Accountability and taxation: Experimental evidence ☆

Ingrid Hoem Sjursen [a,b,*]

[a] *Chr. Michelsen Institute, P.O. Box 6033, N-5892 Bergen, Norway*
[b] *Centre for Applied Research (SNF), Norwegian Centre for Taxation (NoCeT) and FAIR - The Choice Lab, Norwegian School of Economics, Bergen, Norway*

ARTICLE INFO

ABSTRACT

The Rentier State Hypothesis states that taxation promotes government accountability. The argument is that citizens demand more accountability for spending of tax revenue than for spending of windfall revenue (e.g., natural resource revenue and aid). This paper presents causal evidence from a between-subject experiment that tests the effect of taxation on demand for accountability. To investigate the underlying mechanisms of the effect, the design focuses on two main features that distinguish tax from windfall revenue: Tax revenue is produced by citizens' work and has been in their possession before it is collected as tax. The main finding is that taxation causes a higher demand for accountability when both features of taxation are present, consistent with the Rentier State Hypothesis. The paper sheds light on the political economy of government revenues, and contributes to our understanding of how features of the tax system shape behavior.

1. Introduction

The *Rentier State Hypothesis* states that when governments are financed through taxation, as opposed to through windfalls such as natural resources revenue or aid, citizens demand more accountability in government spending. This, in turn, is argued to make politicians more accountable to citizens' demands (Borge et al., 2015; Bräutigam et al., 2008; Karl, 2007; Mahdavy, 1970; Moore et al., 2018; Paler, 2013; Ross, 2001). Thus, the Rentier State Hypothesis suggests that financing public revenue through taxing citizens may promote political accountability and development.

This paper adds to a small, but growing literature providing causal evidence for the positive effect of taxation on accountability, and for the underlying mechanisms explaining this relationship (de la Cuesta et al., 2022; Martin, 2014; Paler, 2013; Weigel, 2020). The focus of this paper is on the first part of the Rentier State Hypothesis, i.e., the effect of taxation on citizens' demand for accountability, with a particular emphasis on identifying the underlying mechanisms of this relationship. Demand for accountability is typically thought to consist of willingness to i) monitor the behavior of decision-makers and ii) impose a cost on decision-makers

when their behavior is not accountable (i.e., when their actions are not responsible and in the best interest of the citizen). This paper studies the latter, i.e., willingness to impose a cost on a decision-maker.

Recent work by de la Cuesta et al. (2022) suggests that psychological feelings of ownership over government revenue is a mechanism through which taxation affects demand for accountability. When citizens contribute tax money, they may have a higher sense of ownership over government revenue, making them more willing to hold the government accountable for misuse. An increase in felt ownership over revenue stemming from paying tax can be thought to affect demand for accountability in three ways. First, it may lead taxpayers to experience misuse of government revenue as a *loss* rather than a reduction in *gain*, which gives higher disutility from misuse ("loss aversion") (de la Cuesta et al., 2022; Kahneman and Tversky, 1979; Martin, 2014). Second, it could affect the norms for what citizens consider fair provision of public goods and services ("fairness norm"). Third, it may affect how aware citizens are and how much they care about receiving what they consider to be a fair amount of public goods and services ("salience of fairness norm").

The present paper sheds light on the three mechanisms by reporting results from a between-subjects online experiment replicating and extending Martin (2014). The participants are randomly assigned to a group with another participant and to the role as a "leader" or a "citizen". The leader decides how much of a group endowment to invest in a common pool and how much to keep for him or herself. The citizen can punish the leader's decision by reducing his or her payoff.[1] The citizen's willingness to punish is the experimental measure of demand for accountability and captures what Fehr et al. (2002, p.3) refer to as *strong negative reciprocity*.[2]

The experiment exogenously varies how the group endowment is financed to capture two major differences between tax and windfall revenue: i) whether it is *earned* money or an *unearned* windfall gain and ii) whether it is financed indirectly through a tax on citizens earnings or directly to the leader, in a two-by-two design with four treatments. In the "Rentier State" treatment, the group endowment is windfall (unearned) and given directly to the leader. In the "Hard Earned & Non-Tax" treatment, the group endowment is produced by the work of the citizen and the leader, and given directly to the leader. In the "Windfall & Tax" treatment, the group endowment is windfall and financed indirectly through a tax on earnings. Finally, in the "Tax State" treatment, the group endowment is produced by the work of the citizen and the leader, and financed indirectly through a tax on earnings. These treatments enable causal tests of the separate effects of the group endowment i) being hard earned ("Hard Earned mechanism") and ii) having been in the possession of the citizen before being collected as tax ("Possession mechanism"), as well as their combined effect, on citizens' willingness to punish.

The two distinguishing features of tax revenue are theorized to increase the salience of fairness considerations to citizens, because they entail that citizens actively contribute to the financing of the group endowment. Increased salience of fairness is, in turn, hypothesized to make citizens care more about how resources are distributed and to increase their willingness to punish the government for unfair behavior. To disentangle *salience* of the fairness norm from a *shift* in the fairness norm, the experiment is designed to keep the fairness norm constant across treatments.

The paper provides evidence that taxation causes a significant increase in citizens' demand for accountability. When the group endowment is produced by the citizens' work and is financed through taxation (Tax State treatment), citizens have a 0.19 standard deviation higher willingness to punish than when the group endowment is windfall and non-tax. The finding is robust across different subgroups of the sample (gender, age, education, employment status, income, political view and political engagement). The study suggests that it is the combination of the Hard Earned and Possession mechanism that makes taxation cause an increase in the demand for accountability. When separately manipulating the Hard Earned mechanism or the Possession mechanism, the estimated effects are not statistically significant, even though the coefficients are in the expected direction. The results add to existing literature by providing evidence on the "Hard Earned mechanism" for the effect of taxation on citizens' demand for accountability, as well as replicating earlier work documenting the "Possession mechanism" in a different experimental set up and sample.

Overall, the results are in line with the Rentier State Hypothesis and suggest that taxation is not only a means to generate government revenue, but may also increase citizens' demand for accountability through a combination of the Hard Earned mechanism and Possession mechanism. The results suggest that, if the government's goal is to increase demand for accountability, the tax system should: i) focus on taxing earned revenues such as labor income rather than unearned income such as prize money from lotteries, and ii) collect taxes in arrears, not as withholding.

The paper is structured as follows: Section 2 discusses related literature. Section 3 describes the experimental design, the sample, and the setting, Section 4 presents the theoretical framework, Section 5 explains the empirical strategy and Section 6 reports the results. Finally, Section 7 provides a discussion of the results and their implications.

2. Related literature

The paper contributes to the broad literature pertaining how source of government revenue affects development. Historical and theoretical accounts argue that taxation forms the basis of a contract between citizens and the state (see for instance Besley (2020), Bräutigam et al. (2008), Levi (1988) and Moore et al. (2018)): Citizens pay their tax conditionally on receiving public services, and can withhold their tax payment (or in other ways exercise demand for accountability) when the government does not deliver - a

---

[1] The experimental design resembles a dictator game with second party punishment, but the amount the leader does not keep is invested in a common pool, not given directly to the citizen. The set up differs some from de la Cuesta et al. (2022) and Martin (2014) in some important aspects - see discussion in Section 3.

[2] Several studies have shown that incentivized lab measures can predict real life behavior, including Barr et al. (2014), who find that cooperation in public good games is correlated with participation in accountability institutions (see also Cohn and Maréchal (2018), Dai et al. (2018) and Hanna and Wang (2017)).

process often referred to as "tax bargaining" (Prichard, 2015). In the absence of taxation, there is no explicit transfer from citizens to the state, and this likely eliminates or weakens the social contract. Several strands of the literature provide empirical evidence in support of the accountability enhancing effects of taxation.

The present paper is related to three recent experiments studying the effect of taxation on demand for accountability by simulating tax payments in the lab. First, the experimental design is based on Martin (2014). She conducts a lab-in-the-field experiment in Uganda and varies whether the group endowment is collected as a tax on a windfall given to the citizen or is given directly to the leader. The leader makes a dictator-decision for how to divide the group endowment between herself and the citizen. Citizens who have been taxed have a significantly higher willingness to punish the leader than citizens who have not been taxed, and it is argued that loss-aversion (Kahneman and Tversky, 1979) is the underlying mechanism. Second, de la Cuesta et al. (2022) conduct lab experiments in Uganda in which they replicate Martin (2014) and implement additional treatments to test the hypothesis that paying tax causes a higher sense of psychological *ownership*, which in turn increases willingness to punish leaders. They find evidence in support of the ownership hypothesis. Third, Paler (2013) shows that a treatment simulating tax payments in a lab-in-the-field experiment in Indonesia increases self-reported willingness to monitor the government and provides evidence suggesting that this effect is explained by citizens having a stronger feeling of ownership over public funds. However, the treatment does not significantly affect actual political participation. The present paper complements Martin (2014), de la Cuesta et al. (2022) and Paler (2013) by shedding further light on the reasons for why taxation causes a higher demand for accountability. First, it investigates the importance of a feature of tax revenue that has not been studied in the previous experiments, namely that tax revenues are *earned* (Hard Earned mechanism), and how this feature interacts with the Possession mechanism. Second, by keeping the fairness norm constant across treatments, the experimental design cleanly disentangles the effect of *salience* of the fairness norm from the effect of a *shift* in the fairness norm.[3]

The paper also relates to the empirical literature using cross- and within-country variation in observational data to investigate the relationship between reliance on tax or windfall revenue and indicators of government accountability such as democratization, provision of public goods, and corruption. This literature shows that good governance is positively associated with reliance on tax revenues and negatively associated with reliance on windfall revenues (Ahmadov, 2014; Besley and Persson, 2014; Gadenne, 2017; Martínez, 2016; Mehlum et al., 2006; Prichard et al., 2018; Ross, 2015).[4] However these studies do not identify a causal relationship between tax or windfall revenue and accountability (with the exceptions of Martínez (2016) and Gadenne (2017)), or provide evidence for the underlying mechanisms explaining it. In particular, the studies based on observational data face two major challenges. First, they typically do not measure citizens' demand for accountability, and thus cannot show that the effect of taxation on government accountability is caused by citizens having a higher demand for accountability. A notable exception is Weigel (2020) who conducts a novel field experiment randomizing property tax registration and collection in D.R. Congo. He combines observational data, surveys and data on political participation and finds that the tax efforts of the government causes a significantly higher attendance in town hall meetings and significantly more submitted evaluations of the government's performance in the taxed neighborhoods. The present study contributes to the literature by providing more causal evidence of the effect of taxation on demand for accountability. Second, observational data studies do not shed light on the reasons why citizens and governments may perceive tax revenues differently. This paper tests whether the fact that money paid in tax is often *earned* by taxpayers contributes to the accountability enhancing effect of taxation. This feature, as well as the fact that money paid as tax has typically been in the taxpayer's possession before being collected as tax is hypothesized to increase demand for accountability by increasing the salience of fairness considerations in government spending.

Finally, the paper relates to the experimental literature on second-party punishment, which shows that parties affected by norm violations are willing to punish the norm violator even when such punishment is costly and yields no material gains now or in the future. Across a large number of studies, receivers in the *ultimatum game* have been shown to punish low offers by rejecting them and leaving both players with zero payoffs (Fehr and Krajbich, 2014). Experiment participants have also been found to punish individuals who contribute less than average in *public goods games* even when group composition changes from period to period and participants only meet each other once (Engelmann and Nikiforakis, 2015; Fehr and Gächter, 2002; Villeval, 2020). The present paper contributes to the literature on second party punishment by investigating such punishment depends on the source of income, and, thus, shedding more light on motivations behind it.

---

3. Experimental design, sample and setting

3.1. Sequence of events

In the experiment, the participants play a two-player investment game in pairs consisting of one citizen and one leader (the instructions for the main parts of the study are provided in Figs. D.1-D.12 in Appendix D). The citizen has the opportunity to punish the leader for his or her investment decision. The set up with one citizen (as opposed to a group of citizens) was chosen to enable identification of the effect of taxation on the individual's punishment decision without the influence of beliefs about punishment behavior of other citizens. Thus, the design abstracts from collective action problems and group dynamics - which are important for some types of citizen demand for accountability (for example the organization of demonstrations and petitions).

The sequence of events for the Rentier State treatment is illustrated in Fig. 1 and can be described as follows. In the first stage (uppermost panel, Fig. 1), the participants are randomly assigned to a treatment, to a pair, and to the role as a either a citizen or a leader. The randomization is done at the individual level. In the second stage (upper middle panel), the citizen and the leader in each pair independently choose whether they would like to work to earn a reward or not. The payment scheme is announced before the participants make their decisions and the instructions emphasize that money the leaders make the decision for belongs to the group (consisting of the citizen and the leader) by using the terms "group money"/"tax revenues" (see Figs. D.4-D.7 in Appendix D for exact wording).[5] To make sure that the participants understand the set-up, they have to correctly answer control questions bout their earnings, the leader's earnings, the source of the group endowment and the task of the leader before they proceeded to the real effort task (see Figs. D.8-D.11 in Appendix D). Only pairs where both participants choose to work are included in the empirical analysis and the remaining explanation will focus on these.[6] The work consists of a five-minute picture categorization task. In the third stage (middle panel), the group receives a group endowment. How the endowment is financed varies between the experimental treatments and is explained in Subsection 3.2 below. In the Rentier State treatment, the group endowment is made up of unearned revenue that does not accrue to the citizen before the leader makes the investment decision for it (non-tax). The size of the group endowment is constant and equal to $2 in all treatments.

In the fourth stage (lower middle panel), the leader decides how much of the group endowment to invest in a common pool. The amount invested is multiplied by a factor of 1.5 and subsequently divided equally between the citizen and the leader.[7] The amount not invested in the common pool is kept by the leader. The leader can invest any share of the group endowment he or she likes in the common pool, in portions of 0.1.

In the fifth stage (lowermost panel), the strategy method is used to elicit the citizen's willingness to punish. For every possible investment decision the leader can make, the citizen decides whether to reduce the payoff to the leader by $0.50 at the cost of $0.05. The citizen's punishment decisions constitute the main outcome of the experiment. The *willingness to punish* is defined as the largest investment share the citizen punishes and proxies demand for accountability: The more willing the citizen is to punish, the higher is his or her demand for accountability.

