wing			0.133*** (0.034)
Non partisans			0.049 (0.059)
Not merged			0.014 (0.086)
Uncertainty	-0.300*** (0.038)	-0.299*** (0.038)	-0.300*** (0.038)
Unequal knowledge	-0.240*** (0.038)	-0.242*** (0.038)	-0.242*** (0.038)
Constant	4.494*** (0.102)	4.420*** (0.109)	4.358*** (0.112)
N	4635	4635	4635
r2	0.037	0.039	0.043

Notes: The table presents the determinants of tax morale from wave 40 of the GIP using OLS regressions of tax morale on various covariates. Tax morale is measured on a 6 point scale based on the question, "How justifiable do you think it is to evade taxes?". Answer categories range from absolutely justifiable (1) to absolutely not justifiable (6). Each column (1)-(3) presents the results of one regression with different sets of covariates. All columns always include dummies for participation in either of the two treatment groups. Robust standard errors are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Variables are defined as follows, *Uncertainty* and *Unequal knowledge* realizations represent the respective allocations of respondents to either group; household size comprises single households and household with 2, 3, 4 and 5+ members; age categories are  $\leq 28$ , 29-38, 39-48, 49-55 and  $\geq 59$ ; Married equals 1 if respondent is married, 0 otherwise; Female equals 1 if respondent is female, 0 otherwise; Full time equals 1 if respondent is full time employed, 0 otherwise; Retired equals 1 if respondent is retired, 0 otherwise; education categories comprise low (secondary schooling, no job training), low to medium education (upper secondary schooling or finished job training), high to medium education (upper secondary schooling and finished job training) and high eduction (tertiary education); household income variables define net monthly personal incomes on a 5-point scale from poor, i.e. 1 ( $\leq 1500$  Euro), 2 ( $1500 \geq x < 2500$  Euro), 3 ( $2500 \geq x < 3500$  Euro), 4 ( $3500 \geq x < 4500$  Euro) to 5 being rich ( $\geq 4500$  Euro) as well as a dummy for no answers (No income stated) and a dummy for those observations which had not been in wave 37 of the GIP where the income question was asked; conservatives equals 1 if >5 on a 11-scale left-right placement variable, for  $\leq 5$  left-wing equals 1. Non partisans did not report a score for the left-right placement variable. I include a dummy for observations which had not been in the GIP wave where the political preference question was asked. Data comes from the German Internet Panel (GIP) wave 40, except for political preferences as well as household incomes (wave 37).

Table 4: Effect of treatments on the distribution of tax morale preferences, LPM models

	(1)	(2)	(3)	(4)	(5)	(6)
Uncertainty	0.004 (0.004)	0.008 (0.005)	0.029* ** (0.009)	0.078* ** (0.014)	0.008 (0.017)	-0.126* ** (0.017)
Unequal knowledge	0.003 (0.004)	0.019* ** (0.006)	0.013 (0.008)	0.065* ** (0.014)	-0.019 (0.017)	-0.081* ** (0.017)
Constant	0.045* ** (0.014)	0.057* ** (0.017)	0.165* ** (0.026)	0.225* ** (0.039)	0.243* ** (0.044)	0.265* ** (0.044)
test (p-val T1 vs. T2)	0.835	0.068	0.091	0.403	0.112	0.007
N	4635	4635	4635	4635	4635	4635
r2	0.008	0.008	0.024	0.013	0.012	0.019
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Household Income	Yes	Yes	Yes	Yes	Yes	Yes
Political Preference	Yes	Yes	Yes	Yes	Yes	Yes

