ORIGINAL ARTICLE

Economic Inquiry

Outcomes or participation? Experimentally testing competing sources of legitimacy for taxation

Christoph Engel[1] | **Luigi Mittone[2]** | **Azzurra Morreale[3,4]**

[1]Max-Planck-Institute for Research on Collective Goods, Bonn, Germany

[2]Economics Department, Trento University, Trento, Italy

[3]Jönköping International Business School, Jönköping University, Jönköping, Sweden

[4]LUT Business School, LUT University, Lappeenranta, Finland

Correspondence
Christoph Engel, Max-Planck-Institute for Research on Collective Goods, Kurt-Schumacher-Straße 10, Bonn D 53113, Germany.
Email: engel@coll.mpg.de

Abstract
Legitimacy may result from support for projects that a government implements. However, legitimacy may also result from the opportunity to participate in the selection process of projects. We tested the strength of these competing sources of legitimacy experimentally and their relationship. We find a straightforward effect of the former: the more projects a participant supports, the higher their taxes. Participants are also willing to pay for participation; if they have had a say, they pay higher taxes. Yet, most of this effect is actually instrumental: participants want participation to ensure that their taxes are used for purposes they deem acceptable.

KEYWORDS
experiment, legitimacy, rule following, tax morale

JEL CLASSIFICATION
C91, D01, D63, D72, D78, D84, D91, H26, H41, K34, K42

1 | INTRODUCTION

In the economic textbooks, crimes are committed when the expected individual utility from breaking the law is higher than the expected disutility of the sanction (Becker, 1968). This utilitarian view has been contrasted with the claim that taxes are paid, as this is the prevailing social norm (Engel et al., 2020). It has further been argued that taxes are paid because taxpayers hold social preferences and support the financing of public projects. Finally, it has been asked whether taxpayers are motivated by the knowledge that a legitimate democratic process is utilized to determine how much taxes are requested and what the money will be used for. With the experiment reported in this paper, we investigate the interplay of the last two motives: support for the substantive policies on which tax money is spent and support for the procedure in which this decision is made.

In earlier work, we showed that a substantial fraction of experimental participants pay taxes even if they do not derive any personal benefit from the way the money is used (i.e., it is not spent for the provision of a public good) and if they do not know what the money will be used for (Engel et al., 2020). From an experimental perspective, this is welcome news, as rule following can be isolated as a motive. Experimental participants must be sensitive to this

---

Managing Editor: Tim Salmon

deontological concern (unless one wishes to postulate rule utilitarianism; Harsanyi (1982)). However, real governments do not often stipulate arbitrary rules. This holds, in particular, for the obligation to pay taxes.

1.1 | Competing sources of legitimacy

Typically, the tax burden is justified with the provision of public goods or with the desire for redistribution. Moreover, the obligation to pay taxes is not inflicted upon the population by a deus ex machina. The famous maxim goes: no taxation without representation.[1] Those who are supposed to pay have had a say in the decision that obliges them. In this experiment, we add the substantive dimension (what is the money used for?) and the procedural dimension (how is it decided what the money is used for?) and investigate how they affect the taxpayers' willingness to abide by an un-enforced rule that asks for the payment of taxes.

We investigate the relationship among these alternative dimensions of legitimacy and how they are related to the general willingness to abide by rules. Tax morale could possibly be low in the first place (even if perhaps not zero), which corresponds with the deterrence story. It could be that tax morale is high across the board (albeit perhaps imperfect) or that participation is critical, which would correspond with the rule following/social norm and democratic participation stories, respectively. It is also possible that the most important determinant of tax morale is harmony between taxpayers' wishes and political choices, which corresponds with the taxpayer preference story. Finding the relative strength of these alternative motives and assessing their relationships are our research questions and have motivated the design of the experiment.

1.2 | Design

In the interest of high internal validity, we implemented a lab experiment. Participants earned an endowment in a real-effort task. They are informed that they are supposed to pay a fraction of the endowment as a tax to finance "public entities." Yet, it is made clear that there is no audit and that transferring less or no money does not have any pecuniary consequences.

Substantive legitimacy is manipulated in the following way: the tax return is transferred to one or several recipients. The participants indicated, for each of 10 public entities, whether they would be fine with money being transferred to the entity. These ratings are our measure of substantive legitimacy. We also incentivize participants to guess the three most popular recipients.

Regimes are our treatments. They reflect different forms and degrees of procedural legitimacy. In the *Imposed* condition, participants are informed that the recipient will be randomly selected from the list of potential recipients. In the *Winner-Takes-All* condition, the money collected from all participants (in that condition) is transferred to the entity approved by most participants. This is in the spirit of the Westminster model of democracy, in which majority rules. In the *Proportional* treatment, the money is split proportionally based on approvals, such that any entity that has been approved by at least one participant receives a positive amount. This is in the spirit of proportional representation.[2] In the *Gold-Silver-Bronze* treatment, half of the money goes to the entity participants ranked highest, a third to the entity ranked second and a sixth to the entity ranked third. This is in the spirit of qualified proportional representation, which is exemplified by the political system in Germany, where a party must collect at least 5% of all votes to be represented in parliament.[3] The *Gold-Silver-Bronze* treatment is also motivated by characteristic constraints of the political process. Putting a new issue on the agenda of a polity typically requires a major concerted effort by multiple authorities and stakeholders (Kingdon, 1995; Rochefort & Cobb, 1994). Every issue must fight for fairly limited attention in a crowded policy space. For legislative change to occur, a narrow window of opportunity must be found (Meyer & Minkoff, 2004). Policymakers are much less likely to invest in a cause they do not consider sufficiently popular.

The *Imposed* regime is the default. Participants made two choices. We first used an incentive-compatible mechanism to elicit their willingness to pay to be in one of the alternative regimes rather than the *Imposed* regime, in which the recipient of the tax return was chosen at random. In each session, one of the three alternative regimes was randomly selected. Participants with a sufficiently high willingness to pay are transferred to that regime. The taxes paid by the remaining participants were handed over to one randomly chosen recipient. For the second choice, we used the strategy method (Selten, 1967). We asked participants to decide how much taxes they were willing to pay, provided that they were in the *Imposed* regime or in any one of the three alternative regimes. For participants who exhibited a sufficiently

high willingness to pay to be transferred into the randomly selected alternative regime, the choices they made were implemented.

1.3 | Preview of results

On average, participants have a substantial willingness to pay to not be in the default regime. Willingness to pay is most pronounced for the *Proportional* regime. Participants pay substantially higher taxes in all the alternative regimes and significantly more in the *Gold-Silver-Bronze* and *Proportional* conditions compared with the *Winner-Takes-All* condition. Thus, we have clear support for legitimacy derived from participation. The more recipients a participant feels comfortable with, the higher their tax payments. This effect is independent of the effect of support for the cause for which the money is used. Thus, we have support for both sources of legitimacy.

However, actually, the picture is more nuanced. We have deliberately constructed the list of potential recipients such that some are unlikely to be popular among our student participants. The results are in line with our expectations. Beliefs are even more extreme. A large majority of participants expected the same two recipients to be the front-runners. The willingness to pay for an alternative regime is almost completely explained by the fact that a participant feels personally comfortable with the recipients whom they believe will be favored by the majority. However, tax declarations are independently explained by treatment *and* the match between beliefs and individual preferences. To the extent that there is a match, we measured the instrumental value of participation. Participation is sought, and participation increases tax payments, as participation makes it more likely for projects that taxpayers deem acceptable to be financed. Yet, we find a separate, additional effect of treatment on tax declarations. Consequently, the mere fact that they have an influence on the decision, even if it does not lead to an outcome they personally desire, also conveys a certain degree of legitimacy.

The remainder of the paper is organized as follows. The next section defines our contribution to the literature. Section 3 explains the design of the experiment. In Section 4, we define our hypotheses. Section 5 reports the results, and Section 6 concludes with a discussion.

2 | CONTRIBUTION

2.1 | Incentives

The canonical model interprets tax evasion as a crime. Similar to theft or assault, individuals are predicted to evade taxes if the utility from committing the crime exceeds the cost (Becker, 1968). If a government wants to collect taxes, it must deter its citizens from evading taxes (Allingham & Sandmo, 1972). For over half a century, this model has been considered a benchmark for tax behavior (Luttmer & Singhal, 2014). If one expects this model to capture the essence of citizens' decision to pay taxes, then the detection probability, tax rate and severity of sanctions are critical (Luttmer & Singhal, 2014). The model assumes that taxpayers are (exclusively) motivated by expected income. If penalties or the probability of being detected are low, the model predicts that citizens will evade taxes. Increasing detection probability and penalties reduces this incentive and promotes tax compliance (Bott et al., 2020). From this angle, it is surprising how much money governments can collect and how little enforcement they require to do so. Previous research has attempted to rationalize payments within the framework of this rational choice model—with pronounced risk aversion, for instance—but the parameters are not quite convincing (Luttmer & Singhal, 2014, p. 152 f.).

2.2 | Tax morale

This insight has led to the emergence of a body of literature on tax morale that examines relevant non-monetary motivations (Luttmer & Singhal, 2014). Tax morale builds on factors such as social norms, social preferences and trust in the tax authority (Jimenez & Iyer, 2016; Kornhauser, 2006). Tax morale has been shown to have a substantial role in affecting tax compliance (Dwenger et al., 2016). It has been traced back to intrinsic motivation, reciprocity, peer effects and culture (Luttmer & Singhal, 2014). With our project, we aim at unpacking this set of motives.