In the sixth and last stage (not illustrated in the figure), participants answer a set of non-incentivized questions intended to capture individual preferences and socioeconomic background. The purpose of collecting this information is to investigate heterogeneity in treatment effects and the underlying mechanisms for the effect of taxation on the willingness to punish. After completion of the experiment, the participants receive their $1 participation fee within three days. To determine bonus earnings from decisions made in the experiment, citizens and leaders are randomly matched in pairs. The earnings are paid to the participants within three weeks of the completion of the experiment.

3.2. Experimental treatments

The experimental treatments exogenously vary the way in which the group endowment is financed based on two main differences between tax and windfall revenue. First, tax revenue is produced by the citizens' work and is thus hard *earned*, whereas windfall revenue is not related to the citizens' effort at all. Second, the revenue collected through taxation has been in the citizens' possession, whereas windfall revenue accrues directly to the government. The financing of the group endowment is varied along these two dimensions in a 2×2 design, giving rise to four treatments, as illustrated in Table 1. An important feature of the design is that the (post-tax) reward for performing the task and the size of the group endowment are constant across treatments.

The structures of the four treatments are illustrated in Fig. 2 and can be described as follows. In the first treatment, the citizen and the leader each earns $1 from their work. In addition, they receive a $2 windfall as group endowment. The group endowment can be considered windfall revenue because the group receives it independently of the citizen and the leader's work. In the second treatment, the citizen and the leader each earns $1 from their work. In addition, their work produces a $2 ($1 each) group endowment. The group endowment can be considered hard earned revenue because it is dependent on the citizen and the leader's work. In the third

---

[5] The argument for announcing the payment scheme in advance is twofold. First, it is important that the participants feel that their earnings (and the group endowment in the Hard Earned treatments) are financed by their work, and they are more likely to do so if they know that their work would be rewarded, and by how much. Second, it gives the participants time to internalize their earnings before paying tax.

[6] The participants that are recruited for the main study and choose not to work, are always paired with another participant that choose to work. To implement these participants' choices without deception, extra participants are recruited after the main data collection.

[7] The investment has a multiplier to represent the social return of investing in public goods, and is smaller than 2 to make the investment in the common pool costly to the leader.

STAGE 1: RANDOM ASSIGNMENT

Random assignment

Treatment, pair, role

STAGE 2: WORK DECISION

Citizen   Leader

Work   Not work   Work   Not work

STAGE 3: GROUP ENDOWMENT

Windfall

Group endowment
($2)

STAGE 4: LEADER DECISION

Leader

$X $(2 - X)

Common pool   Keep
($X × 1.5)

STAGE 5: CITIZEN DECISION

Citizen

Punish   Not
-$0.05   punish
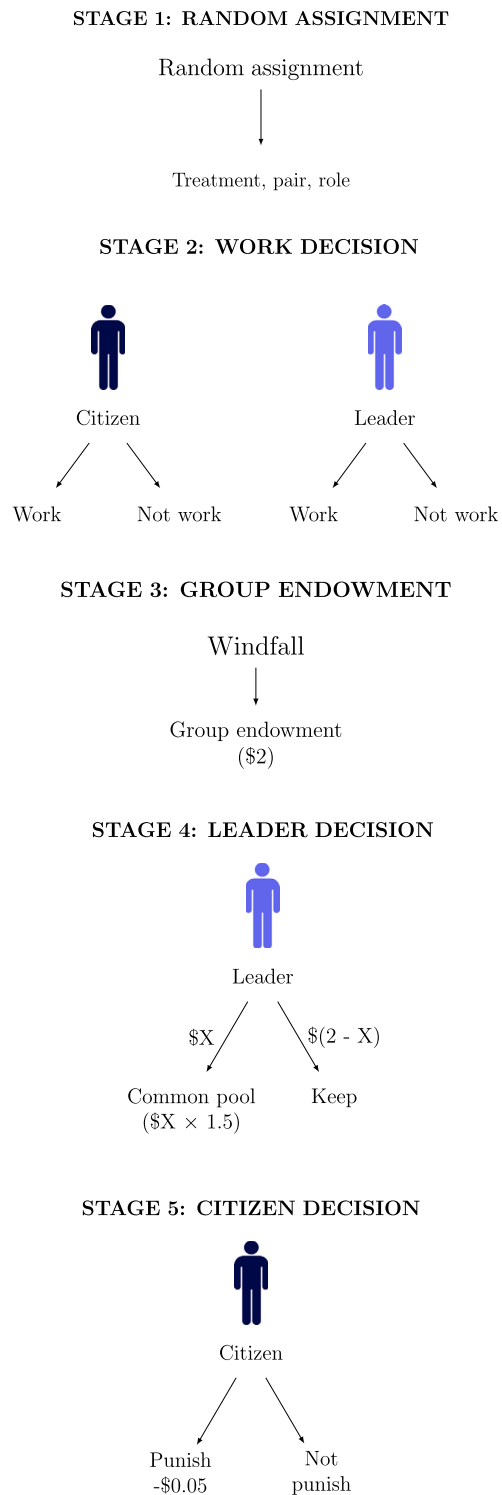
Fig. 1. Sequence of events.

Table 1
Experimental treatments.

	Windfall	**Hard Earned**
Non-tax	Rentier State	Hard Earned & Non-tax
Tax	Winfall & Tax	Tax State
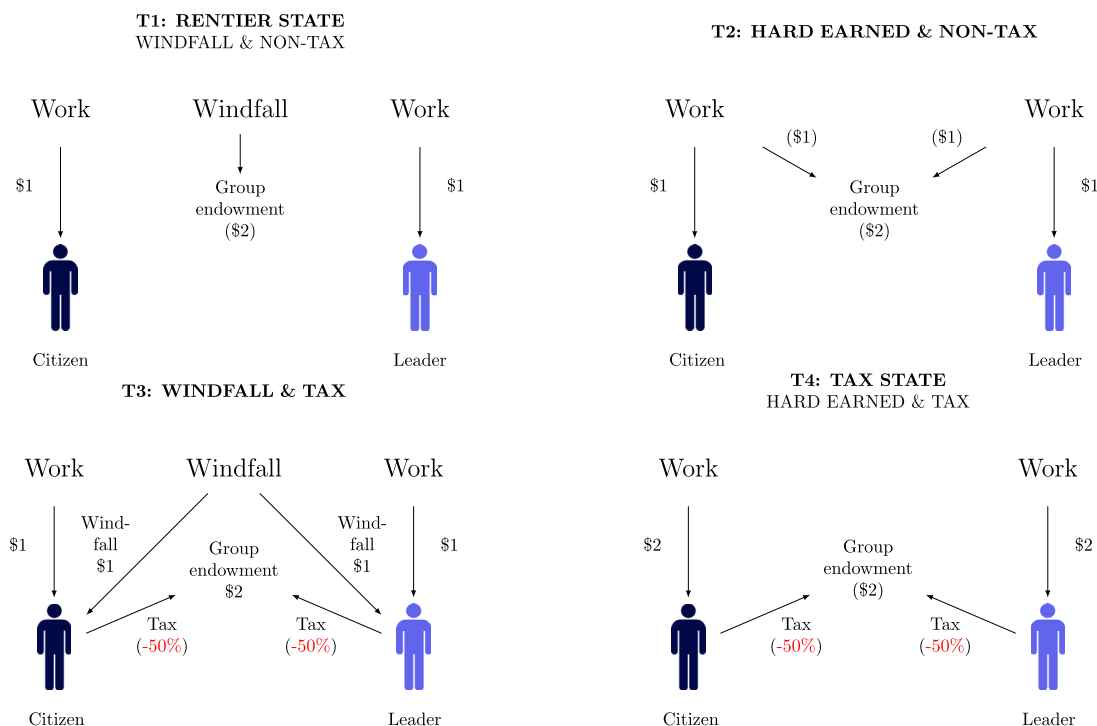
Fig. 2. Treatments: Generation and collection of group endowment.

treatment, the citizen and the leader each earns $1 from their work. In addition, each receives a $1 windfall. The citizen and the leader's total earnings are taxed at 50% and the tax collected finances the $2 group endowment. The group endowment can be considered tax revenue because it is collected from the citizen and the leader's earnings. In the fourth treatment, the citizen and the leader each earns $2 from their work. Their earnings are taxed at 50% and the tax collected finances a $2 group endowment. The group endowment can be considered hard earned revenue because it is produced by the work of the citizen and the leader, as well as financed by tax revenue because it is collected from the citizen and the leader's earnings.[8] The four treatments can be summarized as follows:

- **Rentier State (T1):** Citizen and leader each earn $1. A windfall finances the $2 group endowment.
- **Hard Earned & Non-Tax (T2):** Citizen and leader each earn $1. Their work additionally produces $1 each, financing the $2 group endowment.
- **Windfall & Tax (T3):** Citizen and leader each earn $1. In addition, they each receive a windfall of $1. Their total earnings are taxed at 50%. The $1 tax collected from each finances the $2 group endowment.
- **Tax State (T4):** Citizen and leader each earn $2. Their earnings are taxed at 50%. The $1 tax collected from each finances the $2 group endowment.

The Rentier State treatment can be thought of as a stylization of the situation in a Rentier State where the government is mainly financed by windfall revenue (aid, natural resource revenue) that is not collected through the tax system. The Tax State treatment can be thought of as a stylization of the situation in a tax state, where government revenues are mainly financed through taxing

---

[8] See Appendix B.2.1 for description of a robustness treatment. This treatment resembles the Tax State treatment, but manipulates the Hard Earned Mechanism by making the citizen and the leader work longer (ten minutes instead of five) rather than by making work a more productive activity (pay $2 instead of $1). As the regressions in Table B.12 show, manipulating duration does not yield results different from the main manipulation (productivity) - in line with the pre-specified hypothesis. The two treatments are therefore pooled in all the analyses presented in the paper unless otherwise specified.

hard earned income produced by the citizens' work and the tax is paid in arrears. Comparing the willingness to punish in these two treatments provides a causal test of the Rentier State Hypothesis.

The Hard Earned & Non-Tax treatment can be thought of as a stylization of a state in which government revenues are mainly financed through taxing income, and where income tax is withheld by employers. The Windfall & Tax treatment can be thought of as a stylization of a state in which government revenues are financed through windfall revenues that are distributed to citizens and subsequently (partly) collected as tax, as well as through income tax collected in arrears.[9] Comparing the Rentier State to the Hard Earned & Non-Tax treatment, and the Windfall & Tax to the Tax State treatment provides causal tests for the effect of the group endowment being produced by work when it is not financed through taxation and when it is financed through taxation, respectively. These comparisons test the importance of the Hard Earned mechanism in explaining the Rentier State effect. Similarly, comparing the Rentier State to the Windfall & Tax treatment and the Hard Earned & Non-Tax to the Tax State treatment provides causal tests for the effect of the group endowment being financed through taxation when the money is not produced by work and when the money is produced by work, respectively. These comparisons test the importance of the Possession mechanism in explaining the Rentier State effect.

The treatments are designed to keep what is perceived as a fair investment constant: because the citizen and the leader contribute equally much tax and work in all treatments, the fair thing for the leader to do is always to invest the entire group endowment in the common pool (because this ensures equal pay). The post-experimental survey confirms that this is also how the citizens perceive the situation. When asked how much they think it is fair that the leader invests, the citizens on average answer a share of 0.95 or higher in all treatments.

3.3. Sample and setting

The participants in the experiment were recruited from the online labor market platform, Amazon Mechanical Turk (MTurk). In total, 1996 workers (983 citizens and 1013 leaders) participated in the study. Of these, 110 (54 citizens and 56 leaders) chose not to perform the task and were therefore excluded from the sample. Of the 929 citizens that chose to work, 190 answered the punishment question inconsistently. For these inconsistent punishers, willingness to punish could not be defined.[10] Thus, the main analyses in this paper are based on a sample of 739 citizens. Tables B.1 and B.2 in Appendix B show that there are some significant treatment differences in the share of citizens that choose not to work and the share of citizens that are inconsistent punishers. However, the identified imbalances are not due to non-workers and inconsistent punishers being different from workers and consistent punishers in terms of observable variables (see Table B.3, Table B.4 and Table B.5). This mitigates the concern that differential selection of consistent punishers is driving the treatment effects. It is also worth noting that there is no differential selection of inconsistent punishers between Rentier State and Tax State, which is the main comparison. As a robustness check, the main specifications are also estimated using the number of times punished as the dependent variable, which can be defined for all 929 citizens who chose to work.

Descriptive statistics for the main sample are reported in Table 2. Columns (1)-(5) show means for background variables for each of the four treatments and for the pooled sample, respectively. Column (6) shows the p-value for testing the hypothesis that there is no difference in the background characteristics between each of treatments T2-T4 and the Rentier State treatment. The table shows that the share of males, and the share of respondents that are more politically conservative than the median of the sample, are significantly different between treatments (see Table B.3 in Appendix B for a more detailed balance check). The sample is otherwise balanced.

3.4. Comparison to previous lab experiments

The experimental design builds on Martin (2014) (and de la Cuesta et al. (2022)), but differs in some aspects. Most importantly, the present experiment varies the degree to which the group endowment is earned or not, whereas the money from which citizens pay tax in Martin (2014) and de la Cuesta et al. (2022) is unearned.[11] This study can therefore shed light on whether and how the fact many taxpayers work hard to earn the money the pay tax from shapes their demand for accountability and if it interacts with the "Possession mechanism" studied by Martin (2014) and de la Cuesta et al. (2022).

Second, the present experiment differs in the way the group endowment is generated: in the tax treatments, both the citizen and the leader pay tax, whereas in Martin (2014) and de la Cuesta et al. (2022) the tax is only collected from the citizen. As explained above, this design choice was made to keep what is perceived as a fair investment - the fairness norm - constant across treatment (the citizen and the leader contribute equally much in all treatments, so the leader should always invest everything). In contrast, when

---

[9] The Windfall & Tax treatment corresponds less closely to the actual situation in many countries than the other three treatments. The closest comparison is probably the State of Alaska in the US, where all residents receive a an annual payment from the dividend of oil revenue of approximately $1 000-$3 000 (State of Alaska, 2023). This annual payment is not large enough for residents to get by on, so they still also need to earn their income through wages. An important difference between the experimental set-up and Alaska is that residents in Alaska do not pay taxes on the oil revenue payment or their income, whereas citizens in the Windfall & Tax treatment pay a 50% tax on both incomes.