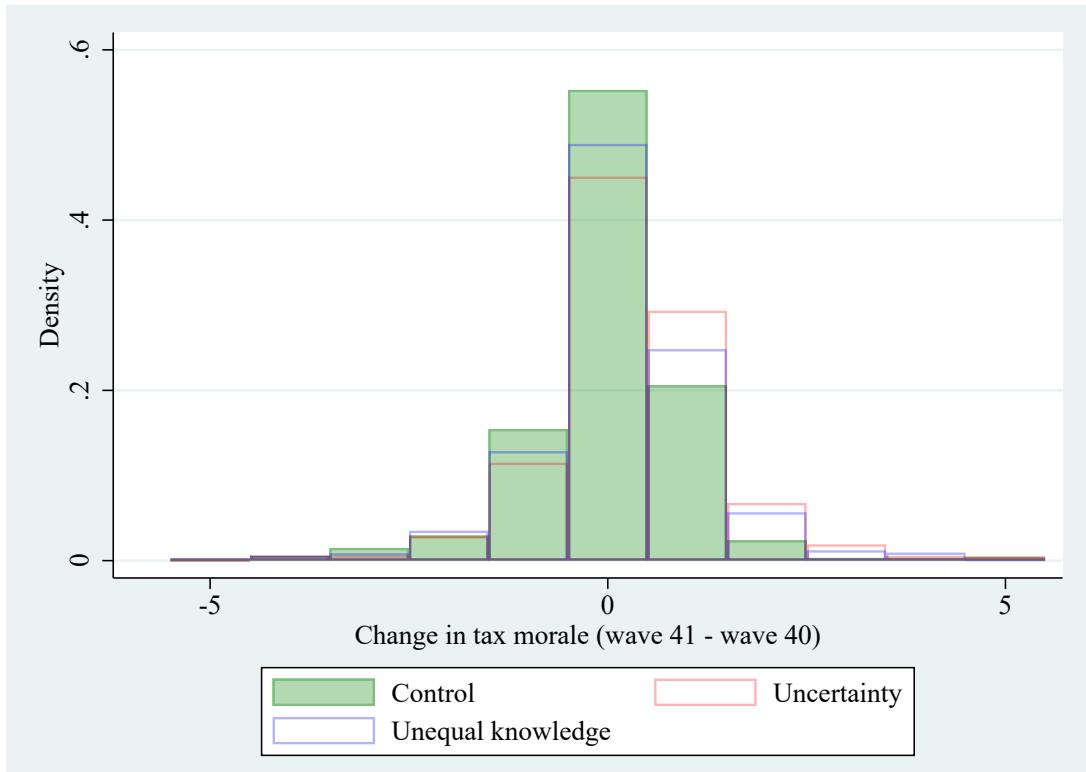
Notes: The table presents the effects of the randomized treatment interventions on the distribution of tax morale preferences. I estimate Linear Probability Model (LPM) estimations of tax morale preferences on treatment dummies. Tax morale dummy variables are derived from a 6 point scale question, “How justifiable do you think it is to evade taxes?” in wave 40 of the GIP. Columns (1)-(6) differ in their outcome variables, namely the response categories to the tax morale question, in Column (1) the outcome equals 1 if the respondent answered that he/she finds tax evasion absolutely justifiable, 0 otherwise; in Column (2), the outcome equals 1 if the respondent answered that he/she finds tax evasion justifiable, 0 otherwise; in Column (3) the outcome equals 1 if the respondent answered that he/she finds tax evasion somewhat justifiable, 0 otherwise; in Column (4) the outcome equals 1 if the respondent answered that he/she finds tax evasion somewhat not justifiable, 0 otherwise; in Column (5) the outcome equals 1 if the respondent answered that he/she finds tax evasion not justifiable, 0 otherwise and in Column (6) the outcome equals 1 if the respondent answered that he/she finds tax evasion absolutely not justifiable, 0 otherwise. Columns (1)-(6) all control for a full set of covariates, including gender, age, marital status, household size, employment status, retirement status, and education, net personal income and political preferences. The experimental groups are Control group, *Uncertainty* and *Unequal Knowledge* group. Control is omitted, implying that the effects are relative to the Control Group. All participants receive the following information, “There are frequent reports on tax evasion in the media.” Participants in the *Uncertainty* group receive the following information, “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration.” Participants in the *Unequal Knowledge* group receive the following information, “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration. Citizens who know tax laws well or use a tax attorney can possibly take advantage of the complexities of the tax law and reduce their tax burden.” The line p-val T1 vs. T2 presents the p-values from t-tests which compare if the regression coefficient for the *Uncertainty* group is different from the regression coefficient for the *Unequal Knowledge* group. Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 5: Heterogeneity of average treatment effects

	(1)	(2)	(3)	(4)
<b>Political orientation. Reference category: Conservative</b>				
Uncertainty × Left-wing	-0.130 (0.081)	-0.126 (0.082)		
Uncertainty × Non partisans	0.071 (0.138)	0.027 (0.136)		
Uncertainty × Not merged	0.121 (0.177)	0.137 (0.185)		
Unequal knowledge × Left-wing	-0.195** (0.082)	-0.210** (0.082)		
Unequal knowledge × Non-partisans	-0.133 (0.139)	-0.236* (0.136)		
Unequal knowledge × Not merged	0.161 (0.185)	0.136 (0.198)		
<b>Age. Reference category = &lt; 28</b>				
Uncertainty × = < 38		-0.094 (0.152)	-0.120 (0.153)	
Uncertainty × = < 48		-0.211 (0.153)	-0.267 * (0.153)	
Uncertainty × = < 58		-0.338** (0.140)	-0.338 ** (0.141)	
Uncertainty × > 58		-0.348*** (0.134)	-0.354 *** (0.135)	
Unequal knowledge × = < 38		-0.065 (0.160)	-0.032 (0.162)	
Unequal knowledge × = < 48		-0.017 (0.155)	-0.022 (0.155)	
Unequal knowledge × = < 58		-0.081 (0.143)	-0.012 (0.144)	
Unequal knowledge × > 58		-0.125 (0.137)	-0.084 (0.139)	
N	4809	4635	4806	4635
r2	0.020	0.045	0.031	0.046
Full-set of controls	No	Yes	No	Yes