2.3 | Legitimacy

A plausible motive is perceived legitimacy. Legitimacy links the exercise of sovereign powers to the attitude of the people toward government. The concept has a normative and positive side. On the normative side, saying that a governmental act is legitimate leads to the expectation that citizens will act in accordance with it. Not only should citizens not be disobedient, but even if the act leaves room for circumvention, they are not supposed to exploit the opportunity. On the positive side, if a governmental act is perceived as legitimate, citizens are believed to act in line with it. Hence, on the positive side, legitimacy is linked to acceptance. Even if the act is not what a person would have chosen autonomously, they go along with the act because they understand the rationale behind it ("output legitimacy") or because they accept that the decision is the result of an adequate decision-making process ("input legitimacy") (Easton, 1965; Norris, 1999; Scharpf, 1999).[4]

2.4 | Social norms

Drawing upon insights from social psychology (Aronson et al., 2007), one may expect that social norms play a significant role in shaping individuals' tax compliance behavior (Abraham et al., 2017; Battiston & Gamba, 2013; Jimenez & Iyer, 2016). Guala and Mittone (2010) define social norms in relation to an extensive debate in the institutional economics literature (Acemoglu & Jackson, 2017; Bicchieri, 2006; Ellickson, 1998; Schultz et al., 2007). They focus on the distinction between social conventions and social norms, with the latter characterized by a conduct rule or strategy, denoted as S. This strategy must meet three conditions: (a) all players prefer to play S when everyone else does the same, (b) all players are aware of this preference and (c) players who choose not to play S expect to experience some costs that reduce their utility. According to this definition, the effectiveness of the social norm that promotes tax compliance is based on the taxpayers' belief that other taxpayers will also fulfill their tax obligations, the shared knowledge of this belief in society and the recognition that tax evasion carries a social stigma and thereby has a cost.

Cialdini and Trost (1998) present a more articulated definition and discussion of the concept of social norm. They follow the traditional social psychology approach that distinguishes among four main norm constructs: injunctive norms, descriptive norms, subjective norms and personal norms. Injunctive norms are closely related to the first condition of Guala and Mittone (2010) because they refer to perceptions of what is approved or disapproved by others. Cialdini et al. (1990) define descriptive norms as rules resulting from the perception of others' behavior. These norms reflect observations of how others typically behave. Cialdini et al. (1990) define subjective norms as the perceived social pressure to engage or not to engage in a specific course of behavior, and personal norms as the individual's internalized standards of behavior. Subjective norms represent individuals' beliefs about whether key people in their lives (e.g., family, co-workers, etc.) approve or disapprove of a given behavior. Applying this definition to our case, one could say that a taxpayer decides to pay taxes because they believe that their peers would disapprove of tax evasion. This mechanism has been experimentally verified by behavioral economics research (e.g., Casal & Mittone, 2016). However, personal norms often differ between individuals and may vary from the prevailing norms in a given society or group.

The role played by social norms in the field of tax evasion has been intensely investigated. For example, in a study conducted by Alm et al. (1999) it was found that almost every participant resorted to cheating when the prevailing social norm, as determined by the voting results on the enforcement regime, tended toward tax evasion. Wenzel and Taylor (2004) investigated how educative communication influences tax compliance and found that the effectiveness of severe sanctions in encouraging compliance is contingent upon favorable social norms.

The aforementioned studies primarily examine external societal reinforcements, even if individual attitudes and personal value systems often mediate these influences. Other streams of research conduct more in-depth investigations of the relationship between personal ethical norms and social norms (Bobek & Hatfield, 2003; Hanno & Violette, 1996). Bobek et al. (2013) found that personal norms and subjective norms directly affect tax compliance decisions, while descriptive and injunctive norms have an indirect impact.

2.5 | Social preferences

A large body of experimental literature shows that many participants hold social preferences (Fehr & Schmidt, 2006). They have a positive willingness to pay for more equitable outcomes (Bolton & Ockenfels, 2000; Charness &

Rabin, 2002; Fehr & Schmidt, 1999). This could translate into the willingness to pay taxes, even in the absence of or with insufficiently strong sanctions (Engel, 2014), provided that the money is used for purposes the taxpayer is willing to support.

Experiments also show that participants give substantial amounts to charity (Engel, 2011). Such individuals may care that their charitable intentions are not perverted. This explains why donors are willing to pay to insure their donation against the risk of getting lost (Buijze et al., 2017). In the same spirit, they might also want to pay for a regime that allows them to influence the choice of the recipient of the taxes collected from them.

Social preferences and social norms can be related. This is the case if a person who individually believes in the desirability of some social outcome (and hence holds the corresponding social preference) additionally perceives a social expectation to bring this outcome about (and hence is motivated by the corresponding social norm). Take, for instance, a devout liberal. Such a person may not only individually be in favor of expanding LGBT+ rights. She may also feel obliged, in relation to her peer group, to stand up in favor of these rights.

2.6 | Support for public projects resulting from social norms and social preferences

Ackert et al. (2007) examined how individuals' aversion to inequality influences their choices concerning tax regimes. Their findings suggest that social preferences, particularly aversion to inequality, affect voting for tax regimes. Notably, even wealthy individuals often vote in favor of progressive tax, knowing that they will be asymmetrically affected themselves. In a way, the expectation that social preferences contribute to tax compliance is also informed by data showing that taxpayers dislike "Okun's leaky bucket"[5] (Beckman et al., 2004; Lambert, 1988; Slemrod, 1994) and, hence, are sensitive to the effectiveness of policies financed by tax payments.

Perceptions of fairness also affect tax compliance (Falkinger, 1995). Torgler (2007) examined the role of fairness attitudes in a theoretical model. Studies by Andreoni et al. (1998) and Bordignon (1993) support this claim. Empirical evidence suggests that compliance is affected by taxpayers' perceptions of their treatment relative to others (Alm, 1991).

Stanley and Hartman (2018) manipulate the degree of transparency about governmental spending. Taxpayers either only learn what fraction is spent on "welfare," are given a breakdown of the spending into four categories (with "health" being the most sizable), or are presented with 25 even more fine-grained categories. Participants were asked about their attitudes. Attitudes toward taxpaying were most positive in the four categories condition, but support for welfare was highest in the non-transparent condition. Sælen and Kallbekken (2011) elicit support for an increase in fuel tax. Support was more pronounced if funds were earmarked for environmental protection.

Results from field experiments have been mixed. Blumenthal et al. (2001) found that explaining to Minnesota citizens what their tax dollars are chiefly spent on did not increase compliance. Castro and Scartascini (2015) found that reminding Argentinian taxpayers of infrastructure investments during the past 6 months did not increase payments as well. A letter that simultaneously pointed Swiss citizens to the purpose of taxation and to the civic duty to pay taxes did not have a significant effect either (Torgler, 2004), nor did informing Israeli corporations "how tax dollars were allocated to finance public commodities" increase payments (Ariel, 2012). In contrast, a letter sent to Norwegian taxpayers who likely had foreign income had an effect. They were told that "your tax payment contributes to the funding of publicly financed services in education, health and other important sectors of society" (Bott et al., 2020). Likewise, a letter sent to UK taxpayers who were late with handing in their tax returns had a small (1.3%) but significant effect. The message read, "Paying tax means we all gain from vital public services like the National Health Service (NHS), roads and schools" or "Not paying tax means we all lose out on vital public services like the NHS, roads and schools."[6]

2.7 | Support for the procedure used to define tax burden and the purposes for which tax money is spent

Studies further show that participants care about procedure (Tyler, 2006) and about voice in particular (Kleine et al., 2016, 2017). More broadly, trust in government has been identified as a key determinant of compliance behavior (Jackson & Milliron, 2002; Torgler, 2003, 2007). This could translate into the willingness to pay taxes if the obligation to pay results from a procedure that taxpayers deem appropriate.

Finally, previous research has shown that participants are willing to fulfill their duties, even if they know that the normative expectation will not be enforced (Kimbrough & Vostroknutov, 2016). This approach is related to the concept

of "civic virtues" (Frey, 1997). Normativity could translate into a willingness to pay taxes simply because there is a rule that makes payment mandatory (Desmet & Engel, 2021; Engel et al., 2020). Casal et al. (2022) report a similar—if not even stronger—result that confirms how personal systems of ethical values can raise the level of tax compliance without any reinforcement by external formal or social rules.

The effect of voting on tax compliance has been tested experimentally. Alm et al. (1993) gave participants a random amount of income and set up a tax scheme in which 30% of the participants' income was taken and used for student financial aid, student health services or the university president's office. An audit was conducted, and a sanction was given if tax evasion was detected. The authors manipulated whether the purpose for which the budget was spent was imposed or chosen by a majority vote. When the purpose was voted upon, tax payments were higher. Giving individuals a vote on procedural issues is more delicate: If the community votes in favor of less stringent enforcement, tax payments go down (Alm et al., 1999). Even if the majority supports enforcement, those who voted against it are very likely to evade taxes (Feld & Tyran, 2002).

In a democracy, the people have a say in public spending. Yet, tax compliance is only correlated with the degree of democracy if the democracy is well established, rather than newly introduced or precarious (Zheng et al., 2019). In the Swiss cantons that have direct democracy, the belief that it is wrong not to report income truthfully is more pronounced (Torgler, 2005). If the budget of a local community is spent only on a public good that serves that community, tax evasion is less pronounced than if the budget of two remote communities is spent on two public goods, of which one serves the other community (Güth et al., 2005). Moreover, when Chinese taxpayers are informed that local representatives are now more responsive to the concerns of citizens, it leads to a decrease in their willingness to comply with tax regulations. This is likely because they realize that the responsiveness is still limited, regardless of the improvements (Kao, 2016).

In political science, the importance of the procedural source of legitimacy is debated. Proponents of political psychology stress the relevance of procedural fairness (Lind & Tyler, 1988; Tyler, 1997). However, sceptics object that participation is only perceived as a technique of "stealth democracy": the powerful use it as a pretext to advance their interests. Citizens see this and care about outcomes (Hibbing & Theiss-Morse, 2002).

In political science, a series of empirical studies investigated the matter, but not with regard to tax compliance. Arnesen (2017) faced participants with a choice between a substantial amount of money being raffled to one of the participants or being donated to a charity. In one condition, participants were told that they "have a say" on the outcome. They were then randomly informed about either outcome and asked how acceptable they found the decision. Only outcome preferences explained acceptance. In a series of vignette studies, Esaiasson et al. (2019) found that the personal effect of a political decision on a participant is the dominant factor in their rating of the acceptability of the decision, irrespective of the procedure employed in making the decision. Strebel et al. (2019) used conjoint analysis to assess how participants in a population survey evaluate a public transport project. They found that cost-effectiveness is by far the most important concern. However, various definitions of public impact also affect approval ratings. The degree to which the individual supported the project in the first place was not manipulated.