[10] A citizen's punishment behavior is defined as inconsistent if, for any level of leader investment, the citizen does not punish that investment, but does punish at least one higher level of investment. There are no good theoretical reasons for why citizens should not punish low, but punish higher investments, and it is difficult to know how to analyze the data from inconsistent punishers.

[11] A related literature shows that individual consumption and investment decisions are affected by both the source of income and how this income is framed/labeled, but has to my knowledge not investigated effects on willingness to punish contributions to a public good (see Hvide et al. (2019) for an overview).

Table 2
Background by treatment.

	Windfall & Non-Tax	Hard Earned & Non-Tax	Windfall & Tax	Hard Earned & Tax	Total	F-test
Male	0.53 (0.04)	0.47 (0.04)	0.57 (0.05)	0.44 (0.03)	0.49 (0.02)	0.09
Above median age	0.50 (0.04)	0.46 (0.04)	0.46 (0.04)	0.49 (0.03)	0.48 (0.02)	0.72
Above median education	0.56 (0.04)	0.53 (0.04)	0.56 (0.04)	0.57 (0.03)	0.56 (0.02)	0.69
Employed full-time	0.60 (0.04)	0.58 (0.04)	0.56 (0.04)	0.57 (0.03)	0.58 (0.02)	0.96
Above median income	0.47 (0.04)	0.51 (0.04)	0.46 (0.04)	0.46 (0.03)	0.47 (0.02)	0.60
Conservative	0.45 (0.04)	0.41 (0.04)	0.34 (0.04)	0.49 (0.03)	0.44 (0.02)	0.01
More politically engaged	0.50 (0.04)	0.42 (0.04)	0.53 (0.04)	0.46 (0.03)	0.47 (0.02)	0.19
Observations	163	158	134	284	739	

Note: The table provides mean values (standard error) for background characteristics in the four respective treatments and for the whole sample. The F-test column provides the p-value for an F-test of no difference in means between the four treatments. *Variables:* "Male": indicator variable taking the value of 1 for males, "Above median age": indicator variable taking the value of 1 if individual is older than the median age of the sample (34 years), "Above median education": indicator taking the value of 1 for individuals who have a 4-year degree or higher education, "Employed full-time": indicator variable for individuals who are full-time employees, "Above median income": indicator variable taking the value of 1 for individuals that have an individual yearly income of USD 40 000 or more, "Conservative": indicator variable taking the value of 1 for individuals that rate themselves 6 or higher on a scale from 0 (strongly liberal) to 10 (strongly conservative) and "More politically engaged" is an indicator variable taking the value of 1 for individuals that report to have participated in more than the median number (two) of political activities during the last year.

only the citizen pays tax in the tax treatment, but the citizen and the leader contribute equally (little) in the non-tax treatment, the citizens' perception of how much is a fair share to receive may differ between the two treatments. More specifically, citizens in the tax treatment may feel entitled to a larger share of the group endowment than in the non-tax treatment, because the endowment is generated by taxes they pay. Thus, in Martin (2014) a higher willingness to punish in the tax treatment could be driven by both a change in the fairness norm and an increased salience of the fairness norm. To investigate the effect of taxation on the fairness norm, Martin (2014) therefore implements a second experiment in which a third, unaffected, party makes the punishment decisions. She finds no difference in punishment behavior of third parties in the tax and non-tax treatment, and argues that this rules out a shift in the fairness norm. This test does not rule out that taxation shifts the fairness norm of affected citizens, but de la Cuesta et al. (2022), who implements the same experimental design, rule out a shift in the fairness norm as the mechanism for the effect of taxation on punishment by measuring citizens' beliefs about fairness. Based on this, the magnitude of an effect of the tax treatments identified here may be expected to be similar to those of Martin (2014) and de la Cuesta et al. (2022).

Third, in the present design, the leaders makes an investment in a common pool which benefits both themselves and the citizens, rather than a transfer decision. This makes it less costly for leaders to share the group endowment with the citizens than in Martin (2014) and de la Cuesta et al. (2022). The investment game was chosen to mimic the social return of investing tax revenue in public goods, but is not expected to give different results than a design in which the leader makes a transfer choice.

Fourth, the experiment is conducted online rather than in person. As a consequence of this, the participants in the tax treatments do not physically have the money they pay taxes from in their possession before it is collected from them, but instead see the amount on the instruction screen. This may make the tax treatment weaker than in-person design, and could potentially lead to smaller treatment effects than previous experiments.

Fifth, the participants in the experiment are Americans, whereas Martin (2014) and de la Cuesta et al. (2022) conduct their experiments with Ugandans. These samples have different socio-economic backgrounds, live in different institutional environments, and are likely to have different knowledge and experiences with paying taxes and may therefore respond differently to the tax treatments.

4. Theoretical framework

This section describes the theoretical framework for the citizen's punishment decision guiding the experimental design.

The one-shot structure and the anonymity of the participants entail that there are no strategic incentives for punishment. However, a large empirical literature has documented that people are willing to sacrifice monetary payoff to punish unfair behavior, even when

they get no monetary or strategic benefits from doing so (Fehr and Fischbacher, 2003, 2004a,b; Henrich et al., 2006; Henrich and Henrich, 2007). Fehr et al. (2002) refer to such behavior as *strong negative reciprocity*.[12]

To capture strong negative reciprocity, this paper assumes that the citizen derives utility from punishing the leader when the leader is perceived to be unfair.

$$V_c = R + \frac{3}{4}I - p \times c - \gamma_c \left(m_c - I\right)^2 + (p \times \beta_c)\left[min\{0, m_c - I\}^2\right]. \tag{1}$$

The utility consists of three components. The first, $R + \frac{3}{4}I - (p \times c)$, represents the citizen's monetary payoff. R is the reward for performing the real effort task, I is the leader's investment in the common pool, $p \in \{0, 1\}$ is an indicator variable for the citizen's binary punishment decision that is taking the value of one if the citizen punishes the leader and zero otherwise, and c is the cost of punishing the leader. The second component, $\gamma_c\left(m_c - I\right)^2$, represents the citizen's disutility from unfair investments. m_c is the citizen's fairness norm and specifies the investment share that he or she perceives to be fair. Thus, $m_c - I$ is how much the leader's investment deviates from the citizen's fairness norm. The parameter γ_c is the weight the citizen attaches to disutility from deviations from the fairness norm. It is assumed that this disutility is increasing in the investment's deviation (both positive and negative) from the fairness norm. The third component, $(p \times \beta_c)\left[min\{0, m_c - I\}^2\right]$, is the citizen's utility from punishing unfair investments. The individual parameter β_c is the weight the citizen attaches to utility from punishment of unfair investments. It is assumed that the citizen only derives utility from punishment of negative deviations from the fairness norm.

Equation (1) shows that the citizen's utility from punishment depends positively on how much the leader's investment deviates from the fairness norm, $m_c - I$, and on the individual parameter β_c, and negatively on the cost of punishment, c. The citizen's punishment behavior is characterized by:

$$c < \beta_c\left[min\{0, m_c - I\}^2\right] \qquad\qquad \text{Punish} \tag{2a}$$

$$c = \beta_c\left[min\{0, m_c - I\}^2\right] \qquad \text{Indifferent, randomize} \tag{2b}$$

$$c > \beta_c\left[min\{0, m_c - I\}^2\right] \qquad\qquad \text{Not punish} \tag{2c}$$

As Equations (2a)-(2c) show, for any given investment level (I), three parameters influence the citizen's punishment in the model; the cost of punishment (c), the level of investment the citizen perceives as fair (m_c), and the utility he or she derives from punishment of unfair investments (β_c). Since the experiment is designed to keep the fairness norm constant across treatments, we assume that the fairness norm, m_c, is the same in all treatments and entail that the leader should invest everything in the common pool. Thus, the effects of treatments must work through changes in the utility derived from punishing unfair investments, β_c.

The basic idea behind the design is that taxation increases the salience of fairness considerations to the citizens. The hypothesis is that, under taxation, citizens actively contribute to the group endowment in two ways; they work hard to earn the money that finances it, and they have the tax money in their possession before it is collected. This active contribution is thought to attract the citizens' attention to the fairness norm, i.e., that the leader should invest everything, which in turn is hypothesized to increase the citizen's utility from punishment (β_c), leading to higher willingness to punish.[13] In the absence of taxation (Rentier State treatment), the citizen does not actively contribute to the group endowment. Even though the citizens' fairness norm is the same as under taxation, the lack of active contribution to the group endowment is hypothesized to make this fairness norm less salient and the citizens less willing to punish unfair investments. Accordingly, the theory predicts willingness to punish to be lower in the Rentier State treatment than in the other three treatments (see Appendix C for a theoretical framework for the leader's investment decision).

The present framework complements de la Cuesta et al. (2022), who theorize that the effect of tax on punishment is caused by a heightened feeling of ownership over the group endowment, by suggesting that stronger feelings of ownership over revenue can affect demand for accountability through fairness considerations.[14]

---

[12] Evolutionary theory posits that the existence of strong negative reciprocity is due to cooperative behavior increasing the likelihood for survival, and that natural selection therefore has favored individuals that are intuitively cooperative and trustworthy (Fehr et al., 2002; Fehr and Fischbacher, 2004a; Jordan et al., 2016; Rand and Nowak, 2013).

[13] This idea is related to a series of recent theoretical papers showing that alternatives that are more salient, i.e., that stand out more because they are different or unusual, receive more of the decision maker's attention and thus influence their decisions more relative to less salient alternatives. Bordalo et al. (2012) theorize that lotteries with payoffs that stand out are over-weighed relative to their objective probabilities in decision-making. Bordalo et al. (2013a) extend this model to demand for risky assets and Bordalo et al. (2013b, 2016) and Köszegi and Szeidl (2013) formulate more general models for salience and choice. These models generally focus on the salience of different alternatives in one particular choice setting. The present argument is slightly different and focuses on how the salience of one particular feature of the choice situation differs between different settings (with and without taxation). The basic mechanism should be the same: When our attention is drawn to a particular product or feature of the choice situation, we care about that product or feature, and put more emphasis on it, compared to when our attention is not drawn to it.

[14] Martin (2014) suggests another theoretical mechanism, namely *loss-aversion* (Kahneman and Tversky, 1979). Appendix B.4 elaborates on the theory and presents exploratory analysis testing one version of the mechanism in the present data.

5. Empirical strategy

5.1. Main analysis

To investigate the effect of taxation on willingness to punish, the following specification is estimated:

$$y_i = \alpha + \beta_{T2}T2_i + \beta_{T3}T3_i + \beta_{T4}T4_i + \beta_X X_i + \beta_Z Z_i + \epsilon_i \tag{3}$$

y_i is the standardized willingness to punish for the citizen in pair i (mean of 0, standard deviation of 1).[15] α is a constant, $T2_i$, $T3_i$ and $T4_i$ are indicator variables taking the value of one for individuals in the Hard Earned & Non-Tax, Windfall & Tax and Tax State treatments, respectively. X_i is a vector of the background variables of the citizen (indicator variables for male, age above median, education above median level, full-time employee and income above median), Z_i is a vector of political view and engagement (indicator variable for above median politically conservative and for above median engaged in political activities) and ϵ_i is an error term. Three versions of Equation (3) are estimated; one with the treatment indicator variables only; one with the treatment indicator variables and X_i; and one with the treatment indicator variables, X_i and Z_i, all OLS regressions with robust standard errors. The Rentier State treatment is the reference category, and the main parameters of interest in the estimation of Equation (3) are β_1, β_2 and β_3. These were all hypothesized to be positive and significant in the pre-analysis plan.

To investigate whether the treatments affect willingness to punish on the intensive or extensive margin Equation (3) is also estimated for two indicator variables: An indicator variable for high punishment, *High punishment dummy*, taking the value of one for citizens that punish all, but the highest level of investment, and zero otherwise; and an indicator variable for positive punishment, *Positive punishment dummy*, taking the value of one for citizens that punish at least one investment share and zero for those who never punish.[16]

As a robustness check, Equation (3) is additionally estimated for two alternative definitions of willingness to punish as the dependent variable that can be defined for all citizens (including those who punish inconsistently): *Number of times punished*, which takes a value between 0 and 11, and *punishment of 0.9 investment share*, which is a dummy taking the value of one for individuals that punish investment shares of 0.9 and zero otherwise. As an additional robustness check, Equation (3) is estimated for *the lowest share invested for which the citizen does not punish the leader* as the dependent variable.

5.2. Leader decisions

To investigate the effect of treatments on leaders' investment decisions, Equation (3) is estimated using the standardized share invested in the common pool as the dependent variable.

6. Results

6.1. Descriptive findings on punishment

The main outcome of interest is the citizens' willingness to pay a cost to punish the leaders' investment decisions by reducing their payoff. The willingness to punish is measured using the strategy method: For every possible investment decision the leader can make, the citizen decides whether he or she would like to punish that decision. The leader can invest as much of the group endowment as he or she likes, in portions of 0.1. Thus, the citizen decides whether to punish an investment share of 0, an investment share of 0.1, an investment share of 0.2 and so on up until an investment share of 1. Fig. 3 illustrates the share of citizens who punish each of the possible investments the leader could make pooled across treatments. It shows that greater investments by the leader make citizens less willing to punish. Whereas more than half of the citizens punish leaders that invested 0 in the common pool, only 16 percent of citizens punish leaders that invest 0.9.

The decisions illustrated in Fig. 3 are used to identify the highest investment level for which the citizen punishes the leader, which is used as a measure for the willingness to punish. To illustrate, imagine that a citizen decides to punish the leader if he or she invests a share of 0.5 or less in the common pool, but does not punish higher investment shares. Then the highest investment share for which the citizen punishes, and the measure for his or her willingness to punish (i.e. his or her demand for accountability), is 0.5. Fig. 4 shows the distribution of willingness to punish in the total sample. It shows that about half of the citizens never punish the leader or punish 0 investments only. The remaining 55% have a positive willingness to punish for at least one level of leader investment, most of whom (about 37% of the sample) are willing to punish high investments, meaning that the highest investment share they punish is between 60% and 90%. A small share (about 15% of the sample) are only willing to punish low investments, meaning that the highest investment share they punish is between 0% and 50%. A minority of three percent always punish the leader, even in the case when he or she invests 100% of the group endowment.