Notes: The table shows the heterogeneous effects of the experimental interventions. Reported are the coefficients and standard errors (in parentheses) from a series of OLS regressions of the form  $y_i = \beta_1 Treat_i + \beta_2 Covariate_i + \beta_3 (Treat_i \times Covariate_i) + \epsilon_i$ . Where  $Covariate_i$  is the respective covariate listed above. For the sake of brevity, only the interaction  $Treat_i \times Covariate_i$  is reported.  $y_i$  represents tax morale from wave 40 of the GIP and is measured on a 6 point scale based on the question, "How justifiable do you think it is to evade taxes?". Answer categories range from absolutely justifiable (1) to absolutely not justifiable (6).  $Treat_i$  represents treatment indicators for either the *Uncertainty* or the *Unequal Knowledge* group, the control group is omitted.  $(Treat_i \times Covariate_i)$  is a full interaction of the treatment indicators with the respective covariate. Specifications (1) to (4) represent heterogeneous effects for different covariates. Column (1)-(2) reports interaction effects of both treatment indicators with measures of political orientation. Column (3)-(4) reports interaction effects of both treatment indicators with age categories. Column (1) and (3) do not include covariates while Column (2) and (4) include a full-set of covariates covering all demographic, income and political preferences used in previous regressions. For the sake of brevity, no heterogeneous effects are listed for which I did not find significant interaction effects (available upon request). Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Figure 2: The persistence of tax morale



Notes: The figure depicts the distribution in change of tax morale responses between the follow-up question in wave 41 and the main experiment in wave 40 for control group respondents. Changes in tax morale are measured only for respondents who answer the tax morale question in both survey waves. Total number of observations is 1,461. Data come from GIP wave 40 and 41 (Blom et al., 2015).

Table 6: Effect of experimental interventions on tax morale in follow-up survey

	(1)	(2)	(3)	(4)
Uncertainty	-0.014 (0.042)	-0.007 (0.042)	-0.005 (0.042)	-0.008 (0.042)
Unequal knowledge	-0.086** (0.043)	-0.066 (0.042)	-0.066 (0.042)	-0.066 (0.042)
Constant	5.107* ** (0.030)	4.400* ** (0.115)	4.403* ** (0.117)	4.309* ** (0.119)
test (p-val T1 vs. T2)	0.086	0.157	0.150	0.174
N	4405	4253	4253	4253
r2	0.001	0.040	0.041	0.048
Demographics	No	Yes	Yes	Yes
Household Income	No	No	Yes	Yes
Political Preference	No	No	No	Yes

Notes: The table presents the effects of the randomized treatment interventions on tax morale preferences in the follow-up wave 41 of the GIP. This is estimated by OLS regressions of tax morale preferences on treatment dummies. I restrict the number of observations to respondents who underwent the experiment in wave 40. Tax morale is measured on a 6 point scale based on the question, “How justifiable do you think it is to evade taxes if a good opportunity to do so presents itself?”. Answer categories range from absolutely justifiable (1) to absolutely not justifiable (6). The experimental groups from wave 40 are: Control group, *Uncertainty* and *Unequal Knowledge* group. Control is omitted, implying that the effects are relative to the Control Group. All participants receive the following information: “There are frequent reports on tax evasion in the media.” The line p-val T1 vs. T2 presents the p-values from t-tests which compare if the regression coefficient for the *Uncertainty* group is different from the regression coefficient for the *Unequal Knowledge* group. Columns (1)-(4) differ in the included sets of covariates. (1): no covariates, (2): gender, age, marital status, household size, employment status, retirement status, and education, (3): (2) plus net personal income, (4): (3) plus political preferences. Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 7: Change of tax morale answer categories in follow-up survey

	(1)	(2)	(3)	(4)	(5)	(6)
Uncertainty	-0.000 (0.004)	0.001 (0.007)	-0.013 (0.009)	0.001 (0.013)	0.042** (0.017)	-0.031* (0.019)
Unequal knowledge	0.009** (0.004)	-0.005 (0.006)	-0.005 (0.009)	0.001 (0.013)	0.052*** (0.017)	-0.052*** (0.018)
Constant	0.018 (0.013)	0.116*** (0.022)	0.164*** (0.025)	0.228*** (0.032)	0.188*** (0.043)	0.286*** (0.046)
test (p-val T1 vs. T2)	0.027	0.288	0.347	0.975	0.569	0.243
N	4253	4253	4253	4253	4253	4253
r <sup>2</sup>	0.009	0.019	0.021	0.008	0.009	0.030
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Household Income	Yes	Yes	Yes	Yes	Yes	Yes
Political Preference	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents the effects of the randomized treatment interventions on tax morale preference answer categories in the follow-up wave 41 of the GIP. This is estimated by OLS regressions of tax morale preferences on treatment dummies. I restrict the number of observations to respondents who underwent the experiment in wave 40. I estimate Linear Probability Model (LPM) estimations of tax morale preferences on treatment dummies. Tax morale dummy variables are derived from a 6 point scale question, “How justifiable do you think it is to evade taxes if a good opportunity to do so presents itself?” in wave 41 of the GIP. Columns (1)-(6) differ in their outcome variables, namely the response categories to the tax morale question, in Column (1) the outcome equals 1 if the respondent answered that he/she finds tax evasion absolutely justifiable, 0 otherwise; in Column (2), the outcome equals 1 if the respondent answered that he/she finds tax evasion justifiable, 0 otherwise; in Column (3) the outcome equals 1 if the respondent answered that he/she finds tax evasion somewhat justifiable, 0 otherwise; in Column (4) the outcome equals 1 if the respondent answered that he/she finds tax evasion somewhat not justifiable, 0 otherwise; in Column (5) the outcome equals 1 if the respondent answered that he/she finds tax evasion not justifiable, 0 otherwise and in Column (6) the outcome equals 1 if the respondent answered that he/she finds tax evasion absolutely not justifiable, 0 otherwise. Columns (1)-(6) all control for a full set of covariates, including gender, age, marital status, household size, employment status, retirement status, and education, net personal income and political preferences. The experimental groups are Control group, *Uncertainty* and *Unequal Knowledge* group. Control is omitted, implying that the effects are relative to the Control Group. All participants receive the following information, “There are frequent reports on tax evasion in the media.” Participants in the *Uncertainty* group receive the following information, “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration.” Participants in the *Unequal Knowledge* group receive the following information, “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration. Citizens who know tax laws well or use a tax attorney can possibly take advantage of the complexities of the tax law and reduce their tax burden.” The line p-val T1 vs. T2 presents the p-values from t-tests which compare if the regression coefficient for the *Uncertainty* group is different from the regression coefficient for the *Unequal Knowledge* group. Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 8: Heterogeneous treatment effects in follow-up