Esaiasson et al. (2012) confronted classmates with the decision to either spend an endowment for their own enjoyment or to have it donated to a charity. Between classes at the same school, they varied the decision-making process, ranging from joint decision-making with full participation to lottery choice. They elicited a rating for the fairness of the procedure and found that the decision-making process involving the direct participation of each classmate was rated highest. In a survey by Neblo et al. (2010), participants stated a pronounced interest in deliberation, but this was not contrasted with outcome preferences.

2.8 | The present experiment

Our experiment chiefly differs from these earlier contributions based on the research question. We do not consider substantive and procedural motives in isolation; instead, we consider them in relation to each other. Which is more important for tax compliance: how strongly an individual supports the purposes for which their money is used, or the degree of their impact on the definition of these purposes? Can lower support for causes be compensated for by higher impact on decision-making? Or, in the language of political science, what is the relationship between input and output legitimacy as determinants of tax compliance?

We also introduce a number of design features that allow for a considerably cleaner test. On the side of independent variables, we do not merely assume preferences but elicit them. We also do not merely compare the presence and

absence of taxpayer impact on the definition of a public good. Rather, we introduce degrees of influence. This, in particular, makes it possible to determine whether lower support for the cause can be compensated for by higher influence on the decision. On the side of the dependent variable, we exploit the fact that we have shown that tax evasion can be meaningfully tested in the complete absence of audits and sanctions in an earlier experiment (Engel et al., 2020). We can therefore isolate the effects on tax morale and need not worry about its interaction with the incentive and behavioral effects of surveillance, and of heterogeneous aversion to the risk of enforcement in particular.

3 | DESIGN

Participants earned an endowment of 100 experimental currency units (ECU) in a real-effort task. Specifically, they were allowed to move 10 sliders to a randomly selected point. For each slider that they move to the correct point, they earn 10 ECU. They are informed that they are supposed to pay a 30% tax. Their declaration is pre-filled, but it is made clear to them that they are perfectly free to change their declaration. They are free to pay any amount of taxes between 0 and 30 ECU. They know that there is no audit and, consequently, no sanction for tax evasion.

Participants are further informed that one of the following 10 public entities may receive taxes collected by the experimenter:

Croce Rossa Italiana (Italian Red Cross)
Ospedale S. Chiara di Trento (local hospital)
Servizio Bus Trento (local bus service)
Pompieri Trento (local fire brigade)
Servizio Pulizie Università Trento (university cleaning service)
Servizio Pulizie Strade Comune di Trento (local street-cleaning service)
Istituto Nazionale Previdenza Sociale (Italian pension system)
Esercito Italiano (Italian army)
Comando Polizia Trento (local police station)
Comando Guardia di Finanza Trento (local tax authority)

The recipients were presented to participants in a randomized order. Participants are asked which of the 10 potential recipients they are fine having their money given to.

Treatments differ in the way recipients are selected. In the *Imposed* condition, participants are told that all taxes collected will be donated to one of the 10 recipients selected at random.[7] In this condition, there is no legitimacy resulting from participation. Additionally, this condition had the lowest possible legitimacy from support for the cause to be financed by the taxes. In the *Winner-Takes-All* treatment, minority interests are not protected at all. The recipient with the strongest support receives all the taxes.[8] The *Proportional* treatment is the opposite. Minority interests have the same weight as majority interests. Each entity receives taxes in proportion to the support for it. The *Gold-Silver-Bronze* treatment strikes a balance. The most popular entity receives half the taxes, while the second and third entities receive one-third and one-sixth of the taxes, respectively.

The *Imposed* condition is the default. All three alternative conditions were explained to all participants. Using a variant of the mechanism by Becker et al. (1964), we elicit willingness to pay under the *Imposed* regime replaced by any of the alternative regimes. Conceptually, the mechanism is a (reverse) second-price auction (Clarke, 1971; Groves, 1973; Vickrey, 1961), which is dominance-solvable and therefore incentive-compatible. As it is known that there are incentive problems in second-prize auctions with multiple heterogeneous items (Ausubel, 2008, pp. 299–301; Clarke, 1971; Groves, 1973), we used the strategy method (Selten, 1967). Participants were informed that in each session, one of the three alternative regimes would be randomly selected, and their choices would be implemented accordingly if this regime was put into effect. They were then asked to state their willingness to pay for each of them.

If the participant states a willingness to pay for a specific regime above the randomly selected cutoff, that regime is applied, but they only pay the cutoff price. The taxes collected from participants with an insufficient willingness to pay go to a randomly selected recipient. Taxes from the remaining participants go to the recipient(s) defined by the regime in question. In selecting the recipient(s), the stated preferences of all participants in the session are taken into account, including the choices made by participants whose tax revenues are paid to a random recipient. These design features were communicated to participants.

We finally administered the Holt/Laury test for risk aversion (Holt & Laury, 2002) and gave participants a demographic questionnaire. The complete instructions are available in Supporting Information S2: Appendix. Feedback is withheld until the end of the entire experiment to ensure that no contamination occurs. The experiment was implemented in oTree (Chen et al., 2016). Screenshots from the slides are in Supporting Information S2: Appendix. Participants were invited with the help of recruitment software developed by the lab.

The experiment was preregistered with the Open Science Framework.[9] We ran the experiment in the Cognitive and Experimental Economic Laboratory of Trento University using its subject pool. We aimed for 134 participants. That would have given us power to detect an effect of medium size (standardized effect size of 0.25) at the conventional levels of $\alpha = .05$ and $\beta = .2$.[10] Unfortunately, we exhausted the subject pool after excluding participants who had already participated in a pilot study.[11] Only 98 participants were recruited.[12] Sixty (61.22%) were female. Participants were, on average, 23.2 years old. Fifty-one were economics students, while seven of them were not students. The remaining participants were from various majors. Participants, on average, earned 123 ECU (€7.39[13]) from the experiment, plus a show-up fee of €3.

4 | HYPOTHESES

If participants only care about their profit from the experiment, treatment effects should not be observed. The participants would not pay taxes regardless of what they would be used for, and the participants would not pay for participation (and, consequently, having participation rights would not induce them to pay taxes). This is the outcome predicted by deterrence theory. If, on the contrary—and relying on our earlier result (Engel et al. (2020)—the critical motive is the social norm of following rules, we expect that a number of participants will pay the requested tax, but their choices should be unrelated to the purpose of tax collection and to their participation in the decision. However, the design of the experiment is motivated by the expectation that either of these two radical versions misses important motives. We predict a main effect of support for outcomes:

H₁—Support for outcomes
The greater the chance for taxes to be used for a public good that an individual supports, the more comprehensively the individual reports their income.

We also predict a main effect of influence on defining outcomes[14]:

H₂ degree of participation
a) Participants have a positive willingness to pay to not be under the *Imposed* condition. This willingness is more pronounced the higher the individual's impact is on the choice of the public good(s) to be provided.
b) The higher the individual's impact is on the choice of the public good(s) to be provided, the more completely the individual reports their income.

Finally, we expect both dimensions of legitimacy to interact:

H₃ interaction
A higher impact on the choice of the public good(s) to be provided compensates for the low subjective desirability of the public good.

5 | RESULTS

5.1 | Support for outcomes

We constructed the list of potential recipients such that support is likely to vary. As Figure 1 shows, this expectation was borne out by the data. While almost all participants declare that they are fine with taxes being spent on the Red Cross or the local hospital, only a few of them deem it acceptable that taxes are spent on the military or tax authorities.

As Figure 2 shows, the degree of acceptance matters for willingness to pay taxes. The figure represents the mean tax payment per participant and pooled over all four alternative regimes. Besides the eight participants who declared being fine with any of the potential recipients, the more recipients a participant deemed acceptable, the higher the amount of taxes they pay on average.

Table 1 shows that the visual impression was supported by statistical analysis. The regression predicts that for any other recipient a participant is fine with, their tax payments increase by close to 1 (of 30) ECU. This holds for the mean tax payment across all four regimes and, specifically, for the regime in which participants had no influence whatsoever on the selection of the recipient.
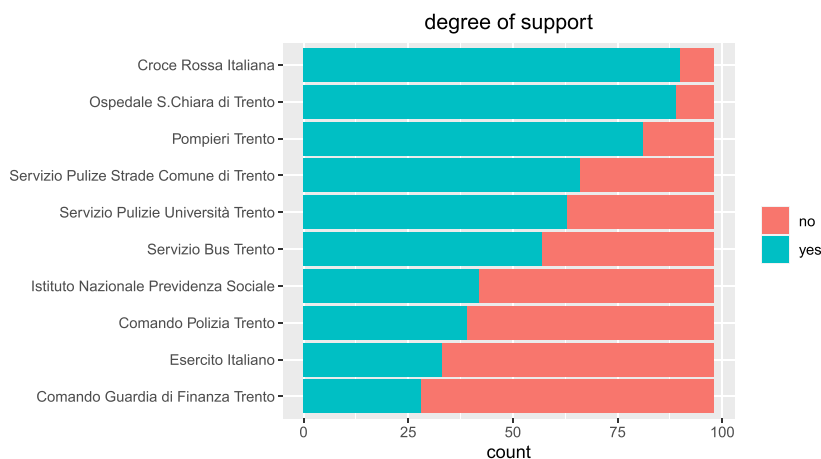

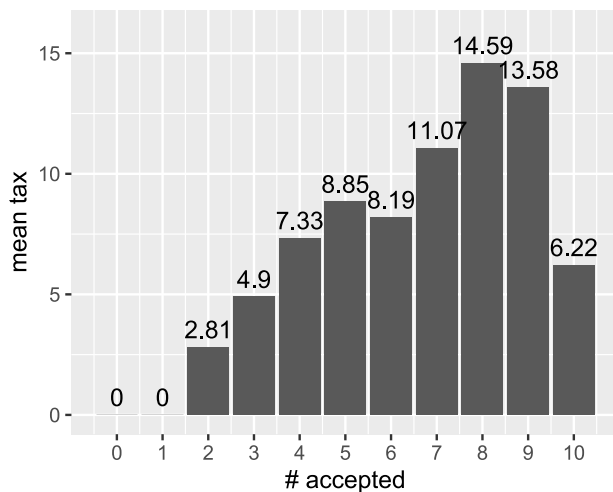**FIGURE 1** Degree of support for potential recipients.