---

[15] The empirical strategy, and the hypotheses to be tested, were specified in the pre-analysis plan submitted to the American Economic Association Randomized Control Trials Registry before the data collection. https://www.socialscienceregistry.org/trials/2233, registration number AEARCTR-0002233. The paper mainly follows the pre-analysis plan, with some minor deviations and a few additional specifications. See Appendix A for an overview.

[16] The regression tables reported in the main paper indicates significance levels for both unadjusted p-vales and p-values adjusted for multiple hypothesis testing. The adjusted p-values are calculated using the False Discovery Rate procedure described by Benjamini and Hochberg (1995).
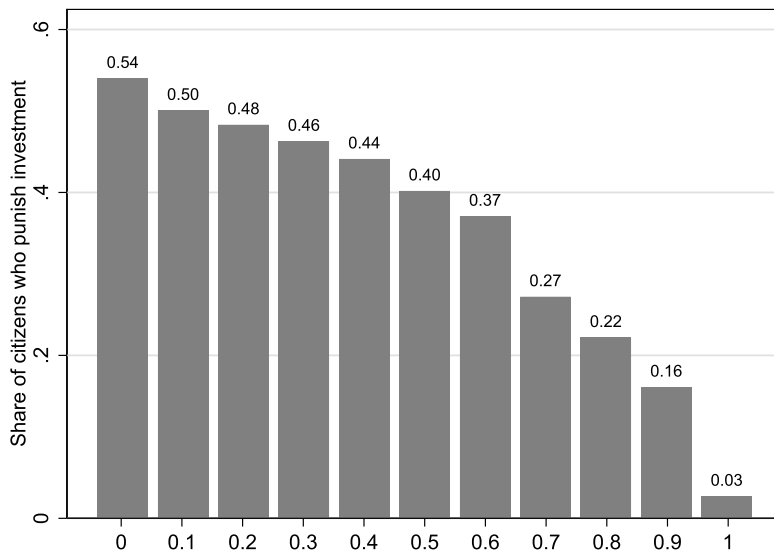
Fig. 3. Punishment by leader's investment level.
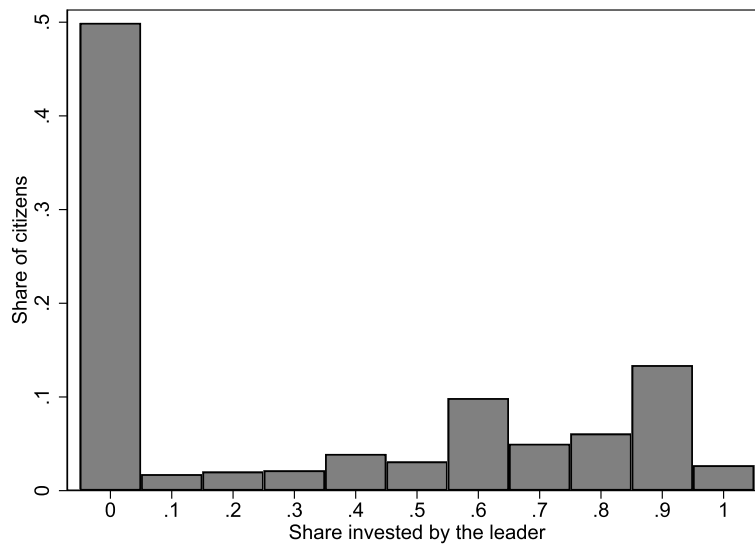
Fig. 4. Distribution of willingness to punish.

6.2. Main analysis: the effect of taxation on punishment

The experiment was designed to study the effect of taxation on punishment. Fig. 5 graphs mean willingness to punish in the four different treatments. It shows that compared to the Rentier State treatment, willingness to punish is higher in all the other treatments. It also shows that the Tax State treatment causes a significant increase in citizens' willingness to punish. Going from a situation where the group endowment is windfall and non-tax to a situation where it is produced by the participants' work and financed through taxation, increases the willingness to punish from 0.30 to 0.37 (p = 0.060). The magnitude of the effect is equivalent to a 0.19 standard deviation increase in the willingness to punish.

Table 3 reports results from estimation of Equation (3) and tests how willingness to punish is affected by the group revenue being hard earned and collected as tax, in a regression framework. Columns (1)-(3) report results for willingness to punish measured by a standardized version of the 11-point punishment scale (mean 0, standard deviation 1). Column (1) reports estimates of Equation (3) with treatment variables and the interaction term only, Column (2) reports estimates for a specification that includes background variables (gender, age, education, income and occupation) and Column (3) reports results for the full specification where indicators for political view and political engagement are also included. The Rentier State treatment is the reference category in all columns. Focusing on the full specification in Column (3), the table shows that the descriptive finding is replicated in the regression analysis, and is robust to inclusion of background variables. Going from the Rentier State to the Tax State treatment significantly increases the
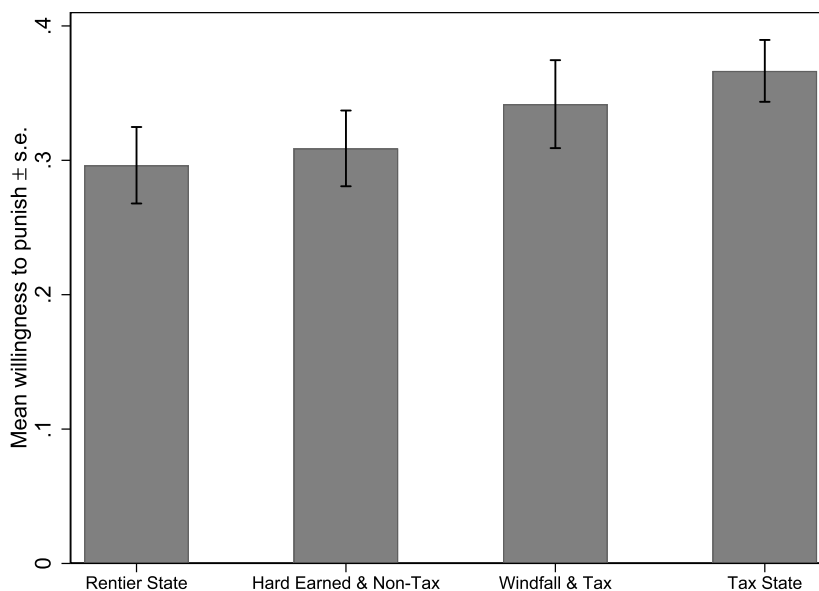
Fig. 5. Effect of taxation on punishment.

willingness to punish (as indicated by the positive and significant coefficient of "Tax State", p = 0.058). The table also shows that estimated effects of the Hard Earned & Non-tax and Windfall & Tax treatments are positive, but not statistically significant. In terms of background variables, only employment status is significantly correlated with punishment; citizens that are employed full-time have a significantly higher willingness to punish than those who are not employed full-time.

Columns (4)-(6) report estimates of Equation (3) where the dependent variable is a standardized version (mean 0, standard deviation 1) of an indicator variable taking the value of one for individuals that punish very high shares invested (0.9): Column (4) reports estimates from the specification with treatment variables only, and Columns (5) and (6) sequentially add the background and political variables. The main result from Column (3) holds when investigating the effect on the indicator for high willingness to punish in Column (6). Table B.8 in Appendix B reports the estimated regressions for the dependent variables that can also be defined for citizens that are inconsistent punishers: number of times punished (Column (1)-(3)) and a dummy variable taking the value of one for citizens that punish an investment share of 0.9 (Column (4)-(6)). It shows that the main results are robust to these specifications. This mitigates the concern the concern the identified treatment effect in Table 3 is driven by non-random attrition in the form of inconsistent punishers.[17]

Based on the regressions in Table 3, the following main results can be formulated:

Result 1 *(Rentier State Hypothesis).* The Tax State treatment significantly increases the willingness to punish compared to the Rentier State treatment ($\beta_{T4} = 0.186$, p = 0.058, Column (3)).

Result 2 *(Mechanisms).* The Hard Earned & Non-Tax and Windfall & Tax treatments do not cause a significant increase in citizens willingness to punish *(Hard Earned & Non-Tax: $\beta_{T2} = 0.040$, p = 0.715, Windfall & Tax: $\beta_{T3} = 0.129$, p = 0.269, Column (3)).*

Result 1 provides evidence in support of the Rentier State Hypothesis. When the financing of the group endowment is characterized by the two distinguishing factors of tax revenue; (i) the money has been in the citizens' possession and then been collected through taxation and (ii) the money is earned income, the citizens have a higher willingness to punish.[18]

Result 2 shows that it is the combination of the Hard Earned and Possession mechanism that makes taxation cause an increase in the demand for accountability. When separately manipulating the Hard Earned mechanism or the Possession mechanism, the estimated effects are not statistically significant, even though the coefficients are in the expected direction.[19]

A direct comparison of effect sizes with de la Cuesta et al. (2022) and Martin (2014) is not possible because the papers do not report standardized outcomes or standard deviations. However, one can compare differences in non-standardized means measured

---

[17] Table B.9 in Appendix B shows the estimated regressions in Table 3, with classical standard errors.

[18] The estimated regressions for the robustness check of the Hard Earned mechanism are reported in Table B.12 in Appendix B. It shows that there is no significant difference in willingness to punish between citizens in the five- and 10-minute Hard Earned treatments.

[19] Table B.11 in Appendix B reports regression results for an indicator variable taking the value of one if the citizen punishes some investment share in Columns (1)-(3). It shows that the treatments have no significant effects on the likelihood of a citizen punishing the leader. Thus, the Tax State treatment increases citizens' willingness to punish very high investments, but does not increase the likelihood of the citizen punishing the leader.

Table 3
Effect of taxation on punishment.

	Willingness to punish			High willingness to punish		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned & Non-tax	0.034	0.031	0.040	0.101	0.103	0.100
	(0.107)	(0.107)	(0.108)	(0.102)	(0.103)	(0.102)
Windfall & Tax	0.121	0.133	0.129	0.168	0.168	0.158
	(0.116)	(0.116)	(0.117)	(0.112)	(0.113)	(0.114)
Tax State	0.188*	0.182⋆	0.186⋆	0.185**	0.176⋆	0.180*
	(0.098)	(0.098)	(0.098)	(0.092)	(0.092)	(0.092)
Male		-0.113	-0.108		-0.056	-0.051
		(0.069)	(0.069)		(0.068)	(0.069)
Above median age		0.001	-0.003		0.002	0.010
		(0.074)	(0.074)		(0.075)	(0.075)
Above median education		-0.004	-0.013		0.104	0.093
		(0.080)	(0.080)		(0.082)	(0.083)
Employed full time		0.137*	0.139*		-0.058	-0.063
		(0.081)	(0.081)		(0.080)	(0.079)
Above median income		0.006	0.007		-0.090	-0.083
		(0.084)	(0.084)		(0.083)	(0.083)
Conservative			0.006			-0.089
			(0.077)			(0.075)
More politically engaged			0.120			0.005
			(0.076)			(0.074)
Constant	-0.101	-0.125	-0.185	-0.123*	-0.076	-0.039
	(0.076)	(0.109)	(0.122)	(0.067)	(0.108)	(0.120)
Observations	739	739	739	739	739	739
R^2	0.006	0.013	0.017	0.005	0.011	0.013

Note: The table reports regressions of the standardized value of willingness to punish on the treatment variables "Hard Earned & Non-tax", "Windfall & Tax" and "Tax State", and background variables. Columns (1)-(3) show the results for the 11-point scale definition of willingness to punish (takes values 0.0, 0.1, 0.2,...,1). Columns (4)-(6) show the results for a dummy taking the value of one if the individual punishes all, but the highest level of investment. "Male": indicator variable taking the value of 1 for males, "Above median age": indicator variable taking the value of 1 if individual is older than the median age of the sample (34 years), "Above median education": indicator taking the value of 1 for individuals who have a 4-year degree or higher education, "Employed full-time": indicator variable for individuals who are full-time employees, "Above median income": indicator variable taking the value of 1 for individuals that have an individual yearly income of USD 40 000 or more, "Conservative": indicator variable taking the value of 1 for individuals that rate themselves 6 or higher on a scale from 0 (strongly liberal) to 10 (strongly conservative) and "More politically engaged": indicator variable taking the value of 1 for individuals that report to have participated in more than the median number (2) of political activities during the last year. Robust standard errors are reported in parenthesis. Statistical significance is indicated by $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ for unadjusted p-values and by $⋆p < 0.1$, $⋆⋆p < 0.05$, $⋆⋆⋆ p < 0.01$ for p-values adjusted for multiple hypotheses testing. When correcting for multiple hypotheses testing, the two outcomes are grouped together in one family for each of the three treatments.

as share of the group endowment transferred to the citizen/invested in the common pool. These are reported in Table B.19 in Appendix B and shows that the difference in means between the Rentier State and the Windfall & Tax treatment (the most direct comparison for the other papers) of 0.043, is smaller than in Martin (2014) (0.054), but larger than in de la Cuesta et al. (2022) (0.022). It can also be noted that the difference in means between the Rentier State and the Tax State treatment (in which the tax paid is also earned by the participants) of is 0.070, is larger than the treatment difference identified in both de la Cuesta et al. (2022) and Martin (2014).

The pre-specified heterogeneity analysis is reported in Tables B.14 and B.15 and shows that i) the Tax State coefficient is positive (but not always significant) for all the reference groups except non-full time employees, and ii) that there is no significant difference in the effect of the Tax State treatment between the subgroups of the sample, except between full-time employed and non-full-time employed. Thus, the effect of taxation on punishment identified in the main analysis does not seem to be driven by a particular subgroup of the sample.

6.3. Leader decisions

This subsection investigates the leaders' decisions for investment in the common pool.[20] The upper panel of Fig. 6 illustrates the distribution of leader investments: each bar indicates the share of leaders that invested the given share in the common pool. It shows that the majority of leaders, 60%, invest everything. About 25% keep the entire group endowment to themselves and 4% invest half.