	(1)	(2)
<b>Political orientation. Reference category: Conservative</b>		
Uncertainty × Left-wing	-0.120 (0.090)	-0.131 (0.084)
Uncertainty × Non-partisans	-0.033 (0.156)	-0.056 (0.139)
Uncertainty × Not-merged	0.037 (0.203)	0.155 (0.229)
Unequal knowledge × Left-wing	-0.237*** (0.091)	-0.213** (0.085)
Unequal knowledge × Non-partisans	-0.187 (0.156)	-0.276** (0.139)
Unequal knowledge × Not merged	-0.481* (0.246)	0.269 (0.232)
N	4405	4247
r2	0.011	0.045
Full-set of controls	No	Yes

Notes: The table shows the heterogeneous effects of the experimental interventions in the follow-up survey (wave 41 of the GIP). Reported are the coefficients and standard errors (in parentheses) from a series of OLS regressions of the form  $y_i = \beta_1 Treat_i + \beta_2 Covariate_i + \beta_3 (Treat_i \times Covariate_i) + \epsilon_i$ . Where  $Covariate_i$  is the respective covariate listed above. For the sake of brevity, only the interaction  $Treat_i \times Covariate_i$  is reported.  $y_i$  represents tax morale from wave 41 of the GIP and is measured on a 6 point scale based on the question, “How justifiable do you think it is to evade taxes?”. Answer categories range from absolutely justifiable (1) to absolutely not justifiable (6).  $Treat_i$  represents treatment indicators for either the *Uncertainty* or the *Unequal Knowledge* group, the control group is omitted.  $(Treat_i \times Covariate_i)$  is a full interaction of the treatment indicators with the respective covariate. Column (1) and (2) represent heterogeneous effects for political ideology. No other heterogeneities were to be found for other covariates. Column (1) does not condition on covariates while Column (2) includes a full-set of covariates covering all demographic, income and political preferences used in previous regressions. For the sake of brevity, no heterogeneous effects are listed for which I did not find significant interaction effects (available upon request). Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

## 6 Appendix

Figure A.1: Screenshot of the survey experiment: Control group

The screenshot shows a survey interface from the 'Gesellschaft im Wandel' project. At the top left is a logo of four stylized figures. At the top right is a 'Hilfe' (Help) link. The main text reads: 'Über Fälle von Steuerhinterziehung wird immer wieder in den Medien berichtet.' Below this is a bold question: 'Für wie vertretbar halten Sie es, Steuern zu hinterziehen?' A note below the question states: 'Bei dieser Frage können Sie nur eine Antwort geben.' A list of six options follows, each preceded by a radio button:

- für absolut vertretbar
- für vertretbar
- für eher vertretbar
- für eher nicht vertretbar
- für nicht vertretbar
- für absolut nicht vertretbar

At the bottom are two buttons: '< Zurück' (Back) and 'Weiter >' (Next).

Notes: The figure depicts a screenshot of the tax morale question. Specifically, the picture shows the screen which is shown to the control group participants. See Section 2 for a description of the survey and the randomized survey experiment. Source: German Internet Panel (GIP), wave 40, including <http://reforms.uni-mannheim.de/internet-panel/Questionnaires/> and Blom et al. (2015).

Figure A.2: Screenshot of the survey experiment: Uncertainty group

The screenshot shows a survey interface from the 'Gesellschaft im Wandel' project. At the top left is a logo of four stylized figures. At the top right is a 'Hilfe' (Help) link. The main text reads: 'Über Fälle von Steuerhinterziehung wird immer wieder in den Medien berichtet.' Below this is a bold question: 'Für wie vertretbar halten Sie es, Steuern zu hinterziehen?' A note below the question states: 'Bei dieser Frage können Sie nur eine Antwort geben.' A list of six options follows, each preceded by a radio button:

- für absolut vertretbar
- für vertretbar
- für eher vertretbar
- für eher nicht vertretbar
- für nicht vertretbar
- für absolut nicht vertretbar

At the bottom are two buttons: '< Zurück' (Back) and 'Weiter >' (Next).