FIGURE 2 Tax declarations as a function of the number of recipients the participant deems acceptable.

TABLE 1 Tax payment as a function of support for potential recipients.

	Mean tax over all four regimes	Tax in the *Imposed* condition
Number of recipients declared acceptable	0.991** (0.340)	0.853** (0.324)
Cons	2.783 (2.182)	−0.793 (2.075)
N	98	98

Note: OLS. Standard errors in parentheses.

***$p < .001$, **$p < .01$, *$p < .05$, +$p < .1$.

Thus, we have clear evidence for an effect of support for the public causes to be financed on tax morale and for H_1. We come to the following conclusion:

> **Result 1 (support for public causes)**: The more recipients a participant declares being comfortable with, the higher their tax payments.

5.2 | Participation

Figure 3 and the constants in the models of Table 2[15] show that participants have a substantial positive willingness to pay for not being in the *Imposed* condition. Participants not only care which projects are financed with their taxes (Result 1); they are also willing to spend additional money to make sure that they have a say in the selection of the recipient(s).

As Figure 3 and the model coefficients presented in Table 2 show, willingness to pay to be in an alternative regime does not differ profoundly between regimes. We only found a small difference between the *Winner-Takes-All* and *Proportional* conditions. In a linear model, this difference is only significant at the 10% level. A Tobit model considers the fact that in the *Winner-Takes-All* condition, a few more participants do not want to pay to leave the *Imposed* condition. In this specification, we find a significant effect of the *Proportional* condition at conventional levels.

With this qualification, we support H_{2a}[16] and conclude the following:

> **Result 2 (preference for participation)**: Participants have a positive willingness to pay to have a say in the selection of recipient(s) that will receive the tax return. This willingness to pay is most pronounced when taxes are paid out in proportion to the declared willingness of participants to finance the respective recipients.
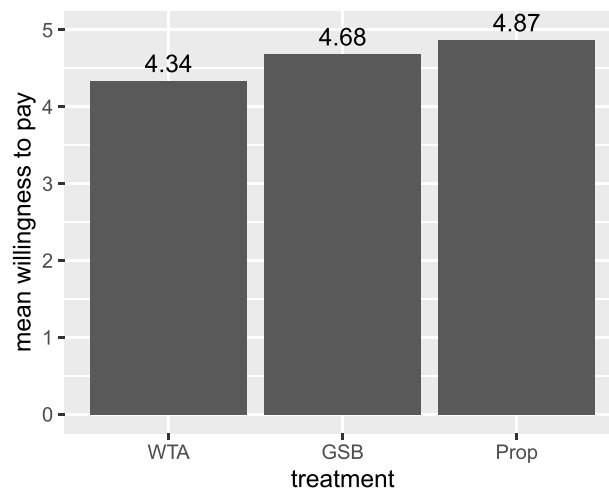

FIGURE 3 Willingness to pay to not be in the imposed condition.

TABLE 2 Willingness to pay to not be in the imposed condition.

	Linear	Tobit
Gold-Silver-Bronze	0.347 (0.276)	0.395 (0.366)
Proportional	0.531+ (0.276)	0.584* (0.280)
Cons (*Winner-Takes-All*)	4.337*** (0.289)	4.078*** (0.388)
N	294	294
N uid	98	98

Note: Model 1: linear model with participant random effect, Hausman test insignificant. Model 2: Tobit model with censoring at 0 and participant random effect. Standard errors in parentheses.

***$p < .001$, **$p < .01$, *$p < .05$, +$p < .1$.

Figure 4 shows that regimes have a dramatic influence on tax payments. In the *Imposed* condition, the large majority of participants paid either no taxes at all or only a very small amount. In contrast, tax payments are substantial under all alternative regimes. The beneficial effects of the *Gold-Silver-Bronze* and *Proportional* conditions are even stronger.

Table 3 presents the results of the statistical tests. Comparing Models 1 and 3, we see that the effect is more pronounced on the intensive rather than the extensive margin. Even in the most favorable condition, the *Proportional* condition, only a small fraction of participants completely paid the requested tax of 30 ECU. In the remaining treatments, Model 3 does not establish a significant difference from the (very low) fraction of full tax payments in the *Imposed* condition.

Yet, Models 1 and 2 show pronounced treatment effects on the intensive margin, that is, on the amount of taxes participants pay. In the linear Model 1, we also find a significant constant: on average, even under the *Imposed* regime, small tax payments were made. Taxes were about twice as high in the *Winner-Takes-All* regime and were even higher in the remaining conditions. Wald tests show that there is a significant difference between the *Winner-Takes-All* regime and the other two alternative regimes. If we replace the linear model with Tobit to take the higher fraction of zero tax payments in the *Imposed* condition into account, the treatment effects are even stronger.

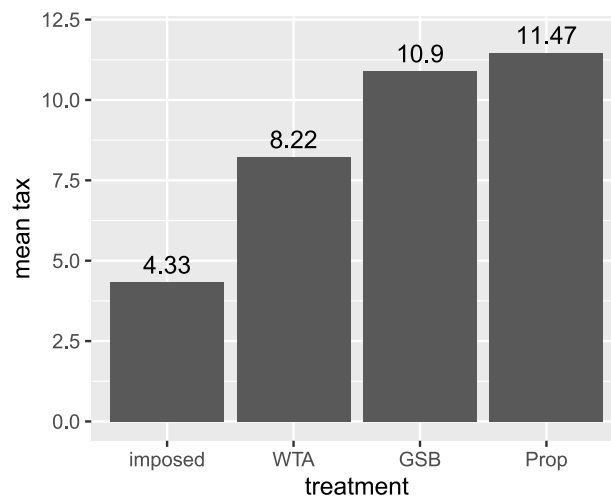Thus, we support H_{2b} and conclude the following:

FIGURE 4   Tax payment as a function of regime.

TABLE 3   Tax payment as a function of regime.

	Amount of taxes		Full taxes
	Model 1 **Linear**	**Model 2** **Tobit**	**Model 3**
Winner-Takes-All	3.898*** (0.732)	7.051*** (1.139)	0.020 (0.027)
Gold-Silver-Bronze	6.571*** (0.732)	10.542*** (1.276)	0.041 (0.027)
Proportional	7.143*** (0.732)	11.112*** (1.004)	0.112*** (0.027)
Cons	4.327*** (0.918)	1.039 (0.878)	0.031 (0.026)
N	392	392	392
N uid	98	98	98
Winner-Takes-All versus *Gold-Silver-Bronze*	$p < .001$	$p = .004$	$p = .442$
Winner-Takes-All versus *Proportional*	$p < .001$	$p < .001$	$p < .001$

Note: Models 1 and 3: linear models with participant random effect, Hausman test insignificant. Model 2: Tobit model with participant random effects. Last two lines: *p*-values from Wald tests. Standard errors in parentheses.

***$p < .001$, **$p < .01$, *$p < .05$, $^{+}p < .1$.

Result 3 (participation): Tax declarations are higher in the *Winner-Takes-All* condition than in the *Imposed* condition and are even higher in the *Gold-Silver-Bronze* and *Proportional* conditions.

5.3 | Relationship between both sources of legitimacy

In a democracy, not everybody sees their wishes fulfilled. However, everyone has a chance to influence the decision-making process, usually through general elections. Ideally, one would hope that those whose wishes remain unfulfilled still do their public duty and pay their taxes. Hence, ideally, democratic influence should compensate for the fact that an individual's preferences remain unsatisfied. The regressions in Table 4 only provide partial support for this expectation. Even when controlling for the number of potential recipients a participant declares acceptable, treatment effects can be established. Comparing Model 1 of Table 4 with Model 1 of Table 3, we see that treatment coefficients are almost perfectly identical whether or not we control for individuals' preferences concerning public goods. This implies that treatment effects are present even if an individual does not feel comfortable with any of the potential recipients.[17]

Model 2 of Table 4 shows that all interactions between treatment and the number of recipients a participant is willing to support are insignificant. Hence, the effects of legitimacy resulting from participation and legitimacy resulting from support for public causes are almost perfectly independent of each other. Participation helps with tax morale, but more pronounced participation does not compensate for lower support for the projects chosen by the polity.

Thus, we only have qualified support for H_3 and note the following:

Result 4 (independent sources of legitimacy): The effects of regime and of the number of recipients a participant declares acceptable on the amount of taxes a participant pays are independent of each other.

We exploit belief data to better understand treatment effects. The comparison of Figure 5 with Figure 1 reveals that beliefs are distributed more unevenly than declared preferences. This should not come as a surprise, since we have asked which recipients participants consider to be the three most popular ones. The fact that so many participants believe the Red Cross and the local hospital to be popular demonstrates that the majority of participants have a good sense of which recipients are most attractive to the majority of the experimental population.

Figure 6 reports willingness to pay to be in any one of the alternative regimes, conditional on the number of recipients for which it holds: the participant feels comfortable with the recipient(s) herself, and expects the recipient(s) to be among the three most popular ones. If this variable is 0, the participant in question is not fine with taxes being handed out to any of the three recipients they expect to be most popular. Hence, these participants either have preferences that strongly diverge from the preferences they expect the majority to hold or only listed a small number of acceptable recipients in the first place. The higher this number, the higher the willingness to pay to not be in the *Imposed* condition.

TABLE 4 Tax payment as a function of regime and support for potential recipients.