---

[20] See Table B.18 in Appendix B for effect of treatment on background variables.
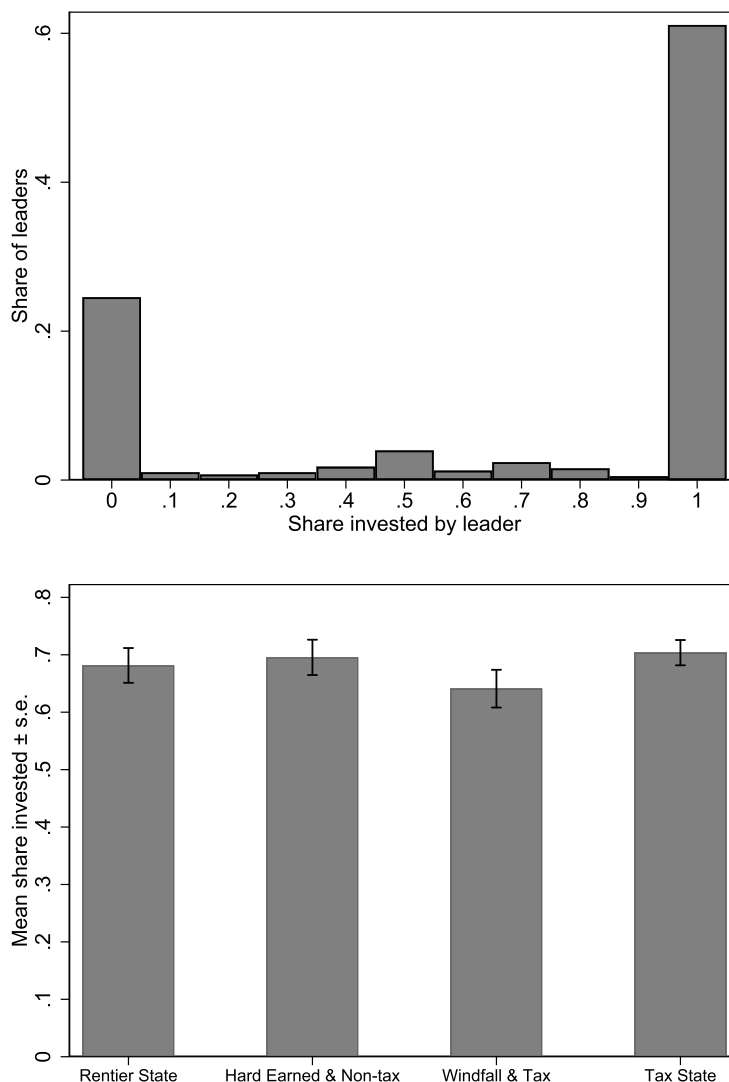
Fig. 6. Leader investment.

The lower panel of the figure illustrates the effect of taxation on leader investments. It shows that mean investment in the Windfall & Tax treatment is lower than in the other treatments (0.64 vs. 0.68-0.70). However, the difference in investment is only significant when comparing the Windfall & Tax treatment to the Tax State treatment.

Table 4 reports the estimated results of Equation (3) taking the standardized value of the leader's investment as the dependent variable.[21] The regressions replicate the descriptive result from Fig. 6, for both the 11-point scale and the binary definition of leader investment (taking the value of one for leaders that invest everything in the common pool and zero otherwise) and show no significant difference in share invested between the Rentier State and the other treatments. In terms of background variables, older and more politically engaged leaders invest more, and full-time employees and more politically conservative leaders invest less in the common pool.

7. Discussion

This paper studies the effect of taxation on demand for accountability in an economic experiment. The experimental design focuses on two features that distinguish tax from other sources of government revenue; tax revenue is produced by the citizens' work and has been in their possession before being collected as tax. The paper shows that when revenue is *earned* and has been in the citizens' possession, this causes a higher demand for accountability, measured as citizens' willingness to punish the leader's investment decision for the group endowment. Citizens have a significantly higher willingness to punish when the group endowment

---

[21] Table B.10 in Appendix B shows the estimated regressions in Table 4, with classical standard errors.

Table 4
Effect of taxation on share invested in the common pool (leaders).

	Share invested			High investment dummy		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned & Non-tax	0.033	0.043	0.050	0.114	0.119	0.124
	(0.100)	(0.099)	(0.099)	(0.101)	(0.100)	(0.100)
Windfall & Tax	-0.094	-0.099	-0.104	-0.040	-0.053	-0.059
	(0.103)	(0.105)	(0.105)	(0.105)	(0.106)	(0.106)
Tax State	0.052	0.058	0.062	0.108	0.105	0.109
	(0.087)	(0.086)	(0.086)	(0.089)	(0.088)	(0.088)
Male		-0.028	-0.035		0.044	0.037
		(0.065)	(0.064)		(0.064)	(0.063)
Above median age		0.227***	0.241***		0.275***	0.287***
		(0.065)	(0.065)		(0.065)	(0.065)
Above median education		0.100	0.074		0.088	0.060
		(0.068)	(0.068)		(0.067)	(0.067)
Employed full time		-0.213***	-0.223***		-0.192***	-0.203***
		(0.070)	(0.070)		(0.070)	(0.070)
Above median income		-0.061	-0.053		-0.071	-0.066
		(0.073)	(0.072)		(0.072)	(0.071)
Conservative			-0.133**			-0.110*
			(0.067)			(0.066)
More politically engaged			0.131**			0.161**
			(0.066)			(0.066)
Constant	-0.009	-0.019	-0.012	-0.059	-0.123	-0.136
	(0.070)	(0.094)	(0.102)	(0.073)	(0.096)	(0.104)
Observations	957	957	957	957	957	957
R^2	0.003	0.032	0.042	0.004	0.035	0.046

Note: The table reports regressions of the standardized value of share invested by the leader on the treatment variables "Hard Earned & Non-tax", "Windfall & Tax" and "Tax State" and a set of explanatory variables (see Table 3 for definitions of variables). Columns (1)-(3) show the results for the discrete definition of share invested (takes values 0.1, 0.2,...,1). Columns (4)-(6) show the results for a dummy taking the value of one if the leader invests everything in the common pool. Robust standard errors are reported in parenthesis. Statistical significance is indicated by $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ for unadjusted p-values and by $\star p < 0.1$, $\star\star p < 0.05$, $\star\star\star p < 0.01$ for p-values adjusted for multiple hypotheses testing. When correcting for multiple hypotheses testing, the two outcomes are grouped together in one family for each of the three treatments.

is produced by work and financed through taxation compared to when it is windfall and non-tax. This finding provides evidence in support of the Rentier State Hypothesis. The heterogeneity analysis shows that the finding is robust across all, but one, subgroups of the sample. Furthermore, it is the combination of the Hard Earned mechanism and the Possession mechanism that is creating the effect on demand for accountability. Each of them alone results in positive, but not statistically significant, effects. However, the direction of the result for the Possession mechanism is in line with de la Cuesta et al. (2022) and Martin (2014) and the results in the present paper can be interpreted as a replication of these findings with a different experimental setup and different sample.

The paper suggests that fairness considerations may contribute to explain why taxes may increase the demand for accountability, which is consistent with the ownership mechanism suggested by de la Cuesta et al. (2022). Specifically, the argument is that fairness considerations become more salient when taxes are imposed on income that is hard earned and has been in the possession of the citizen. In line with this, the paper shows that taxation causes more punishment in a setting where the fairness norm is constant across treatments, meaning that the observed treatment effects are not driven by a shift in fairness norms. Thus, the paper suggests a potential reason for why increased feelings of ownership over revenue, the more fundamental mechanism suggested by de la Cuesta et al. (2022), caused by taxation may increase punishment.

The results provide important implications for our understanding of citizens' accountability behavior and for policy. Taxing citizens is not only an instrument for generating government revenue, it may also promote demand for accountability, which in turn is generally assumed to increase government accountability. A tax system designed to enhance demand for accountability should have the following features: First, tax should be paid in arrears, not as withholding. A related argument is that direct taxes may cause more demand for accountability than indirect taxes such as VAT, because they are more visible (Bräutigam et al., 2008; de la Cuesta et al., 2021). Second, tax should mainly be levied on income that is *earned*, such as employment income, not on unearned (windfall) income such as lottery prizes. The argument is that when paying taxes in arrears and on earned income, citizens actively contribute to tax revenues, which increases the salience of fairness in resource distribution and that this in turn increases demand for accountability. However, collecting tax in arrears might conflict with other policy goals, such as increasing tax compliance (see for instance Bagchi and Duöek (2021), Dhami and al Nowaihi (2007), Engström et al. (2015), or Vossler et al. (2021)). Using observational data to study the effect of a shift in collection from arrears-based to withholding, or vice-versa, on both demand for accountability and compliance is an interesting avenue for further research. Furthermore, the results imply that if the government's aim is to improve accountability

to all groups in the population, the tax base should be broadly defined and also include those who, from a revenue perspective, it is not worth collecting taxes from.

This paper studies the effect of taxation on demand for accountability in an experimental setting that tightly controls for factors that are not the focus of the study, but that might affect demand for accountability. This enables a clean causal test of the effect of taxation on demand for accountability and of the micro-founded mechanisms that might explain it. Testing these mechanisms in a field setting is an interesting topic for future research. Furthermore, testing the causal effect of taxation and the effect of demand for accountability on accountability in government spending will shed more light on the political effects of taxation.

Declaration of competing interest

I hereby declare that I have no relevant or material financial interests that relate to the research described in this paper.

Data availability

Data will be made available on request.

Appendix A. Deviations from pre-analysis plan

A.1. Deviations from pre-analysis plan

Theoretical framework
The theoretical framework presented in the paper states that taxation causes a higher demand for accountability through an increased salience of fairness. The framework presented in the pre-analysis plan included both the salience of fairness considerations and deviations from reference payoff.

Dependent variable
The dependent variable was defined as "Highest investment share not punished" in the pre-analysis plan, but is defined as "Highest investment share punished" in the specifications reported in the paper. The latter definition was chosen for pedagogical reasons, but does not qualitatively or significantly change the results.

Main specification
The pre-analysis plan specified the main regression to be estimated with indicator variables for Tax treatments (equal to one for participants in Windfall & Tax and Tax State), Hard Earned treatments (equal to one for participants in Hard Earned & Non-tax and Tax State) and an interaction term between these two indicators. The specification reported in the paper is with indicator variables for the Windfall & Tax, Hard Earned & Non-tax and Tax State treatments. This specification was preferred for pedagogical reasons and does not affect the results.

The pre-analysis plan also specified that negative reciprocity, positive reciprocity and risk aversion should be included as controls in the main regressions. However, as noted in the pre-analysis plan, the questions to measure these preferences were asked *after* the treatment to avoid priming effects. This design choice comes at the cost that the measures may be affected by the treatments. Table B.3 shows this indeed the case for negative reciprocity, which is significantly higher for citizens in the Tax Treatment. Thus, negative reciprocity is not included as a co-variate in the regressions reported in this paper. Positive reciprocity and risk aversion are not included as co-variates in the main tables of the paper for pedagogical reasons, but as shown i Table B.13, their inclusions do not change the results.

Heterogeneity analysis
Employment status was not pre-specified as a dimension for the heterogeneity analysis in the pre-analysis plan, but results for this are presented in the paper. This is because the emission of employment status from the pre-analysis plan was not intentional. Given that the heterogeneity results are mostly nulls, the regression tables for the heterogeneity analysis are reported in Appendix B.3.

Sample size and power
The pre-analysis plan specified a sample size of 1 000 citizens - 200 in each treatment, which, based on previous literature would give an estimated power of 86% to identify an effect size of 0.3 standard deviation at 5% level of significance. 983 citizens were recruited for the experiment. Out of these 54 did not perform the real effort task, and 190 were inconsistent punishers, resulting in a sample of 739 citizens for the main analysis (162 in Rentier State, 158 in Hard Earned & Windfall, 134 in Windall & Tax and 284 in Tax State). This gives an estimated power of 72%, 76% and 86%, to identify a 0.3 standard deviation effect between the Rentier State and the Hard Earned & Windfall, the Windall & Tax and the Tax State treatment, respectively.

A.2. Additional analyses reported

Citizen behavior
The following results are reported in the paper, but were not specified in the pre-analysis plan and should be considered exploratory.

- Regressions with binary dependent variable reported in Columns (4)-(6) in Table 3, and in Columns (1)-(3) in Table B.11
- Regressions in Table B.8

Leader behavior
The pre-analysis specified regressions to investigate the effect of the Hard Earned treatments on the share invested in the common pool in the leader. The paper additionally reports treatment effects of the Tax treatments.

Appendix B. Additional figures and tables

B.1. Balance tables

Table B.1
Effect of treatments on decision to work.

	(1)	(2)
Hard Earned & Non-Tax	0.029	0.032*
	(0.018)	(0.018)
Windfall & Tax	−0.062**	−0.059**
	(0.027)	(0.027)
Tax State	0.001	0.003
	(0.019)	(0.019)
Male		−0.010
		(0.014)
Above median age		0.027*
		(0.014)
Above median education		0.006
		(0.017)
Employed full-time		−0.011
		(0.016)
Above median income		0.004
		(0.018)
Conservative		−0.017
		(0.016)
More politically engaged		0.040***
		(0.015)
Constant	0.951***	0.932***
	(0.015)	(0.022)
P-value F-test	0.001***	0.001***
Observations	983	983
R^2	0.017	0.033

Robust standard errors in parentheses
$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Note: Regressions showing the relationship between the decision to work (dependent variable is indicator variable taking the value of one for citizen who choose to work) and treatments. The Rentier State treatment is the reference category in all regressions. "Tax State" is an indicator variable taking the value of one for individuals in the Tax State treatment, "Hard Earned & Non-Tax" is an indicator variable taking the value of one for individuals in the Hard Earned & Non-Tax treatment and "Windfall & Tax" is an indicator variable taking the value of one for individual in the Windfall & Non-Tax treatment. "P-value of F-test" reports the p-value for the test of "Tax State" = "Hard Earned & Non-Tax" = "Windfall & Tax". The background variables are as defined in Table 3.

Table B.2

Effect of treatments on inconsistent punishment behavior.

	(1)	(2)
Hard Earned & Non-Tax	0.026	0.024
	(0.038)	(0.039)
Windfall & Tax	0.043	0.047
	(0.041)	(0.041)
Tax State	0.079**	0.073**
	(0.034)	(0.035)
Male		−0.040
		(0.025)
Above median age		−0.005
		(0.027)
Above median education		−0.020
		(0.029)
Employed full-time		0.027
		(0.030)
Above median income		−0.006
		(0.031)
Conservative		0.036
		(0.028)
More politically engaged		−0.009
		(0.027)
Constant	0.160***	0.170***
	(0.026)	(0.043)
P-value F-test	0.300	0.384
Observations	929	929
R^2	0.006	0.012

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Regressions showing the relationship between switching the wrong way (not punish low levels of investments and punish high levels of investments) or multiple times and treatments. The dependent variable is indicator variable taking the value of one for citizens who switch the wrong way. The Rentier State treatment is the reference category in all regressions. "Tax State" is an indicator variable taking the value of one for individuals in the Tax State treatment, "Hard Earned & Non-Tax" is an indicator variable taking the value of one for individuals in the Hard Earned & Non-Tax treatment and "Windfall & Tax" is an indicator variable taking the value of one for individual in the Windfall & Non-Tax treatment. "P-value of F-test" reports the p-value for the test of "Tax State" = "Hard Earned & No-Tax" = "Windfall & Tax". The background variables are as defined in Table 3.