Notes: The figure depicts a screenshot of the tax morale question. Specifically, the picture shows the screen which is shown to the *Uncertainty* group participants. See Section 2 for a description of the survey and the randomized survey experiment. Source: German Internet Panel (GIP), wave 40, including <http://reforms.uni-mannheim.de/internet-panel/Questionnaires/> and Blom et al. (2015).

Figure A.3: Screenshot of the survey experiment: Unequal knowledge group

Über Fälle von Steuerhinterziehung wird immer wieder in den Medien berichtet.

Steuergesetze sind aufgrund vieler Ausnahmeregelungen und Abzugsmöglichkeiten oftmals kompliziert. Viele Bürger sind sich angesichts der Komplexität des Steuersystems nicht sicher, ob sie alle Einkommen korrekt bei der Steuer angeben. Bürger, die die Steuergesetzgebung gut kennen oder einen Steuerberater in Anspruch nehmen, können die komplizierten Steuergesetze jedoch unter Umständen auch zu ihrem Vorteil nutzen und so ihre Steuerlast minimieren.

**Für wie vertretbar halten Sie es, Steuern zu hinterziehen?**

Bei dieser Frage können Sie nur eine Antwort geben.

für absolut vertretbar  
 für vertretbar  
 für eher vertretbar  
 für eher nicht vertretbar  
 für nicht vertretbar  
 für absolut nicht vertretbar

< Zurück      Weiter >



Notes: Notes: The figure depicts a screenshot of the tax morale question. Specifically, the picture shows the screen which is shown to the *Unequal knowledge* group participants. See Section 2 for a description of the survey and the randomized survey experiment. Source: German Internet Panel (GIP), wave 40, including <http://reforms.uni-mannheim.de/internet-panel/Questionnaires/> and Blom et al. (2015).

Table A.1: Balancing tests of experimental groups

Variable	Control	Uncertainty	Unequal knowledge
<b>Gender. Reference category Male</b>			
Sex	-0.006 (0.013)	-0.002 (0.013)	0.004 (0.013)
<b>Marital status: Reference category: Not married</b>			
Married	0.015 (0.014)	-0.033** (0.014)	0.022 (0.014)
<b>Full time employment: Reference category: Not full time employed</b>			
Unemployed	-0.006 (0.014)	0.001 (0.014)	0.003 (0.014)
<b>Retirement status: Reference category: Not retired</b>			
Retired	0.009 (0.017)	-0.001 (0.017)	0.004 (0.017)
<b>Household size: Reference category Single households</b>			
2	0.020 (0.019)	-0.061*** (0.019)	0.042** (0.019)
3	0.004 (0.023)	-0.025 (0.023)	0.016 (0.023)
4	0.026 (0.024)	-0.056 (0.024)	0.027 (0.024)
5+	0.043 (0.033)	-0.054 (0.033)	-0.006 (0.033)
<b>Education: Reference category: Low Education</b>			
2	-0.024 (0.035)	0.005 (0.035)	0.026 (0.035)
3	-0.009 (0.037)	-0.022 (0.037)	0.039 (0.037)
high educ	-0.009 (0.036)	0.000 (0.036)	0.014 (0.036)
<b>Household net income: Reference category poor</b>			
2	0.017 (0.027)	-0.053* (0.027)	0.037 (0.027)
3	-0.004 (0.027)	-0.068** (0.027)	0.079*** (0.027)
4	0.005 (0.028)	-0.031 (0.028)	0.028 (0.028)
rich	0.021 (0.027)	-0.065** (0.027)	0.042 (0.027)
6	0.010 (0.044)	-0.047 (0.044)	0.042 (0.044)
100	0.003 (0.027)	-0.055** (0.027)	0.042 (0.027)
<b>Age category: Reference category: &lt; 34</b>			
34 – 49	-0.012 (0.025)	0.001 (0.025)	0.030 (0.025)

> 49	0.019 (0.025)	0.003 (0.025)	-0.003 (0.025)
<b>Political orientation: Reference category: Conservatives</b>			
Left-wing	-0.009 (0.015)	0.014 (0.015)	-0.009 (0.015)
3	0.008 (0.023)	0.001 (0.023)	-0.009 (0.023)
100	-0.014 (0.031)	0.032 (0.031)	-0.061** (0.031)

Notes: The table shows the coefficients and standard errors (in parentheses) from a series of regressions of the form  $y_i = \beta Covariate_i + \epsilon_i$ . Where *Covariate<sub>i</sub>* is the respective covariate listed above. In Column (1)  $y_i$  equals 1 if participant i is in the control group and 0 otherwise. In Column (2),  $y_i$  equals 1 if participant i is in the *Uncertainty* group and 0 otherwise. In Column (3),  $y_i$  equals 1 if participant i is in the *Unequal Knowledge* group and 0 otherwise. Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.2: Effect of experimental interventions on tax morale, LPM models