	Model 1	Model 2
Winner-Takes-All	3.898*** (0.732)	5.481** (2.057)
Gold-Silver-Bronze	6.571*** (0.732)	4.345* (2.057)
Proportional	7.143*** (0.732)	4.476* (2.057)
Number of recipients declared acceptable	0.991** (0.340)	0.853* (0.393)
Winner-Takes-All × number of recipients declared acceptable		−0.264 (0.321)
Gold-Silver-Bronze × number of recipients declared acceptable		0.371 (0.321)
Proportional × number of recipients declared acceptable		0.444 (0.321)
Cons	−1.620 (2.228)	−0.793 (2.520)
N	392	392
N uid	98	98

Note: Linear models with participant random effect, Hausman test insignificant. Standard errors in parentheses.
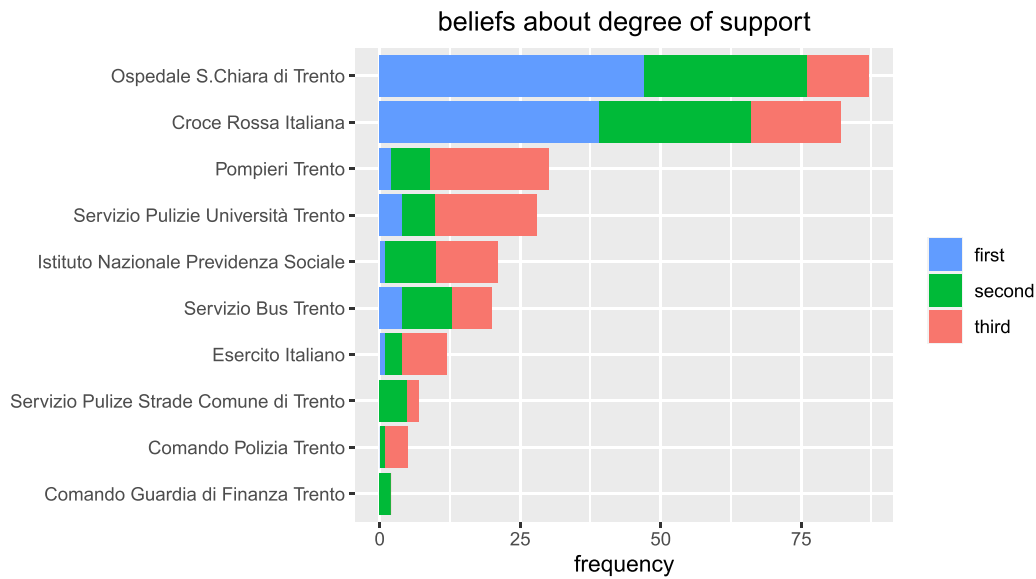
***$p < .001$, **$p < .01$, *$p < .05$, $^+p < .1$.
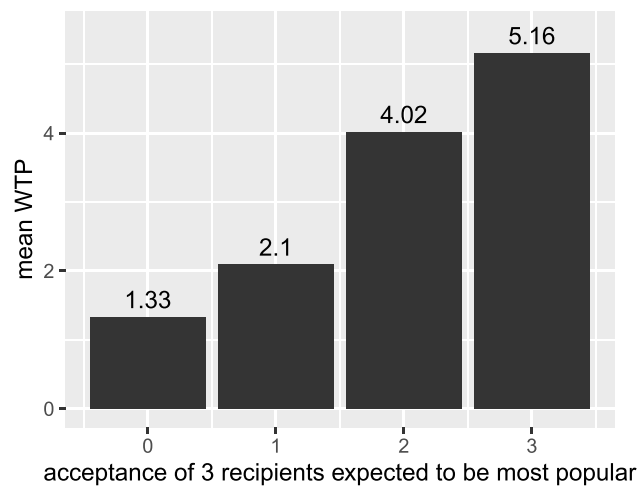
FIGURE 5 Beliefs.

FIGURE 6 Willingness to pay to not be in the imposed condition, conditional on support for the recipients a participant expects to be most popular.

Table 5 presents the statistical tests. Model 1 repeats Model 1 of Table 2 for comparison. Model 2 shows that the expectation that others will desire the same public goods as oneself strongly increases the willingness to pay for any of the alternative regimes. Model 3 shows that this expectation has an effect on willingness to pay that is independent of the treatment. When controlling for the *Gold-Silver-Bronze* and *Proportional* regimes, the coefficient and standard error of the expectation remain constant. Only a slight reduction in the effect of expectations is observed when the interaction terms between expectations and treatments are considered. The interaction effects were insignificant. This yields a clear interpretation of willingness to pay for one of the alternative regimes; it is more pronounced when the participant has a higher degree of expectations regarding their democratic influence to channel tax revenue to the recipients they deem acceptable. Participation is desired as a means to achieving an acceptable substantive outcome.

This gives us the following:

Result 5 (participation as a means to achieve a desirable outcome): Willingness to pay for alternative regimes increases with the number of expected popular recipients a participant finds acceptable.

The previous result could imply that influence on the process of selecting public goods is exclusively desired for its instrumental value. If that interpretation were to hold, the effect of being in one of the alternative regimes on tax payments would have to vanish once we controlled for the number of recipients that the participant expects to be most popular and found acceptable. However, as Model 2 of Table 6 shows, this is not the case. In fact, treatment effects stay

TABLE 5 Willingness to pay to not be in the imposed condition, conditional on support for the recipients a participant expects to be most popular.

	Model 1	Model 2	Model 3	Model 4
Gold-Silver-Bronze	0.347 (0.276)		0.347 (0.276)	−0.159 (0.992)
Proportional	0.531$^+$ (0.276)		0.531$^+$ (0.276)	0.613 (0.992)
Support for recipients the participant expects to be most popular		1.356*** (0.288)	1.356*** (0.289)	1.302*** (0.360)
Gold-Silver-Bronze × support for recipients the participant expects to be most popular				0.195 (0.368)
Proportional × support for recipients the participant expects to be most popular				−0.032 (0.368)
Cons	4.337*** (0.289)	1.114 (0.778)	0.822 (0.797)	0.963 (0.971)
N	294	294	294	294
N uid	98	98	98	98

Note: Linear models with participant random effect, Hausman test insignificant. Standard errors in parentheses.

***$p < .001$, **$p < .01$, *$p < .05$, $^+p < .1$.

TABLE 6 Tax payment as a function of regime and support for the recipients a participant expects to be most popular.

	Model 1	Model 2	Model 3
Winner-Takes-All	3.898*** (0.732)	3.898*** (0.732)	3.898*** (0.732)
Gold-Silver-Bronze	6.571*** (0.732)	6.571*** (0.732)	6.571*** (0.732)
Proportional	7.143*** (0.732)	7.143*** (0.732)	7.143*** (0.732)
Support for recipients the participant expects to be most popular		3.079** (1.026)	3.072** (1.020)
Risk attitude			−0.459 (0.487)
Age			0.128 (0.270)
Female			3.501* (1.666)
Law major			4.422$^+$ (2.279)
Engineering major			3.041 (3.401)
Humanities major			−0.293 (4.332)
Sociology major			−0.368 (3.831)
Psychology major			−7.869 (7.412)
Mathematics or physics major			−0.331 (3.503)
Other major			7.680* (3.437)
Not student			−2.071 (3.307)
Cons	4.327*** (0.918)	1.039 (0.878)	−6.832 (7.204)
N	392	392	392
N uid	98	98	98

Note: Linear models with participant random effect, Hausman test insignificant. Standard errors in parentheses.

***$p < .001$, **$p < .01$, *$p < .05$, $^+p < .1$.

perfectly the same when this control variable is added. The constant is affected, not treatment effects. As Model 3 shows, treatment effects also remain perfectly identical if we add further controls, namely risk attitude, age, gender and major. Hence, participation has an independent effect. It is desired as such, not only for its instrumental value.

We conclude the following:

> **Result 6 (independent effect of participation)**: Participants who are in a regime that allows them to influence the selection of public goods even declare higher taxes when controlling for the instrumental value of participation.

From Figure 3, we know that the willingness to pay to not be in the *Imposed* condition is pronounced. From Table 5, we know that this willingness to pay is mostly instrumental. However, from Table 6, we know that there is also a residual willingness to pay for participation, even if it is not explained by the expectation that participation will secure a desirable policy outcome. Either way, one may wonder whether the opportunity to invest in participation is backfiring: do the participants who cherish participation react with low tax payments if they are deprived of influence on the choice of recipient?[18]

The graphs in Figure 7 do not convey this impression. Even if willingness to pay to be in the respective participatory regime is pronounced (the value on the *x*-axis is high), there is still a discernible fraction of participants who pay substantial taxes if the recipient is *Imposed* (the value on the *y*-axis is high). We have made this visible by color-coding: the closer to red a dot is, the more it contradicts that missing the opportunity for participation is counterproductive. In the *Proportional* treatment, a high willingness to pay for participation and high tax payments in the *Imposed* condition most frequently coincide. The positive association even turns out to be significant at conventional levels (Table 7).

This is remarkable, as not only is the participation option not backfiring, but those most keen on participation are also the ones who are most willing to pay their taxes in the absence of participation. This suggests that the intrinsic willingness to pay taxes is most pronounced in participants who care most about individual influence (which is strongest in the proportional condition). These participants care about public goods in general (they pay even if they must accept the random selection of recipients), but they also care about directing tax money to causes they consider worthy (they pay for participation).

6 | DISCUSSION

Governmental oversight tends to be imperfect. Citizens often stand a chance to escape audits. Moreover, audits and enforcement are costly. If the only reason why citizens pay taxes is deterrence, tax evasion should be observed much more broadly than anecdotal evidence suggests. If, by contrast, the main source of tax morale is deontological, the only information governments would need is a personality trait: the more tax authorities expect a citizen to be conscientious and hence willing to abide by the rules they know to be in force, the more they could expect these citizens to pay diligently. The present experiment is motivated by the expectation that both these behavioral effects are moderated by citizens' policy preferences and by their preferences for participation. Our experiment is thus designed as a horse race between four stories: a pessimistic deterrence story, a very optimistic social norm story, an intermediate policy preference story, and a democracy preference story.