Table B.3
Effect of treatments on background variables, main sample of citizens.

	Male	Older	Higher educ	Full time	Higher income	Conservative	Pol. engaged	High neg. reci.	High pos. reci.	Risk averse	Loss averse	More upset
Hard Earned & Non-Tax	-0.053	-0.047	-0.027	-0.025	0.040	-0.043	-0.073	0.020	0.009	-0.047	-0.107[*]	-0.017
	(0.058)	(0.056)	(0.056)	(0.055)	(0.056)	(0.055)	(0.056)	(0.056)	(0.056)	(0.056)	(0.055)	(0.056)
Windfall & Tax	0.047	-0.048	0.001	-0.042	-0.004	-0.105[*]	0.033	0.035	0.090	0.008	-0.018	0.082
	(0.068)	(0.058)	(0.058)	(0.058)	(0.058)	(0.057)	(0.058)	(0.058)	(0.058)	(0.058)	(0.057)	(0.058)
Tax State	-0.087[*]	-0.014	0.016	-0.031	-0.009	0.042	-0.032	0.114[**]	-0.013	0.024	-0.097[**]	0.073
	(0.051)	(0.049)	(0.049)	(0.048)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)
Constant	0.528[***]	0.503[***]	0.558[***]	0.601[***]	0.466[***]	0.448[***]	0.497[***]	0.436[***]	0.485[***]	0.521[***]	0.607[***]	0.448[***]
	(0.042)	(0.039)	(0.039)	(0.038)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.038)	(0.039)
P-value of F-test	0.088[*]	0.718	0.694	0.961	0.602	0.012[**]	0.193	0.110	0.139	0.353	0.233	0.131
Observations	739	739	739	739	739	739	739	739	739	739	739	739

Robust standard errors in parentheses, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Note: The table displays coefficients from estimated regressions of each of the background indicator variables as the dependent variable on indicator variables for the treatments (see Table 3 for definitions of treatments, Male, Older, Higher educ. Full time, Higher income, Conservative and Pol. engaged). The Rentier State treatment is the reference category. "High negative Reciprocity": indicator variable taking the value of one if negative reciprocity index is higher than the median. In accordance with Falk et al. (2018), the index for negative reciprocity was constructed using the following three items *If I am treated very unjustly, I will take revenge at the first occasion, even if it is a cost to do so* (0 = Strongly disagree, ..., 10 = Strongly agree)), *How willing are you to punish someone who treats you unfairly, even if there may be costs for you?* (0 = Completely unwilling, ..., 10 = Completely willing) and *How willing are you to punish someone who treats others unfairly, even if there may be costs for you?* (0 = Completely unwilling, ..., 10 = Completely willing). From these questions, the index was constructed as follows: negative reciprocity = 0.2631× Willingness to punish if oneself is treated unfairly + 0.2631× Willingness punish if other is treated unfairly + 0.3738× Willingness to take revenge. These weights were obtained by Falk et al. (2018) by running a regression of observed risk behavior in the lab on responses to these survey questions. "High positive reciprocity": indicator variable taking the value of 1 for individuals that answered a number higher than the median reported (10) to the question *When someone does me a favor, I am willing to return it* (0 = Strongly disagree, ..., 10 = Strongly agree), and zero otherwise. "Risk averse": indicator variable taking the value of 1 for individuals answering a number lower than median reported (7) to the question *How willing or unwilling you are to take risks?* (0 = Completely unwilling, ..., 10 = Completely willing), and zero otherwise. "Loss averse": indicator variable taking the value of one for individuals that to the following question *If you could choose between the following hypothetical scenarios, which would you choose? 1. Lottery: win $80 with probability 1/2, lose $50 with probability 1/2 or 2. Receive $0 for sure* answer "2. Receive $0 for sure". "More upset": indicator variable taking the value of 1 for individuals answering a number higher than the median reported (6) to the question *How upset would you feel if the leader invests less than what you think is fair in the common pool?* (0 = Not upset at all, ..., 10 = Very upset). Reported p-values tests the hypothesis that all the treatments have the same effect on the background indicator variables.

Table B.4
Effect of treatments on background variables, all citizens who worked.

	Male	Older	Higher educ	Full time	Higher income	Conservative	Pol. engaged	High neg. reci.	High pos. reci.	Risk averse	Loss averse	More upset
Hard Earned & Non-Tax	-0.041	-0.031	-0.052	-0.010	0.067	-0.026	-0.098[*]	-0.005	0.015	-0.046	-0.093[*]	-0.010
	(0.052)	(0.051)	(0.051)	(0.050)	(0.051)	(0.050)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.048)
Windfall & Tax	0.067	-0.070	-0.002	0.013	0.015	-0.079	-0.038	0.030	0.089[*]	0.006	-0.048	0.030
	(0.060)	(0.053)	(0.052)	(0.052)	(0.053)	(0.052)	(0.053)	(0.053)	(0.053)	(0.053)	(0.052)	(0.051)
Tax State	-0.089[*]	-0.030	-0.009	-0.006	0.021	0.048	-0.059	0.070	0.016	0.028	-0.105[**]	0.097[**]
	(0.046)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.043)
Constant	0.510[***]	0.510[***]	0.567[***]	0.582[***]	0.443[***]	0.448[***]	0.521[***]	0.464[***]	0.459[***]	0.500[***]	0.577[***]	0.345[***]
	(0.038)	(0.036)	(0.036)	(0.035)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.034)
P-value of F-test	0.012[**]	0.669	0.558	0.892	0.513	0.015[**]	0.503	0.229	0.249	0.241	0.457	0.034[**]
Observations	929	929	929	929	929	929	929	929	929	929	929	929

Robust standard errors in parentheses, [*] $p < 0.10$, [**] $p < 0.05$, [***] $p < 0.01$

Note: The table displays coefficients from estimated regressions of each of the background indicator variables as the dependent variable on indicator variables for the treatments (see Table 3 for definitions of treatments, Male, Older, Higher educ. Full time, Higher income, Conservative and Pol. engaged). The Rentier State treatment is the reference category. "High negative Reciprocity": indicator variable taking the value of one if negative reciprocity index is higher than the median. In accordance with Falk et al. (2018), the index for negative reciprocity was constructed using the following three items *If I am treated very unjustly, I will take revenge at the first occasion, even if it is a cost to do so* (0 = Strongly disagree, ..., 10 = Strongly agree)), *How willing are you to punish someone who treats you unfairly, even if there may be costs for you?* (0 = Completely unwilling, ..., 10 = Completely willing) and *How willing are you to punish someone who treats others unfairly, even if there may be costs for you?* (0 = Completely unwilling, ..., 10 = Completely willing). From these questions, the index was constructed as follows: negative reciprocity = 0.2631× Willingness to punish if oneself is treated unfairly + 0.2631× Willingness punish if other is treated unfairly + 0.3738× Willingness to take revenge. These weights were obtained by Falk et al. (2018) by running a regression of observed risk behavior in the lab on responses to these survey questions. "High positive reciprocity": indicator variable taking the value of 1 for individuals that answered a number higher than the median reported (10) to the question *When someone does me a favor, I am willing to return it* (0 = Strongly disagree, ..., 10 = Strongly agree), and zero otherwise. "Risk averse": indicator variable taking the value of 1 for individuals answering a number lower than median reported (7) to the question *How willing or unwilling you are to take risks?* (0 = Completely unwilling, ..., 10 = Completely willing), and zero otherwise. "Loss averse": indicator variable taking the value of one for individuals that to the following question *If you could choose between the following hypothetical scenarios, which would you choose? 1. Lottery: win $80 with probability 1/2, lose $50 with probability 1/2 or 2. Receive $0 for sure* answer "2. Receive $0 for sure". "More upset": indicator variable taking the value of 1 for individuals answering a number higher than the median reported (6) to the question *How upset would you feel if the leader invests less than what you think is fair in the common pool?* (0 = Not upset at all, ..., 10 = Very upset). Reported p-values tests the hypothesis that all the treatments have the same effect on the background indicator variables.

Table B.5
Effect of treatments on background variables, all citizens.

	Male	Older	Higher educ	Full time	Higher income	Conservative	Pol. engaged	High neg. reci.	High pos. reci.	Risk averse	Loss averse	More upset
Hard Earned & Non-Tax	−0.040	−0.025	−0.058	−0.022	0.054	−0.027	−0.091[*]	0.009	0.024	-0.055	-0.079	-0.029
	(0.051)	(0.050)	(0.050)	(0.049)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.048)
Windfall & Tax	0.067	−0.071	−0.023	0.005	−0.001	−0.086[*]	−0.045	0.046	0.089[*]	0.014	-0.040	0.067
	(0.056)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.051)	(0.050)	(0.050)
Tax State	−0.089[**]	−0.026	−0.017	−0.027	0.003	0.047	−0.051	0.067	0.018	0.031	-0.089[**]	0.092[**]
	(0.045)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.042)
Constant	0.515[***]	0.500[***]	0.574[***]	0.598[***]	0.456[***]	0.456[***]	0.505[***]	0.451[***]	0.451[***]	0.510[***]	0.564[***]	0.377[***]
	(0.037)	(0.035)	(0.035)	(0.034)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.034)
P-value of F-test	0.007[***]	0.542	0.632	0.758	0.445	0.008[***]	0.583	0.409	0.259	0.135	0.533	0.016[**]
Observations	983	983	983	983	983	983	983	983	983	983	983	983

Robust standard errors in parentheses, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Note: The table displays coefficients from estimated regressions of each of the background indicator variables as the dependent variable on indicator variables for the treatments (see Table 3 for definitions of treatments, Male, Older, Higher educ. Full time, Higher income, Conservative and Pol. engaged). The Rentier State treatment is the reference category. "High negative Reciprocity": indicator variable taking the value of one if negative reciprocity index is higher than the median. In accordance with Falk et al. (2018), the index for negative reciprocity was constructed using the following three items *If I am treated very unjustly, I will take revenge at the first occasion, even if it is a cost to do so* (0 = Strongly disagree, …, 10 = Strongly agree)), *How willing are you to punish someone who treats you unfairly, even if there may be costs for you?* (0 = Completely unwilling, …, 10 = Completely willing) and *How willing are you to punish someone who treats others unfairly, even if there may be costs for you?* (0 = Completely unwilling, …, 10 = Completely willing). From these questions, the index was constructed as follows: negative reciprocity = 0.2631× Willingness to punish if oneself is treated unfairly + 0.2631× Willingness punish if other is treated unfairly + 0.3738× Willingness to take revenge. These weights were obtained by Falk et al. (2018) by running a regression of observed risk behavior in the lab on responses to these survey questions. "High positive reciprocity": indicator variable taking the value of 1 for individuals that answered a number higher than the median reported (10) to the question *When someone does me a favor, I am willing to return it* (0 = Strongly disagree, …, 10 = Strongly agree), and zero otherwise. "Risk averse": indicator variable taking the value of 1 for individuals answering a number lower than median reported (7) to the question *How willing or unwilling you are to take risks?* (0 = Completely unwilling, …, 10 = Completely willing), and zero otherwise. "Loss averse": indicator variable taking the value of one for individuals that to the following question *If you could choose between the following hypothetical scenarios, which would you choose? 1. Lottery: win $80 with probability 1/2, lose $50 with probability 1/2 or 2. Receive $0 for sure* answer "2. Receive $0 for sure". "Above median upset": indicator variable taking the value of 1 for individuals answering a number higher than the median reported (6) to the question *How upset would you feel if the leader invests less than what you think is fair in the common pool?* (0 = Not upset at all, …, 10 = Very upset). Reported p-values tests the hypothesis that all the treatments have the same effect on the background indicator variables.

Table B.6

Descriptive statistics for willingness to punish by treatment.

	Mean	Std. Dev	Min	Max	Observations
Windfall & Non-Tax	0.296	0.363	0.0	1.0	163
Hard Earned & Non-Tax	0.309	0.354	0.0	0.9	158
Windfall & Tax	0.342	0.379	0.0	1.0	134
Hard Earned & Tax	0.367	0.388	0.0	1.0	284
Total	0.334	0.374	0.0	1.0	739

Note: The table reports mean, standard deviation, min, max and number of observations for citizens' willingness to punish measured as the largest investment share the citizen punishes (takes values 0.0, 0.1, 0.2, ..., 1), for each of the four treatments as well as the total sample.

B.2. Robustness

Table B.7

Effect of taxation on punishment, non-standardized outcome.