	(1)	(2)	(3)	(4)
Uncertainty	-0.045*** (0.010)	-0.044*** (0.010)	-0.044*** (0.010)	-0.044*** (0.010)
Unequal knowledge	-0.039*** (0.010)	-0.039*** (0.010)	-0.039*** (0.010)	-0.039*** (0.010)
Constant	0.928*** (0.006)	0.770*** (0.028)	0.757*** (0.031)	0.742*** (0.031)
test (p-val T1 vs. T2)	0.578	0.620	0.669	0.632
N	4890	4710	4710	4710
r2	0.004	0.028	0.030	0.033
Demographics	No	Yes	Yes	Yes
Household Income	No	No	Yes	Yes
Political Preference	No	No	No	Yes

Notes: The table presents the effects of the randomized treatment interventions on tax morale preferences. I estimate Linear Probability Model (LPM) estimations of tax morale preferences on treatment dummies. Tax morale is coded as 1 if responses are 4,5 or 6 for the question, “How justifiable do you think it is to evade taxes?” which is measured on a 6 point scale, 0 otherwise. The experimental groups are Control group, *Uncertainty* and *Unequal Knowledge* group. Control is omitted, implying that the effects are relative to the Control Group. All participants receive the following information, “There are frequent reports on tax evasion in the media.” Participants in the *Uncertainty* group receive the following information, “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration.” Participants in the *Unequal Knowledge* group receive the following information, “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration. Citizens who know tax laws well or use a tax attorney can possibly take advantage of the complexities of the tax law and reduce their tax burden.” The line p-val T1 vs. T2 presents the p-values from t-tests which compare if the regression coefficient for the *Uncertainty* group is different from the regression coefficient for the *Unequal Knowledge* group. Columns (1)-(4) differ in the included sets of covariates. (1): no covariates, (2): gender, age, marital status, household size, employment status, retirement status, and education, (3): (2) plus net personal income, (4): (3) plus political preferences. The scale of the outcome variable is 1 (absolutely justifiable) to 6 (absolutely not justifiable). Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.3: Effect of experimental interventions on tax morale, Ordered Probit models

	(1)	(2)	(3)	(4)
Uncertainty	-0.324*** (0.038)	-0.325*** (0.039)	-0.324*** (0.039)	-0.327*** (0.039)
Unequal knowledge	-0.252*** (0.039)	-0.251*** (0.040)	-0.253*** (0.040)	-0.254*** (0.040)
test (p-val T1 vs. T2)	0.057	0.054	0.063	0.057
N	4809	4635	4635	4635
Pseudo r2	0.006	0.013	0.014	0.015
Demographics	No	Yes	Yes	Yes
Household Income	No	No	Yes	Yes
Political Preference	No	No	No	Yes

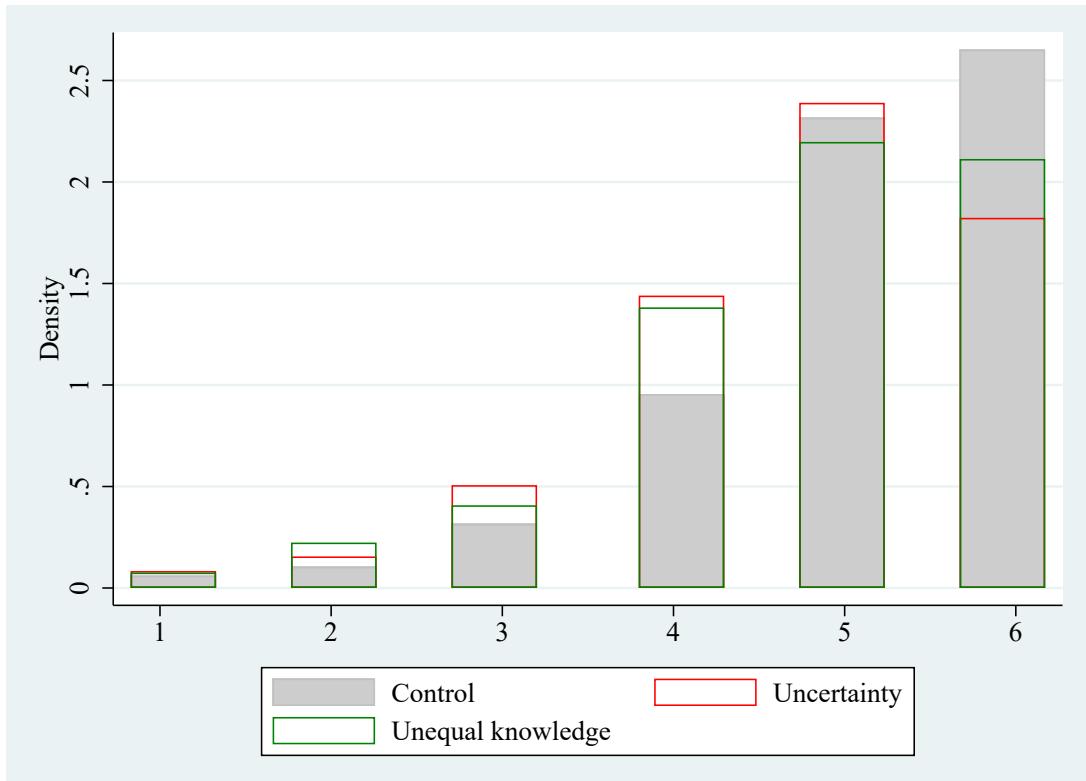
Notes: The table presents the effects of the randomized treatment interventions on tax morale preferences. I estimate ordered probit regressions of tax morale preferences on treatment dummies. Tax morale is measured on a 6 point scale based on the question: "How justifiable do you think it is to evade taxes?". Answer categories range from 1 (absolutely justifiable) to 6 (absolutely not justifiable). The experimental groups are: Control group, *Uncertainty* and *Unequal Knowledge* group. Control is omitted, implying that the effects are relative to the Control Group. All participants receive the following information: "There are frequent reports on tax evasion in the media." Participants in the *Uncertainty* group receive the following information: "Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration." Participants in the *Unequal Knowledge* group receive the following information: "Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration. Citizens who know tax laws well or use a tax attorney can possibly take advantage of the complexities of the tax law and reduce their tax burden." The line p-val T1 vs. T2 presents the p-values from t-tests which compare if the regression coefficient for the *Uncertainty* group is different from the regression coefficient for the *Unequal Knowledge* group. Columns (1)-(4) differ in the included sets of covariates. (1): no covariates, (2): gender, age, marital status, household size, employment status, retirement status, and education, (3): (2) plus net personal income, (4): (3) plus political preferences. The scale of the outcome variable is 1 (absolutely justifiable) to 6 (absolutely not justifiable). Robust standard errors are in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.4: Keeping the numbers of observations constant, Baseline results