We found a straightforward effect of support for the substantive causes to be financed on taxes. The more potential public goods a participant declares to be comfortable with, the higher the amount of taxes they pay. For two reasons, this is a remarkable finding. It was made absolutely clear during the experiment that tax evasion would not have any pecuniary consequences. The mere fact that participants do not object to a potential use of the public budget already increases their payments: We did not elicit preferences for public goods. We merely asked whether a participant would be comfortable with taxes being used to finance certain public projects. This is a considerably weaker statement, but it still had a measurable effect. Arguably, this weaker measure also comes closer to legitimacy theory; it may suffice that taxpayers understand why their taxes are used for meaningful purposes, even if they would not have prioritized all of these purposes themselves.

The effect of legitimacy from participation is more nuanced. Participants have a pronounced willingness to pay for any procedure that gives them a chance to influence the selection of public projects. If they are in a regime that gives them influence, they pay more taxes. Tax payments are highest if the tax return is spent in proportion to the support for public projects in the (experimental) population. This effect of legitimacy through participation is independent of the effect of legitimacy through support for the causes to be financed. However, we do not find what policymakers would
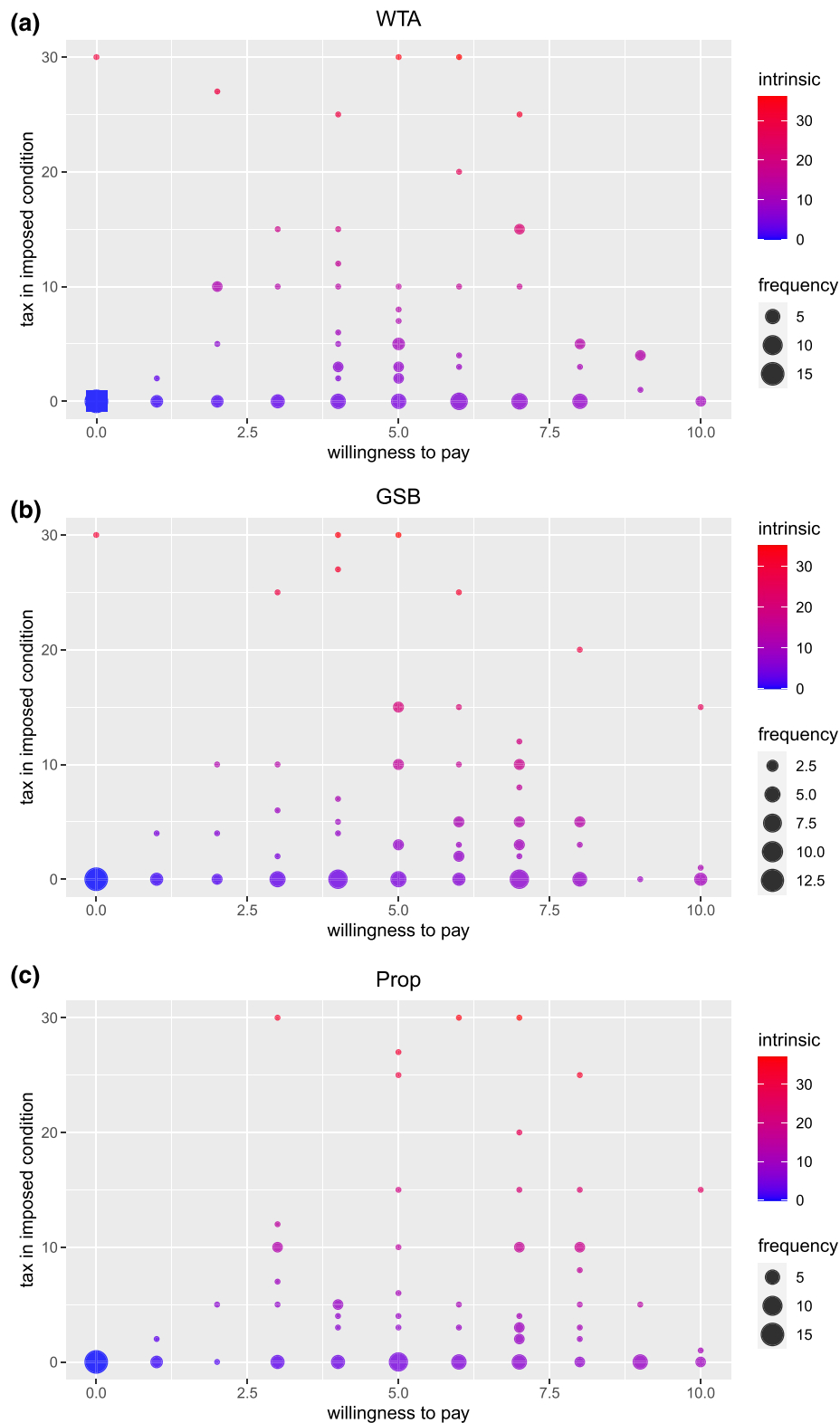
FIGURE 7 Tax payments in imposed condition as a function of willingness to pay for participation. (a) WTA: winner take all. (b) GSB: gold, silver, bronze. (c) Prop: proportional.

likely find most appealing: there is no interaction between democratic procedure and support for public projects. Participation does not compensate for a lack of support for the policies to be implemented.

Willingness to pay for a democratic regime, as well as tax declarations conditional on regime, are explained by the fact that a participant is fine with those public projects that they expect to be most popular. To the extent that this is the

TABLE 7 Tax payments in imposed condition as a function of WTP.

WTP for *Winner-Takes-All*	−0.002 (1.645)
WTP for *Gold-Silver-Bronze*	−0.236 (0.380)
WTP for *Proportional*	0.646* (0.317)
Cons	2.297 (1.645)
N	98

Note: OLS. Standard errors in parentheses.

Abbreviation: WTP, willingness to pay for participation.

*$p < .05$.

case, participation is instrumental. Participants cherish participation, and they react with higher tax declarations, because participation increases the probability of the tax returns being used for purposes they deem acceptable. Hence, for the most part, legitimacy from participation is potential legitimacy from support for the causes to be financed. Participants do predominantly appreciate participation as a channel for achieving an outcome in line with their views. We do, however, also establish a small independent positive effect of participation on the willingness to pay taxes.

Every empirical project has its limitations. The largest advantage of our project is, at the same time, its major limitation. We have run an experiment, enabling us to rely on random assignment to treatments. This approach ensures the reliability of our findings. Secondly, we have eliminated the confounding factor of beliefs and attitudes toward audits and enforcement, further enhancing the reliability of our results. Additionally, participants engage their own money. All of this makes the evidence credible. However, as usual, the increase in internal validity comes at the expense of lower external validity. We believe the ensuing limitation not to be severe, but there are, of course, differences between real-life taxation and our experiment. In real life, much more money is at stake. While we credibly promise no enforcement and anonymity, participants have reason to expect that their choices are recorded, such that the degree of tax evasion becomes known. In reality, governments finance many more than 10 projects. The amount of money that goes to individual projects depends on needs, not only on political will. Additionally, the portfolio of projects is much less transparent. Another limitation is that our participants were mostly students. Arguably, they are better educated and possibly also more public-minded than the general population.

The within-subject design, using the strategy method, gives us more data (each participant decides about each of the three alternative regimes) and cleaner data (we do not have to worry that differences between regimes are driven by unobserved idiosyncratic tendencies of participants, as panel data methods take such idiosyncrasies out of the equation). However, the strategy method is, of course, more "cold," which is why participants might have taken choices less seriously. The strategy method also frames the choice as a comparative one, which it would not if, in political reality, one channel of participation was implemented. We acknowledge this limitation but note that, in general, the strategy method has been found to be behaviorally valid (Fischbacher et al., 2012).

In our instructions, we used the names of colors (but not the colors themselves) to label the alternative regimes (calling *Winner-Takes-All* "YELLOW," *Gold-Silver-Bronze* "GREEN" and *Proportional* "RED"). We deliberately matched colors and regimes such that the colors did not indicate a logical sequence. We cannot completely rule out that the colors' names triggered real-world associations, as with a traffic light. Nonetheless, the data do not suggest that this was the case. Participants did not favor the GREEN regime (*Gold-Silver-Bronze*) nor were they hesitant to choose the RED regime (*Proportional*).

Notwithstanding these limitations, we believe that our experiment makes a valuable contribution to the understanding of taxpayer decisions. In our earlier experiment (Engel et al., 2020), we showed that the willingness to pay taxes was surprisingly pronounced, even if participants were not informed about the use of their money. This experiment made the situation one step more realistic. In the current study, participants are led to understand that public projects need not be met with unanimous support. We show that the degree of support is an important determinant of tax morale. To the degree possible, policymakers should care about making citizens understand why their money is needed. The current study also found that participation increases tax morale. It does so predominantly on the instrumental channel. Even if acceptable outcomes are not guaranteed, individuals appreciate the opportunity to influence the choice of public projects and expect this influence to lead to the definition of projects they deem acceptable. However, in the experiment, the effect of participation is not exclusively instrumental. Participants also have a certain willingness to pay for participation per se and reciprocate with higher tax payments. Hence, legitimacy through voice is

more fragile than legitimacy through support for a policy, but the former remains a real force on which policymakers can count. Democracy matters.

DATA AVAILABILITY STATEMENT

The replication package for this project is available at https://doi.org/10.3886/E193522V4, Engel et al. (2023).

ENDNOTES

[1] This famous slogan by Jonathan Mayhew fueled American independence from England.

[2] For example, a pure system of (party list) proportional representation is implemented in the Netherlands; every party that receives at least 1/150 of all votes is assigned one of the 150 seats in the House of Representatives, Art. 53 I Dutch Constitution.

[3] § 6 III Bundeswahlgesetz.