	Willingness to punish			High willingness to punish		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned & Non-tax	0.013	0.011	0.015	0.035	0.035	0.034
	(0.040)	(0.040)	(0.040)	(0.035)	(0.035)	(0.035)
Windfall & Tax	0.045	0.050	0.048	0.057	0.057	0.054
	(0.043)	(0.044)	(0.044)	(0.038)	(0.038)	(0.039)
Tax State	0.070*	0.068*	0.070*	0.063**	0.060*	0.061*
	(0.037)	(0.037)	(0.037)	(0.031)	(0.031)	(0.031)
Male		-0.042	-0.040		-0.019	-0.017
		(0.026)	(0.026)		(0.023)	(0.024)
Above median age		0.000	-0.001		0.001	0.003
		(0.028)	(0.028)		(0.026)	(0.026)
Above median education		-0.002	-0.005		0.035	0.032
		(0.030)	(0.030)		(0.028)	(0.028)
Employed full time		0.051*	0.052*		-0.020	-0.021
		(0.030)	(0.030)		(0.027)	(0.027)
Above median income		0.002	0.003		-0.031	-0.028
		(0.032)	(0.031)		(0.028)	(0.028)
Conservative			0.002			-0.030
			(0.029)			(0.026)
More politically engaged			0.045			0.002
			(0.028)			(0.025)
Constant	0.296***	0.287***	0.265***	0.092***	0.108***	0.121***
	(0.028)	(0.041)	(0.046)	(0.023)	(0.037)	(0.041)
Observations	739	739	739	739	739	739
R^2	0.006	0.013	0.017	0.005	0.011	0.013

Note: The table reports regressions of willingness to punish on the treatment variables "Hard Earned & Non-tax", "Windfall & Tax" and "Tax State", and background variables. Columns (1)-(3) show the results for the 11-point scale definition of willingness to punish (takes values 0.0, 0.1, 0.2,...,1). Columns (4)-(6) show the results for a dummy taking the value of one if the individual punishes all, but the highest level of investment. "Male": indicator variable taking the value of 1 for males, "Above median age": indicator variable taking the value of 1 if individual is older than the median age of the sample (34 years), "Above median education": indicator taking the value of 1 for individuals who have a 4-year degree or higher education, "Employed full-time": indicator variable for individuals who are full-time employees, "Above median income": indicator variable taking the value of 1 for individuals that have an individual yearly income of USD 40 000 or more, "Conservative": indicator variable taking the value of 1 for individuals that rate themselves 6 or higher on a scale from 0 (strongly liberal) to 10 (strongly conservative) and "More politically engaged": indicator variable taking the value of 1 for individuals that report to have participated in more than the median number (2) of political activities during the last year. Robust standard errors are reported in parenthesis. Statistical significance is indicated by $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ for unadjusted p-values

Table B.8

Effect of taxation on punishment, extended sample.

	Number of times punished			High punishment dummy		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned & Non-tax	0.084	0.087	0.099	-0.000	0.006	0.007
	(0.099)	(0.100)	(0.100)	(0.097)	(0.098)	(0.098)
Windfall & Tax	0.093	0.097	0.099	0.035	0.038	0.036
	(0.105)	(0.105)	(0.105)	(0.102)	(0.102)	(0.102)
Tax State	0.146*	0.140	0.150*	0.158*	0.154*	0.156*
	(0.087)	(0.088)	(0.088)	(0.088)	(0.088)	(0.088)
Male		-0.089	-0.082		-0.064	-0.063
		(0.063)	(0.063)		(0.061)	(0.062)
Above median age		-0.007	-0.006		-0.016	-0.014
		(0.066)	(0.067)		(0.067)	(0.067)
Above median education		0.055	0.038		0.072	0.068
		(0.070)	(0.070)		(0.071)	(0.072)
Employed full time		0.159**	0.160**		0.112	0.111
		(0.072)	(0.071)		(0.072)	(0.072)
Above median income		-0.031	-0.026		-0.070	-0.068
		(0.074)	(0.073)		(0.076)	(0.076)
Conservative			-0.035			-0.023
			(0.068)			(0.069)
More politically engaged			0.141**			0.010
			(0.067)			(0.068)
Constant	-0.093	-0.154	-0.209*	-0.070	-0.104	-0.098
	(0.070)	(0.098)	(0.110)	(0.069)	(0.100)	(0.111)
Observations	929	929	929	929	929	929
R^2	0.003	0.011	0.017	0.005	0.011	0.011

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, robust standard error in parenthesis.

Note: The table reports regressions of the standardized value of willingness to punish on the treatment variables "Hard Earned & Non-tax", "Windfall & Tax" and "Tax State" and background variables. Columns (1)-(3) show the results for the number of times the citizen punished the leader (takes values 0, 1, 2, …, 11). Columns (4)-(6) show the results for a dummy taking the value of one if the individual punishes the leader for an investment of 0.9 and zero if the citizen does no punish an investment of 0.9. "Male": indicator variable taking the value of 1 for males, "Above median age": indicator variable taking the value of 1 if individual is older than the median age of the sample (34 years), "Above median education": indicator taking the value of 1 for individuals who have a 4-year degree or higher education, "Employed full-time": indicator variable for individuals who are full-time employees, "Above median income": indicator variable taking the value of 1 for individuals that have an individual yearly income of USD 40 000 or more, "Conservative": indicator variable taking the value of 1 for individuals that rate themselves 6 or higher on a scale from 0 (strongly liberal) to 10 (strongly conservative) and "More politically engaged": indicator variable taking the value of 1 for individuals that report to have participated in more than the median number (2) of political activities during the last year.

Table B.9

Robustness check, main regressions with classical standard errors.

	Willingness to punish			High willingness to punish		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned & Non-tax	0.034	0.031	0.040	0.101	0.103	0.100
	(0.112)	(0.112)	(0.112)	(0.112)	(0.112)	(0.112)
Windfall & Tax	0.121	0.133	0.129	0.168	0.168	0.158
	(0.116)	(0.117)	(0.117)	(0.117)	(0.117)	(0.117)
Tax State	0.188*	0.182*	0.186*	0.185*	0.176*	0.180*
	(0.098)	(0.098)	(0.098)	(0.098)	(0.098)	(0.099)
Male		-0.113	-0.108		-0.056	-0.051
		(0.071)	(0.072)		(0.072)	(0.072)
Above median age		0.001	-0.003		0.002	0.010
		(0.075)	(0.075)		(0.075)	(0.075)
Above median education		-0.004	-0.013		0.104	0.093
		(0.079)	(0.080)		(0.079)	(0.080)
Employed full time		0.137*	0.139*		-0.058	-0.063
		(0.082)	(0.082)		(0.082)	(0.082)
Above median income		0.006	0.007		-0.090	-0.083
		(0.084)	(0.084)		(0.084)	(0.085)
Conservative			0.006			-0.089
			(0.077)			(0.077)
More politically engaged			0.120			0.005
			(0.076)			(0.076)
Constant	-0.101	-0.125	-0.185	-0.123	-0.076	-0.039
	(0.078)	(0.110)	(0.123)	(0.078)	(0.110)	(0.123)
Observations	739	739	739	739	739	739
R^2	0.006	0.013	0.017	0.005	0.011	0.013

$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, classical standard error in parenthesis.
Note: The table reports regressions of a standardized dummy taking the value of one for citizens that punish some investment and zero for citizens that never punish on the treatment variables. The variables are as defined in Table 3.

Table B.10

Effect of taxation on share invested in the common pool (leaders), with classical standard errors.

	Share invested			High investment dummy		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned & Non-tax	0.033	0.043	0.050	0.114	0.119	0.124
	(0.101)	(0.100)	(0.100)	(0.101)	(0.100)	(0.099)
Windfall & Tax	-0.094	-0.099	-0.104	-0.040	-0.053	-0.059
	(0.103)	(0.102)	(0.102)	(0.103)	(0.102)	(0.101)
Tax State	0.052	0.058	0.062	0.108	0.105	0.109
	(0.088)	(0.087)	(0.087)	(0.088)	(0.087)	(0.087)
Male		-0.028	-0.035		0.044	0.037
		(0.061)	(0.061)		(0.061)	(0.061)
Above median age		0.227***	0.241***		0.275***	0.287***
		(0.065)	(0.065)		(0.064)	(0.065)
Above median education		0.100	0.074		0.088	0.060
		(0.068)	(0.068)		(0.068)	(0.068)
Employed full time		-0.213***	-0.223***		-0.192***	-0.203***
		(0.070)	(0.070)		(0.070)	(0.070)
Above median income		-0.061	-0.053		-0.071	-0.066
		(0.072)	(0.072)		(0.072)	(0.072)
Conservative			-0.133**			-0.110*
			(0.066)			(0.066)
More politically engaged			0.131**			0.161**
			(0.066)			(0.066)
Constant	-0.009	-0.019	-0.012	-0.059	-0.123	-0.136
	(0.072)	(0.094)	(0.101)	(0.072)	(0.093)	(0.101)
Observations	957	957	957	957	957	957
R^2	0.003	0.032	0.042	0.004	0.035	0.046

$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, classical standard error in parenthesis.
Note: The table reports regressions of the standardized value of share invested by the leader on the treatment variables "Hard Earned & Non-tax", "Windfall & Tax" and "Tax State" and a set of explanatory variables (see Table 3 for definitions of variables). Columns (1)-(3) show the results for the discrete definition of share invested (takes values 0.1, 0.2,...,1). Columns (4)-(6) show the results for a dummy taking the value of one if the leader invests everything in the common pool.

Table B.11
Robustness check, positive punishment indicator.

	(1)	(2)	(3)
Hard Earned & Non-tax	0.070	0.058	0.072
	(0.112)	(0.112)	(0.113)
Windfall & Tax	0.086	0.098	0.098
	(0.117)	(0.117)	(0.116)
Tax State	0.080	0.070	0.073
	(0.099)	(0.098)	(0.098)
Male		-0.175**	-0.172**
		(0.069)	(0.069)
Above median age		-0.092	-0.102
		(0.075)	(0.075)
Above median education		-0.038	-0.043
		(0.079)	(0.079)
Employed full time		0.176**	0.182**
		(0.082)	(0.082)
Above median income		0.048	0.045
		(0.084)	(0.083)
Conservative			0.056
			(0.077)
More politically engaged			0.159**
			(0.075)
Constant	-0.062	-0.030	-0.130
	(0.079)	(0.110)	(0.122)
Observations	739	739	739
R^2	0.001	0.018	0.024

$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, robust standard error in parenthesis.
Note: The table reports regressions of a standardized dummy taking the value of one for citizens that punish some investment and zero for citizens that never punish on the treatment variables. The variables are as defined in Table 3.

B.2.1. Robustness, hard earned mechanism

The basic idea behind the Hard Earned mechanism is that the group endowment is generated by the citizen and the leader's performance of the real effort task. There are two possible ways to finance the group endowment through conduction of the real effort task while keeping the size of the group endowment and reward constant across treatments. First, the duration of the real effort task can be held constant and equal to five minutes, and the productivity of the real effort task can be increased from $1 to $2 going from the Windfall to the Hard Earned treatments. Second, the productivity of the real effort task can be held constant and equal to $1 per minute, and the duration of the task can b increased from five to ten minutes between the Windfall and Hard Earned treatments.

In the main treatments, the former type of manipulation of the Hard Earned mechanism is used. To investigate whether the effect of taxation on the willingness to punish is sensitive to the way the Hard Earn mechanism is manipulated, a robustness treatment, "Tax State Extra Hard Earned" is implemented. The treatment is identical to the Rentier State treatment in all respects apart from the real effort task being ten minutes long in stead of five. The treatment can be summarized as follows:

Tax State Extra Hard Earned (Extra Hard Earned & Tax (T4)): Citizen and leader each earn $2 from performing a **10** minute real effort task. Their earnings are taxed at 50 percent. The $1 tax collected from each finances the $2 group endowment.

Empirical strategy

To test whether the Hard Earned mechanism is sensitive to whether the productivity or the duration of the task is manipulated, the following regression is estimated on a sample restricted to the two Tax State treatments:

$$y_i = \alpha + \beta_{\text{T4b}}\text{T4b}_i + \beta^X X_i + \beta^Z Z_i + \epsilon_i, \tag{4}$$

where T4b_i and an indicator variable taking the value of one for citizens in the Tax State Extra Hard Earned treatment and zero for citizens in the Tax State treatment.

Results

Table B.12
Robustness check for Hard Earned treatments.

	(1)	(2)	(3)
Extra Hard Earned	0.038	0.038	0.068
	(0.123)	(0.123)	(0.122)
Male		-0.081	-0.106
		(0.132)	(0.131)
Above median age		0.076	0.050
		(0.125)	(0.124)
Above median education		0.007	0.023
		(0.130)	(0.130)
Employed full time		0.354**	0.353**
		(0.141)	(0.142)
Above median income		-0.105	-0.100
		(0.142)	(0.142)
Conservative			0.263**
			(0.127)
More politically engaged			0.264**
			(0.127)
Observations	284	284	284
R^2	0.000	0.024	0.047

$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, robust standard error in parenthesis.
Note: The table reports regressions for the standardized willingness to punish variable on "Extra Hard Earned" an indicator variable taking the value of one for individuals in the Hard Earned treatment that worked 10 minutes instead of 5 minutes, and a range of background variables. The background variables are as defined in Table 3.

Table B.13

Inclusion of positive reciprocity and risk aversion as co-variates.

	Willingness to punish			High willingness to punish		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned & Non-tax	0.034	0.031	0.034	0.101	0.103	0.093
	(0.107)	(0.107)	(0.108)	(0.102)	(0.103)	(0.102)
Windfall & Tax	0.121	0.133	0.130	0.168	0.168	0.156
	(0.116)	(0.116)	(0.116)	(0.112)	(0.113)	(0.114)
Tax State	0.188*	0.182*	0.188*	0.185**	0.176*	0.183**
	(0.098)	(0.098)	(0.098)	(0.092)	(0.092)	(0.091)
Male		-0.113	-0.121*		-0.056	-0.062
		(0.069)	(0.069)		(0.068)	(0.068)
Above median age		0.001	0.008		0.002	0.018
		(0.074)	(0.076)		(0.075)	(0.077)
Above median education		-0.004	-0.005		0.104	0.107
		(0.080)	(0.081)		(0.082)	(0.084)
Employed full time		0.137*	0.141*		-0.058	-0.062
		(0.081)	(0.081)		(0.080)	(0.079)
Above median income		0.006	-0.004		-0.090	-0.095
		(0.084)	(0.085)		(0.083)	(0.084)
Conservative			0.005			-0.091
			(0.077)			(0.076)
More politically engaged			0.114			-0.004
			(0.076)			(0.074)
High positive reciprocity			0.012			0.053
			(0.075)			(0.076)
Risk averse			-0.113			-0.135*
			(0.075)			(0.074)
Constant	-0.101	-0.125	-0.127	-0.123*	-0.076	0.010
	(0.076)	(0.109)	(0.135)	(0.067)	(0.108)	(0.130)
Observations	739	739	739	739	739	739
R^2	0.006	0.013	0.020	0.005	0.011	0.018

$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$, robust standard error in parenthesis.

Note: The table reports regressions for the standardized willingness to punish variable on the treatment variables "Hard Earned & Non-tax", "Windfall & Tax" and "Tax State" and the background defined in Table 3, as well as "High positive reciprocity": indicator variable taking the value of 1 for individuals that answered a number higher than the median reported (10) to the question *When someone does me a favor, I am willing to return it* (0 = Strongly disagree, ..., 10 = Strongly agree), and zero otherwise and "Risk averse": indicator variable taking the value of 1 for individuals answering a number lower than median reported (7) to the question *How willing or unwilling you are to take risks?* (0 = Completely unwilling, ..., 10 = Completely willing), and zero otherwise.