	(1)	(2)	(3)	(4)
Uncertainty	-0.309*** (0.038)	-0.300*** (0.038)	-0.299*** (0.038)	-0.300*** (0.038)
Unequal knowledge	-0.246*** (0.039)	-0.240*** (0.038)	-0.242*** (0.038)	-0.242*** (0.038)
Constant	5.086*** (0.026)	4.494*** (0.102)	4.420*** (0.109)	4.358*** (0.112)
test (p-val T1 vs. T2)	0.117	0.132	0.153	0.143
N	4635	4635	4635	4635
r2	0.015	0.037	0.039	0.043
Demographics	No	Yes	Yes	Yes
Household Income	No	No	Yes	Yes
Political Preference	No	No	No	Yes

Notes: The table presents the effects of the randomized treatment interventions on tax morale preferences while keeping the number of observations through specifications (1) to (4) constant. I fix the number of observations at the level of the full-set of covariates of model (4). I estimate the models by OLS regressions of tax morale preferences on treatment dummies. Tax morale is measured on a 6 point scale based on the question: “How justifiable do you think it is to evade taxes?”. Answer categories range from absolutely justifiable (1) to absolutely not justifiable (6). The experimental groups are: Control group, *Uncertainty* and *Unequal Knowledge* group. Control is omitted, implying that the effects are relative to the Control Group. All participants receive the following information: “There are frequent reports on tax evasion in the media.” Participants in the *Uncertainty* group receive the following information: “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration.” Participants in the *Unequal Knowledge* group receive the following information: “Tax laws are often complicated due to many possible deductions and allowances. Given these complexities of the tax system, many citizens are often not sure, whether they report all incomes correctly in their tax declaration. Citizens who know tax laws well or use a tax attorney can possibly take advantage of the complexities of the tax law and reduce their tax burden.” The line p-val T1 vs. T2 presents the p-values from t-tests which compare if the regression coefficient for the *Uncertainty* group is different from the regression coefficient for the *Unequal Knowledge* group. Columns (1)-(4) differ in the included sets of covariates. (1): no covariates, (2): gender, age, marital status, household size, employment status, retirement status, and education, (3): (2) plus net household income, (4): (3) plus political preferences. The scale of the outcome variable is 1 (absolutely justifiable) to 6 (absolutely not justifiable). Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Figure A.4: Tax morale distributions by experimental group



Notes: Tax morale response categories by experimental group. The outcome variable is the survey question about tax morale preferences as described in Section 2.1. Tax morale is measured on a 6 point scale based on the question: "How justifiable do you think it is to evade taxes?" Treatment groups as described in Section 2.2. Total number of observations is 4,809 with even distribution across experimental groups. Data come from GIP wave 40 (Blom et al., 2015).