[4] In our context, "throughput legitimacy," that is, the perceived quality of governmental institutions and process, does not play a role; see Héritier (2003). "Composite democracy in Europe. The role of transparency and access to information." *Journal of European Public Policy* **10**: 814–833, Schmidt (2013). "Democracy and legitimacy in the European Union revisited. Input, output and 'throughput'." *Political Studies* **61**(1): 2–22.

[5] The "leaky bucket" concept, developed by Arthur Okun in 1975, emphasizes the trade-off that redistributive policies make between economic efficiency and income equality. According to Okun, while these policies transfer resources from the wealthy to the poor, they also create significant deadweight loss for society, analogous to "leaks" in a bucket.

[6] There was no difference between the gain and the loss frame version.

[7] In a preregistered pilot, the default choice was the Italian Red Cross, which in a preparatory survey had been the most popular recipient. The attractiveness of the default recipient masked treatment effects. The data from this pilot are available from the authors upon request.

[8] In the case of a tie, the recipient is randomly selected between the entities with the strongest support.

[9] https://osf.io/93sy7/?view_only=b2887d0ac9414d3dad8958e524191064.

[10] We conducted power calculations using G*Power and added this information to the preregistration document.

[11] See footnote.[7]

[12] Due to this, we could not reach 0.25 power, assuming $\alpha = .05$ and $\beta = .8$. Fortunately, we were still able to test our hypotheses, indicating that this potential limitation did not materialize. The only partial exception is in Table 2, but see Table 3.

[13] Equivalent to $8.35 on the first day of the experiment.

[14] Since, in the pilot, we found that tax declarations are highest in the *Gold-Silver-Bronze* condition, we preregistered this ranking of outcomes.

[15] Willingness to pay to be in any of the three alternative regimes was measured for each participant. We thus have panel data. We take the dependence within participants into account by way of a participant random effect. We check, with the help of the Hausman test, whether there is bias, in which case we would have to switch to a model with a participant fixed effect. However, the test never turned out significant.

[16] We have no support for the preregistered hypothesis that *Gold-Silver-Bronze* outperforms the remaining treatments; this expectation was based on a finding from the pilot experiment.

[17] Recall that the treatment coefficients are predictions conditional on the fact that the number of recipients a participant has declared acceptable being 0.

[18] We are grateful to an anonymous referee for bringing this question to our attention.

REFERENCES

Abraham, M., Lorek, K., Richter, F. & Wrede, M. (2017) Collusive tax evasion and social norms. *International Tax and Public Finance*, 24(2), 179–197. Available from: https://doi.org/10.1007/s10797-016-9417-0

Acemoglu, D. & Jackson, M.O. (2017) Social norms and the enforcement of laws. *Journal of the European Economic Association*, 15(2), 245–295.

Ackert, L.F., Martinez-Vazquez, J. & Rider, M. (2007) Social preferences and tax policy design: some experimental evidence. *Economic Inquiry*, 45(3), 487–501. Available from: https://doi.org/10.1111/j.1465-7295.2007.00048.x

Allingham, M.G. & Sandmo, A. (1972) Income tax evasion. A theoretical analysis. *Journal of Public Economics*, 1(3-4), 323–338. Available from: https://doi.org/10.1016/0047-2727(72)90010-2

Alm, J. (1991) A perspective on the experimental analysis of taxpayer reporting. *The Accounting Review*, 66(3), 577–593.

Alm, J., Jackson, B.R. & McKee, M. (1993) Fiscal exchange, collective decision institutions, and tax compliance. *Journal of Economic Behavior & Organization*, 22(3), 285–303. Available from: https://doi.org/10.1016/0167-2681(93)90003-8

Alm, J., McClelland, G.H. & Schulze, W.D. (1999) Changing the social norm of tax compliance by voting. *Kyklos*, 52(2), 141–171. Available from: https://doi.org/10.1111/1467-6435.00079

Andreoni, J., Erard, B. & Feinstein, J. (1998) Tax compliance. *Journal of Economic Literature*, 36(2), 818–860.

Ariel, B. (2012) Deterrence and moral persuasion effects on corporate tax compliance. Findings from a randomized controlled trial. *Criminology*, 50(1), 27–69. Available from: https://doi.org/10.1111/j.1745-9125.2011.00256.x

Arnesen, S. (2017) Legitimacy from decision-making influence and outcome favourability. Results from general population survey experiments. *Political Studies*, 65(1_suppl), 146–161. Available from: https://doi.org/10.1177/0032321716667956

Aronson, E., Wilson, T.D. & Akert, R.M. (2007) *Social psychology*. Upper Saddle River: Pearson Prentice-Hall.

Ausubel, L.M. (2008) Auctions (theory). In: Durlauf, S.N. & Blume, L. (Eds.) *The new Palgrave dictionary of economics*. New York: Springer, pp. 290–302.

Battiston, P. & Gamba, S. (2013) *Is tax compliance a social norm? A field experiment*. University of Milan Bicocca Department of Economics, Management and Statistics Working Paper.

Becker, G.S. (1968) Crime and punishment. An economic approach. *Journal of Political Economy*, 76(2), 169–217. Available from: https://doi.org/10.1086/259394

Becker, G.M., DeGroot, M.H. & Jacob, M. (1964) Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3), 226–232. Available from: https://doi.org/10.1002/bs.3830090304

Beckman, S.R., Formby, J.P. & Smith, W.J. (2004) Efficiency, equity and democracy: experimental evidence on Okun's leaky bucket. In: Cowell, F. (Ed.) *Inequality, welfare and income distribution: experimental approaches*. Bingley: Emerald Group Publishing Limited, pp. 17–42.

Bicchieri, C. (2006) *The grammar of society. The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Blumenthal, M., Christian, C., Slemrod, J. & Smith, M.G. (2001) Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *National Tax Journal*, 54(1), 125–138. Available from: https://doi.org/10.17310/ntj.2001.1.06

Bobek, D.D, M Hageman, A. & Kelliher, C.F. (2013) Analyzing the role of social norms in tax compliance behavior. *Journal of Business Ethics*, 115(3), 451–468. Available from: https://doi.org/10.1007/s10551-012-1390-7

Bobek, D.D. & Hatfield, R.C. (2003) An investigation of the theory of planned behavior and the role of moral obligation in tax compliance. *Behavioral Research in Accounting*, 15(1), 13–38. Available from: https://doi.org/10.2308/bria.2003.15.1.13

Bolton, G.E. & Ockenfels, A. (2000) ERC: a theory of equity, reciprocity and competition. *The American Economic Review*, 90(1), 166–193. Available from: https://doi.org/10.1257/aer.90.1.166

Bordignon, M. (1993) A fairness approach to income tax evasion. *Journal of Public Economics*, 52(3), 345–362. Available from: https://doi.org/10.1016/0047-2727(93)90039-v

Bott, K.M., Cappelen, A.W., Sorensen, E. & Tungodden, B. (2020) You've got mail. A randomised field experiment on tax evasion. *Management Science*, 66(7), 2801–2819. Available from: https://doi.org/10.1287/mnsc.2019.3390

Buijze, R., Engel, C. & Hemels, S. (2017) Insuring your donation. An experiment. *Journal of Empirical Legal Studies*, 14(4), 858–885. Available from: https://doi.org/10.1111/jels.12166

Casal, S., Faillo, M. & Mittone, L. (2022) I want to pay!-Identifying the Unconditional Tax Propensity (UTP). *Journal of Economic Behavior & Organization*, 197, 103–114. Available from: https://doi.org/10.1016/j.jebo.2022.02.023

Casal, S. & Mittone, L. (2016) Social esteem versus social stigma. The role of anonymity in an income reporting game. *Journal of Economic Behavior & Organization*, 124, 55–66. Available from: https://doi.org/10.1016/j.jebo.2015.09.014

Castro, L. & Scartascini, C. (2015) Tax compliance and enforcement in the pampas. Evidence from a field experiment. *Journal of Economic Behavior & Organization*, 116, 65–82. Available from: https://doi.org/10.1016/j.jebo.2015.04.002

Charness, G. & Rabin, M. (2002) Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3), 817–869. Available from: https://doi.org/10.1162/003355302760193904

Chen, D.L., Schonger, M. & Wickens, C. (2016) oTree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88–97. Available from: https://doi.org/10.1016/j.jbef.2015.12.001

Cialdini, R.B., Reno, R.R. & Kallgren, C.A. (1990) A focus theory of normative conduct. Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6), 1015–1026. Available from: https://doi.org/10.1037/0022-3514.58.6.1015

Cialdini, R.B. & Trost, M.R. (1998) Social influence, social norms, conformity, and compliance. In: Gilbert, D.T., Fiske, S.T. & Lindzey, G. (Eds.) *The handbook of social psychology*. Boston: McGraw, pp. 151–192.

Clarke, E.H. (1971) Multipart pricing of public goods. *Public Choice*, 11(1), 17–33. Available from: https://doi.org/10.1007/bf01726210

Desmet, P. & Engel, C. (2021) People are conditional rule followers. *Journal of Economic Psychology*, 85, 102384. Available from: https://doi.org/10.1016/j.joep.2021.102384

Dwenger, N., Kleven, H., Rasul, I. & Rincke, J. (2016) Extrinsic vs intrinsic motivations for tax compliance. Evidence from a randomized field experiment in Germany. *American Economic Journal: Applied Economics*, 8(3), 203–232.

Easton, D. (1965) *A systems analysis of political life*. New York: Wiley.