Table A.5: Testing selective attrition to follow-up survey

	(1)	(2)	(3)	(4)
Uncertainty	-0.001 (0.003)	-0.003 (0.003)	-0.003 (0.003)	-0.003 (0.003)
Unequal knowledge	-0.003 (0.003)	-0.004 (0.003)	-0.004 (0.003)	-0.005 (0.003)
Constant	0.994*** (0.002)	0.990*** (0.006)	0.990*** (0.006)	0.992*** (0.006)
test (p-val T1 vs T2)	0.706	0.725	0.700	0.691
N	4438	4283	4283	4283
r2	0.000	0.004	0.005	0.006
Demographics	No	Yes	Yes	Yes
Household Income	No	No	Yes	Yes
Political Preference	No	No	No	Yes

Notes: The table presents the effects of the randomized treatment interventions on the likelihood to answer the follow-up question in wave 41 of the GIP. This is estimated by OLS regressions of an indicator variable which equals 1 if the respondent has answered that question on treatment dummies and is zero otherwise. I restrict the number of observations to respondents who underwent the experiment in wave 40. 4,405 of these respondents answered the tax morale question in wave 41 and 33 did not. The experimental groups from wave 40 are: Control group, *Uncertainty* and *Unequal Knowledge* group. Control is omitted, implying that the effects are relative to the Control Group. All participants receive the following information: “There are frequent reports on tax evasion in the media.” The line p-val T1 vs. T2 presents the p-values from t-tests which compare if the regression coefficient for the *Uncertainty* group is different from the regression coefficient for the *Unequal Knowledge* group. Columns (1)-(4) differ in the included sets of covariates. (1): no covariates, (2): gender, age, marital status, household size, employment status, retirement status, and education, (3): (2) plus net personal income, (4): (3) plus political preferences. Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A.6: Balancing tests of experimental groups in follow-up

Variable	Control	Uncertainty	Unequal knowledge
<b>Gender. Reference category Male</b>			
Sex	0.001 (0.014)	0.001 (0.014)	-0.001 (0.014)
<b>Marital status: Reference category: Not married</b>			
Married	0.019 (0.014)	-0.035** (0.014)	0.015 (0.014)
<b>Full time employment: Reference category: Not full time employed</b>			
Full-time employed	-0.003 (0.014)	-0.003 (0.014)	0.006 (0.014)
<b>Retirement status: Reference category: Not retired</b>			
Retired	0.005 (0.018)	-0.001 (0.018)	-0.004 (0.018)
<b>Household size: Reference category Single households</b>			
2	0.023 (0.020)	-0.062*** (0.021)	0.038* (0.020)
3	0.016 (0.024)	-0.032 (0.024)	0.016 (0.024)
4	0.026 (0.025)	-0.057 (0.025)	0.031 (0.025)
5+	0.073** (0.036)	-0.069 (0.035)	-0.004 (0.035)
<b>Education: Reference category: Low Education</b>			
2	-0.026 (0.039)	0.002 (0.039)	0.024 (0.039)
3	0.000 (0.040)	-0.039 (0.040)	0.039 (0.039)
high educ	-0.012 (0.040)	0.001 (0.040)	0.011 (0.039)
<b>Household net income: Reference category poor</b>			
2	0.007 (0.019)	0.005 (0.019)	-0.011 (0.019)
3	-0.016 (0.022)	0.019 (0.023)	-0.003 (0.023)
4	-0.001 (0.033)	0.016 (0.033)	-0.015 (0.033)
rich	0.052 (0.036)	-0.005 (0.035)	-0.048 (0.034)
6	0.020 (0.024)	-0.005 (0.023)	-0.048 (0.023)
100	0.004 (0.038)	0.027 (0.038)	-0.048 (0.037)
<b>Age category: Reference category: &lt; 34</b>			
34 – 49	-0.018 (0.027)	-0.000 (0.027)	0.018 (0.027)

> 49	0.019 (0.027)	-0.005 (0.027)	-0.014 (0.027)
<b>Political orientation: Reference category: Conservatives</b>			
Left-wing	-0.009 (0.015)	0.015 (0.015)	-0.005 (0.015)
3	0.009 (0.025)	0.006 (0.024)	-0.005 (0.024)
100	-0.002 (0.037)	0.027 (0.038)	-0.025 (0.037)

Notes: The table shows the coefficients and standard errors (in parentheses) from a series of regressions of the form:  $y_i = \beta Covariate_i + \epsilon_i$ . Where  $Covariate_i$  is the respective covariate listed above. In Column (1)  $y_i$  equals 1 if participant i is in the control group and 0 otherwise. In Column (2),  $y_i$  equals 1 if participant i is in the *Uncertainty* group and 0 otherwise. In Column (3),  $y_i$  equals 1 if participant i is in the *Unequal Knowledge* group and 0 otherwise. The table is based on 4,405 respondents who answered the tax morale question in wave 41 and underwent the experiment in wave 40. Robust standard errors are in parentheses \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.



Download ZEW Discussion Papers from our ftp server:

<http://ftp.zew.de/pub/zew-docs/dp/>

or see:

<https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html>

<https://ideas.repec.org/s/zbw/zewdip.html>



## IMPRINT

**ZEW – Leibniz-Zentrum für Europäische  
Wirtschaftsforschung GmbH Mannheim**

ZEW – Leibniz Centre for European  
Economic Research

L 7,1 · 68161 Mannheim · Germany

Phone +49 621 1235-01

[info@zew.de](mailto:info@zew.de) · [zew.de](http://zew.de)

Discussion Papers are intended to make results of ZEW

research promptly available to other economists in order  
to encourage discussion and suggestions for revisions.

The authors are solely responsible for the contents which  
do not necessarily represent the opinion of the ZEW.