Ellickson, R.C. (1998) Law and economics discovers social norms. *Journal of Legal Studies*, 27(S2), 537–552. Available from: https://doi.org/10.1086/468033

Engel, C. (2011) Dictator games. A meta-study. *Experimental Economics*, 14(4), 583–610. Available from: https://doi.org/10.1007/s10683-011-9283-7

Engel, C. (2014) Social preferences can make imperfect sanctions work. Evidence from a public good experiment. *Journal of Economic Behavior & Organization*, 108, 343–353. Available from: https://doi.org/10.1016/j.jebo.2014.02.015

Engel, C., Mittone, L. & Morreale, A. (2020) Tax morale and fairness in conflict. An experiment. *Journal of Economic Psychology*, 81, 102314. Available from: https://doi.org/10.1016/j.joep.2020.102314

Engel, C., Mittone, L. & Morreale, A. (2023) *ECIN replication package for "Outcomes or participation? Experimentally testing competing sources of legitimacy for taxation"* [distributor]. Ann Arbor: Inter-University Consortium for Political and Social Research. Available from: https://doi.org/10.3886/E193522V4

Esaiasson, P., Gilljam, M. & Persson, M. (2012) Which decision-making arrangements generate the strongest legitimacy beliefs? Evidence from a randomised field experiment. *European Journal of Political Research*, 51(6), 785–808. Available from: https://doi.org/10.1111/j.1475-6765.2011.02052.x

Esaiasson, P., Persson, M., Gilljam, M. & Lindholm, T. (2019) Reconsidering the role of procedures for decision acceptance. *British Journal of Political Science*, 49(1), 291–314. Available from: https://doi.org/10.1017/s0007123416000508

Falkinger, J. (1995) Tax evasion, consumption of public goods and fairness. *Journal of Economic Psychology*, 16(1), 63–72. Available from: https://doi.org/10.1016/0167-4870(94)00038-c

Fehr, E. & Schmidt, K.M. (1999) A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3), 817–868. Available from: https://doi.org/10.1162/003355399556151

Fehr, E. & Schmidt, K.M. (2006) The economics of fairness, reciprocity and altruism. Experimental evidence and new theories. *Handbook on the Economics of Giving, Reciprocity and Altruism*, 1, 615–691.

Feld, L.P. & Tyran, J.-R. (2002) Tax evasion and voting. An experimental analysis. *Kyklos*, 55(2), 197–222. Available from: https://doi.org/10.1111/1467-6435.00183

Fischbacher, U., Gächter, S. & Quercia, S. (2012) The behavioural validity of the strategy method in public good experiments. *Journal of Economic Psychology*, 33(4), 897–913. Available from: https://doi.org/10.1016/j.joep.2012.04.002

Frey, B.S. (1997) A constitution for knaves crowds out civic virtues. *Economic Journal*, 107(443), 1043–1053. Available from: https://doi.org/10.1111/j.1468-0297.1997.tb00006.x

Groves, T. (1973) Incentives in teams. *Econometrica*, 41(4), 617–631. Available from: https://doi.org/10.2307/1914085

Guala, F. & Mittone, L. (2010) How history and convention create norms. An experimental study. *Journal of Economic Psychology*, 31(4), 749–756. Available from: https://doi.org/10.1016/j.joep.2010.05.009

Güth, W., Levati, V. & Sausgruber, R. (2005) Tax morale and (de-) centralization. An experimental study. *Public Choice*, 125(1-2), 171–188. Available from: https://doi.org/10.1007/s11127-005-3414-7

Hanno, D.M. & Violette, G.R. (1996) An analysis of moral and social influences in tax compliance. *Behavioral Research in Accounting*, 8, 57–75.

Harsanyi, J.C. (1982) Rule utilitarianism, rights, obligations and the theory of rational behavior. In: Harsanyi, J. (Ed.) *Papers in game theory*. Heidelberg: Springer, pp. 235–253.

Héritier, A. (2003) Composite democracy in Europe. The role of transparency and access to information. *Journal of European Public Policy*, 10(5), 814–833. Available from: https://doi.org/10.1080/1350176032000124104

Hibbing, J.R. & Theiss-Morse, E. (2002) *Stealth democracy. Americans′ beliefs about how government should work*. Cambridge University Press.

Holt, C.A. & Laury, S.K. (2002) Risk aversion and incentive effects. *The American Economic Review*, 92(5), 1644–1655. Available from: https://doi.org/10.1257/000282802762024700

Jackson, B.R. & Milliron, V.C. (2002) Tax compliance research: findings, problems and prospects. *Journal of Accounting Literature*, 5, 125–165.

Jimenez, P. & Iyer, G.S. (2016) Tax compliance in a social setting: the influence of social norms, trust in government, and perceived fairness on taxpayer compliance. *Advances in Accounting*, 34, 17–26. Available from: https://doi.org/10.1016/j.adiac.2016.07.001

Kao, C. (2016) No taxation with information on representation. A survey experiment on tax morale in China.

Kimbrough, E.O. & Vostroknutov, A. (2016) Norms make preferences social. *Journal of the European Economic Association*, 14(3), 608–638. Available from: https://doi.org/10.1111/jeea.12152

Kingdon, J.W. (1995) Agenda setting. In: Teodoulou, S. & Cahn, M. (Eds.) *Public policy. The essential readings*. Prentice Hall, pp. 105–113.

Kleine, M., Pascal, L. & Zhurakhovska, L. (2016) Fairness and persuasion: how stakeholder communication affects impartial decision making. *Economics Letters*, 141, 173–176. Available from: https://doi.org/10.1016/j.econlet.2016.02.019

Kleine, M., Langenbach, P. & Zhurakhovska, L. (2017) How voice shapes reactions to impartial decision-makers: an experiment on participation procedures. *Journal of Economic Behavior & Organization*, 143, 241–253. Available from: https://doi.org/10.1016/j.jebo.2017.08.011

Kornhauser, M.E. (2006) A tax morale approach to compliance: recommendations for the IRS. *Florida Tax Review*, 8(6), 599–640. Available from: https://doi.org/10.5744/ftr.2007.1061

Lambert, P.J. (1988) Okun's bucket: a leak and two splashes? *Journal of Economic Studies*, 15(1), 71–78. Available from: https://doi.org/10.1108/eb002666

Lind, E.A. & Tyler, T.R. (1988) *The social psychology of procedural justice*. New York: Plenum Press.

Luttmer, E.F.P. & Singhal, M. (2014) Tax morale. *The Journal of Economic Perspectives*, 28(4), 149–168. Available from: https://doi.org/10.1257/jep.28.4.149

Meyer, D.S. & Minkoff, D.C. (2004) Conceptualizing political opportunity. *Social Forces*, 82(4), 1457–1492. Available from: https://doi.org/10.1353/sof.2004.0082

Neblo, M.A., Esterling, K.M., Kennedy, R.P., Lazer, D.M.J. & Sokhey, A.E. (2010) Who wants to deliberate—and why? *American Political Science Review*, 104(3), 566–583. Available from: https://doi.org/10.1017/s0003055410000298

Norris, P. (1999) *Critical citizens. Global support for democratic government*. Oxford: OUP.

Rochefort, D.A. & Cobb, R.W. (1994) *The politics of problem definition. Shaping the policy agenda*. Lawrence: University Press of Kansas.

Sælen, H. & Kallbekken, S. (2011) A choice experiment on fuel taxation and earmarking in Norway. *Ecological Economics*, 70(11), 2181–2190. Available from: https://doi.org/10.1016/j.ecolecon.2011.06.024

Scharpf, F.W. (1999) *Governing in Europe. Effective and democratic?* Oxford: Oxford University Press.

Schmidt, V.A. (2013) Democracy and legitimacy in the European Union revisited. Input, output and 'throughput'. *Political Studies*, 61(1), 2–22. Available from: https://doi.org/10.1111/j.1467-9248.2012.00962.x

Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J. & Griskevicius, V. (2007) The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5), 429–434. Available from: https://doi.org/10.1111/j.1467-9280.2007.01917.x

Selten, R. (1967) Die Strategiemethode Zur Erforschung Des Eingeschränkt Rationalen Verhaltens Im Rahmen Eines Oligopolexperiments. In: Sauermann, E. (Ed.) *Beiträge Zur Experimentellen Wirtschaftsforschung*. Tübingen: Mohr, pp. 136–168.

Slemrod, J. (1994) Fixing the leak in Okun's bucket optimal tax progressivity when avoidance can be controlled. *Journal of Public Economics*, 55(1), 41–51. Available from: https://doi.org/10.1016/0047-2727(94)90079-5

Stanley, L. & Hartman, T.K. (2018) Tax preferences, fiscal transparency, and the meaning of welfare. An experimental study. *Political Studies*, 66(4), 830–850. Available from: https://doi.org/10.1177/0032321717731661

Strebel, M.A., Kübler, D. & Frank, M. (2019) The importance of input and output legitimacy in democratic governance. Evidence from a population-based survey experiment in four West European countries. *European Journal of Political Research*, 58(2), 488–513. Available from: https://doi.org/10.1111/1475-6765.12293

Torgler, B. (2003) Tax morale, rule-governed behaviour and trust. *Constitutional Political Economy*, 14(2), 119–140. Available from: https://doi.org/10.1023/a:1023643622283

Torgler, B. (2004) Moral suasion. An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. *Economics of Governance*, 5(3), 235–253. Available from: https://doi.org/10.1007/s10101-004-0077-7

Torgler, B. (2005) Tax morale and direct democracy. *European Journal of Political Economy*, 21(2), 525–531. Available from: https://doi.org/10.1016/j.ejpoleco.2004.08.002

Torgler, B. (2007) *Tax compliance and tax morale: a theoretical and empirical analysis*. Edward Elgar Publishing.

Tyler, T.R. (1997) The psychology of legitimacy. *Personality and Social Psychology Review*, 1, 323–344.

Tyler, T.R. (2006) *Why people obey the law*. New Haven: Yale University Press.

Vickrey, W. (1961) Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1), 8–37. Available from: https://doi.org/10.1111/j.1540-6261.1961.tb02789.x

Wenzel, M. & Taylor, N. (2004) An experimental evaluation of tax-reporting schedules. A case of evidence-based tax administration. *Journal of Public Economics*, 88(12), 2785–2799. Available from: https://doi.org/10.1016/j.jpubeco.2003.10.001

Zheng, B., Gu, Y. & Xu, D. (2019) No taxation without representation? An empirical examination of the relationship between democratization and tax compliance. *Chinese Political Science Review*, 5(2), 1–22. Available from: https://doi.org/10.1007/s41111-019-00131-3

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.