B.3. Heterogeneity analysis

B.3.1. Empirical strategy

To investigate whether different subgroups of the sample respond differently to the treatments, the following specification with interaction terms between treatment and indicator variables is estimated in Table B.14:

$$y_i = \alpha + \beta_{\text{Var}}\text{Var}_i + \beta_{\text{T2}}\text{T2}_i + \theta_{\text{T2}}\text{T2}_i \times \text{Var}_i + \beta_{\text{T3}}\text{T3}_i + \theta_{\text{T3}}\text{T3}_i \times \text{Var}_i$$

$$+\beta_{\text{T4}}\text{T4}_i + \theta_{\text{T4}}\text{T4}_i \times \text{Var}_i + \beta_{\text{X}}X_i + \beta_{\text{Z}}Z_i + \epsilon_i. \tag{5}$$

Where Var_i are indicator variables for the subgroups of respondents that are of interest (male, above median age, above median education, full-time employee, above median income, conservative, more than median engaged in political activities, and above median upset). $\text{T2}_i \times \text{Var}_i$ is an interaction term between T2_i and Var_i, $\text{T3}_i \times \text{Var}_i$ is an interaction term between T3_i and Var_i, and $\text{T4}_i \times \text{Var}_i$ is interaction term between T4_i and Var_i. The reference category is the subgroup for which Var_i takes the value of zero in the Rentier State treatment. When estimating Equation (5) for each background variable, θ_{T2}, θ_{T3} and θ_{T4} test whether the treatments deferentially affect different subgroups of the sample.

B.3.2. Results

Table B.14
Heterogeneity in effect of taxation on punishment.

	Var						
	Male	Older	Higher educ	Fulltime	Higher income	Conservative	Pol. engaged
Var	-0.190	0.180	-0.273[*]	-0.081	-0.101	-0.099	0.117
	(0.135)	(0.153)	(0.156)	(0.161)	(0.158)	(0.153)	(0.154)
Hard Earned & Non-tax	0.017	0.195	-0.045	-0.036	0.007	0.078	0.135
	(0.150)	(0.148)	(0.160)	(0.173)	(0.149)	(0.146)	(0.150)
Windfall & Tax	0.052	0.289[*]	-0.243	-0.025	-0.004	0.105	0.078
	(0.158)	(0.158)	(0.169)	(0.180)	(0.160)	(0.152)	(0.168)
Tax State	0.127	0.256[*]	0.024	-0.039	0.136	0.036	0.145
	(0.133)	(0.134)	(0.149)	(0.157)	(0.133)	(0.135)	(0.138)
Hard Earned & Non-tax X Var	0.038	-0.323	0.144	0.122	0.072	-0.105	-0.227
	(0.203)	(0.216)	(0.216)	(0.221)	(0.217)	(0.216)	(0.217)
Windfall & Tax X Var	0.140	-0.333	0.668[***⋆]	0.260	0.287	0.035	0.099
	(0.187)	(0.233)	(0.231)	(0.237)	(0.233)	(0.237)	(0.234)
Tax State X Var	0.117	-0.137	0.287	0.382[*]	0.107	0.316	0.088
	(0.183)	(0.197)	(0.197)	(0.200)	(0.197)	(0.195)	(0.197)
Background vars	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Political vars	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	739	739	739	739	739	739	739
R^2	0.018	0.021	0.028	0.022	0.019	0.024	0.021

Note: The table reports regressions of the standardized value of willingness to punish on the treatment variables "Hard Earned & Non-tax", "Windfall & Tax" and "Tax State" and the interaction between these and indicator variables for different subgroups of the sample ("Var"). "Male": indicator variable taking the value of 1 for males, "Above median age':' indicator variable taking the value of 1 if individual is older than the median age of the sample (34 years), "Above median education": indicator taking the value of 1 for individuals who have a 4-year degree or higher education, "Employed full-time": indicator variable for individuals who are full-time employees, "Above median income": indicator variable taking the value of 1 for individuals that have an individual yearly income of USD 40 000 or more, "Conservative": indicator variable taking the value of 1 for individuals that rate themselves 6 or higher on a scale from 0 (strongly liberal) to 10 (strongly conservative) and "More politically engaged": indicator variable taking the value of 1 for individuals that report to have participated in more than the median number (2) of political activities during the last year. Robust standard errors are reported in parenthesis. Statistical significance is indicated by [*] $p < 0.10$, [**] $p < 0.05$, [***] $p < 0.01$ for unadjusted p-values and by $\star p < 0.1$, $\star\star p < 0.05$, $\star\star\star p < 0.01$ for p-values adjusted for multiple hypotheses testing. When correcting for multiple hypotheses testing, p-values for all the heterogeneity dimensions in this table as well as Table B.15 (Male, Older, Higher educ., Fulltime, Higher income, Conservative, Pol. engaged, Risk averse, More altruistic, High pos. reci., High neg. reci., Loss averse and More upset) are grouped together in one family for each of the three treatments and the three interaction terms.

Table B.15

Heterogeneity in effect of taxation on punishment, by preferences.

	Var					
	Risk averse	More altruistic	High pos. reci.	High neg. reci.	Loss averse	More upset
Var	-0.135	0.198	0.237	0.847***	-0.094	0.516***
	(0.153)	(0.163)	(0.155)	(0.143)	(0.159)	(0.152)
Hard Earned & Non-tax	-0.037	0.201	0.255*	-0.081	0.064	0.043
	(0.160)	(0.132)	(0.150)	(0.111)	(0.169)	(0.129)
Windfall & Tax	0.207	0.219	0.309*	-0.003	0.032	0.025
	(0.176)	(0.141)	(0.163)	(0.122)	(0.183)	(0.145)
Tax State	0.167	0.280**	0.267**	0.065	0.162	0.126
	(0.147)	(0.118)	(0.133)	(0.112)	(0.155)	(0.125)
Hard Earned & Non-tax X Var	0.149	-0.457**	-0.442**	0.234	-0.069	0.012
	(0.216)	(0.226)	(0.217)	(0.194)	(0.220)	(0.213)
Windfall & Tax X Var	-0.143	-0.242	-0.349	0.239	0.163	0.115
	(0.233)	(0.248)	(0.232)	(0.209)	(0.238)	(0.227)
Tax State X Var	0.038	-0.263	-0.165	0.028	0.028	0.048
	(0.197)	(0.210)	(0.197)	(0.181)	(0.202)	(0.194)
Background vars	Yes	Yes	Yes	Yes	Yes	Yes
Political vars	Yes	Yes	Yes	Yes	Yes	Yes
Observations	739	739	739	739	739	739
R^2	0.022	0.022	0.023	0.232	0.019	0.092

Note: The table reports regressions of the standardized value of willingness to punish on the treatment variables "Hard Earned & Non-tax", "Windfall & Tax" and "Tax State". and interaction between these and "Var", indicator variables for different subgroups of the sample. "Risk averse": indicator variable taking the value of 1 for individuals answering a number lower than median reported (7) to the question *How willing or unwilling you are to take risks?* (0 = Completely unwilling, ..., 10 = Completely willing), and zero otherwise. "More altruistic": indicator variable taking the value of one for individuals answering a number higher than the median reported (9) to the question *How willing are you to give to good causes without expecting anything in return?* (0 = Completely unwilling, ..., 10 = Completely willing), and zero otherwise. "High negative Reciprocity": indicator variable taking the value of one if negative reciprocity index is higher than the median. In accordance with Falk et al. (2018), the index for negative reciprocity was constructed using the following three items *If I am treated very unjustly, I will take revenge at the first occasion, even if it is a cost to do so* (0 = Strongly disagree, ..., 10 = Strongly agree (10)), *How willing are you to punish someone who treats you unfairly, even if there may be costs for you?* (0 = Completely unwilling, ..., 10 = Completely willing) and *How willing are you to punish someone who treats others unfairly, even if there may be costs for you?* (0 = Completely unwilling, ..., 10 = Completely willing). From these questions, the index was constructed as follows: negative reciprocity = 0.2631× Willingness to punish if oneself is treated unfairly + 0.2631× Willingness punish if other is treated unfairly + 0.3738× Willingness to take revenge. These weights were obtained by Falk et al. (2018) by running a regression of observed risk behavior in the lab on responses to these survey questions. "High positive reciprocity": indicator variable taking the value of 1 for individuals that answered a number higher than the median reported (10) to the question *When someone does me a favor, I am willing to return it* (0 = Strongly disagree, ..., 10 = Strongly agree), and zero otherwise. "Loss averse": indicator variable taking the value of one for individuals that to the following question *If you could choose between the following hypothetical scenarios, which would you choose? 1. Lottery: win $80 with probability 1/2, lose $50 with probability 1/2 or 2. Receive $0 for sure* answer "2. Receive $0 for sure". "Above median upset": indicator variable taking the value of 1 for individuals answering a number higher than the median reported (6) to the question *How upset would you feel if the leader invests less than what you think is fair in the common pool?* (0 = Not upset at all, ..., 10 = Very upset). Robust standard errors are reported in parenthesis. Statistical significance is indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ for unadjusted p-values and by ⋆$p < 0.1$, ⋆⋆$p < 0.05$, ⋆⋆⋆$p < 0.01$ for p-values adjusted for multiple hypotheses testing. When correcting for multiple hypotheses testing, p-values for all the heterogeneity dimensions in this table as well as Table B.14 (Male, Older, Higher educ., Fulltime, Higher income, Conservative, Pol. engaged, Risk averse, More altruistic, High pos. reci., High neg. reci., Loss averse and More upset) are grouped together in one family for each of the three treatments and the three interaction terms.

B.4. Testing treatment effects in the loss and gain domain

The present data can shed light on the importance of loss-aversion in explaining willingness to punish, by comparing behavior in Non-Tax to Tax-treatments. Martin (2014) derives a model for taxation, loss aversion and accountability built on Sandbu (2006) and Kahneman and Tversky (1979). She argues that citizens evaluate their utility in reference to pre-tax income. According to Prospect Theory (Kahneman and Tversky, 1979) the utility function of citizens is steeper below the reference point than above the reference point. Thus, low investments in the common pool leads to a larger utility loss for citizens in the loss domain than an equally low investments in the common pool for citizens in the gain domain.

Loss aversion can be included in the theoretical framework presented in Equation (1) by adding a term for utility from punishment of deviations from the reference payoff:

$$V_c = R + \frac{3}{4}I - p \times c - \gamma_c (m_c - I)^2 + p \big(\beta_{c,m} [min\{0, m_c - I\}^2] + \beta_{c,r} [min\{0, r_c - I\}^2] \big) \tag{6}$$

where r is the leader investment required for the citizen to get his reference payoff (pre-tax income) and $\beta_{c,r}$ is the weight the citizen attaches to utility from punishment of investments that lead to payoffs lower than the reference payoff.

Equation (6) shows that the citizen's utility from punishment depends positively on how much the leader's investment deviates from the fairness norm, $m_c - I$, how much the leader's investment deviates from the investment required for the citizen to get his reference payoff, $r_c - I$, and on the individual parameters $\beta_{c,m}$ and $\beta_{c,r}$, and negatively on the cost of punishment, c. The citizen's punishment behavior is then characterized by:

$$c < \big(\beta_{c,m}[min\{0, m_c - I\}^2] + \beta_{c,r}[min\{0, r_c - I\}^2] \big) \qquad \text{Punish} \tag{7a}$$

$$c = \big(\beta_{c,m}[min\{0, m_c - I\}^2] + \beta_{c,r}[min\{0, r_c - I\}^2] \big) \qquad \text{Indifferent, randomize} \tag{7b}$$

$$c > \big(\beta_{c,m}[min\{0, m_c - I\}^2] + \beta_{c,r}[min\{0, r_c - I\}^2] \big) \qquad \text{Not punish} \tag{7c}$$

Thus, when adding loss aversion to the theoretical framework, the citizen's punishment is influenced by his or her reference payoff, r, and the utility he or she derives from punishment of investments that gives a payoff lower than their reference, β_{r_c}, in addition to m_c, $\beta_{c,m}$ and c.

If the assumption that the pre-tax income is the citizens' reference point is correct, the theory implies citizens in the tax treatments in my experiment are in the realm of losses as long as the leader invests less 0.7 of the group endowment in the common pool (because it leaves them with a lower payoff than their \$2 pre-tax income). Citizens in the non-tax treatments, on the other hand, are always in the gain domain, because they do not pay tax and always get their "pre-tax" income of \$1. Thus, for levels of investment between 0 and 0.6, citizens in the tax treatments experience larger utility losses than citizens in the non-tax treatments, and if loss aversion is a mechanism for punishment, we would expect to see more punishment of investments between 0 and 0.6 in the tax treatments than in the non-tax treatments, but similar punishment behavior for investments of 0.7 and higher.

To investigate if this is the case in the present data, Fig. B.1 illustrates the effect of going from a Non-Tax treatment to a Tax treatment on punishment of investments shares for which citizens in the Tax treatments are below their reference point (upper panel) and on punishment of investment shares for which all citizens are above their reference point (lower panel), respectively. Both panels show the effect for Windfall treatments only (T1 and T3), for Hard Earned treatments (T2 and T4) only, and for the pooled sample of Windfall and Hard Earned treatments. The upper panel illustrates that going from a Non-Tax to Tax treatment increases willingness to punish investment shares that leave citizens in the Tax treatments below their reference point for Windfall treatments, but that the effect is not significant for any of the three comparisons (see Table B.16 for corresponding regressions). The lower panel shows that going from Non-Tax to Tax treatments increases punishment of investment shares that leave citizens in all treatments above their reference point and that the effect is significant for both the Hard Earned treatment and the pooled sample of Windfall and Hard Earned treatments (see Table B.17 for corresponding regressions). The finding that the Tax manipulation increases punishment of investments that leave citizens in both Tax and Non-Tax treatments above their reference point, but does not significantly affect punishment of investment levels that leave the citizens in the Tax treatments below their reference point does not support the reference dependent model in which pre-tax income is the reference payoff for citizens and they only derive utility from punishment of investments that leave them in the realm of losses. There are several possible interpretations of this result. For instance, the citizens' reference point may not be defined by their pre-tax income, punishment may not solely be determined by the reference point, but can also be influenced by other factors such as fairness and the utility from punishment may be not be zero, for positive deviations from the reference point.
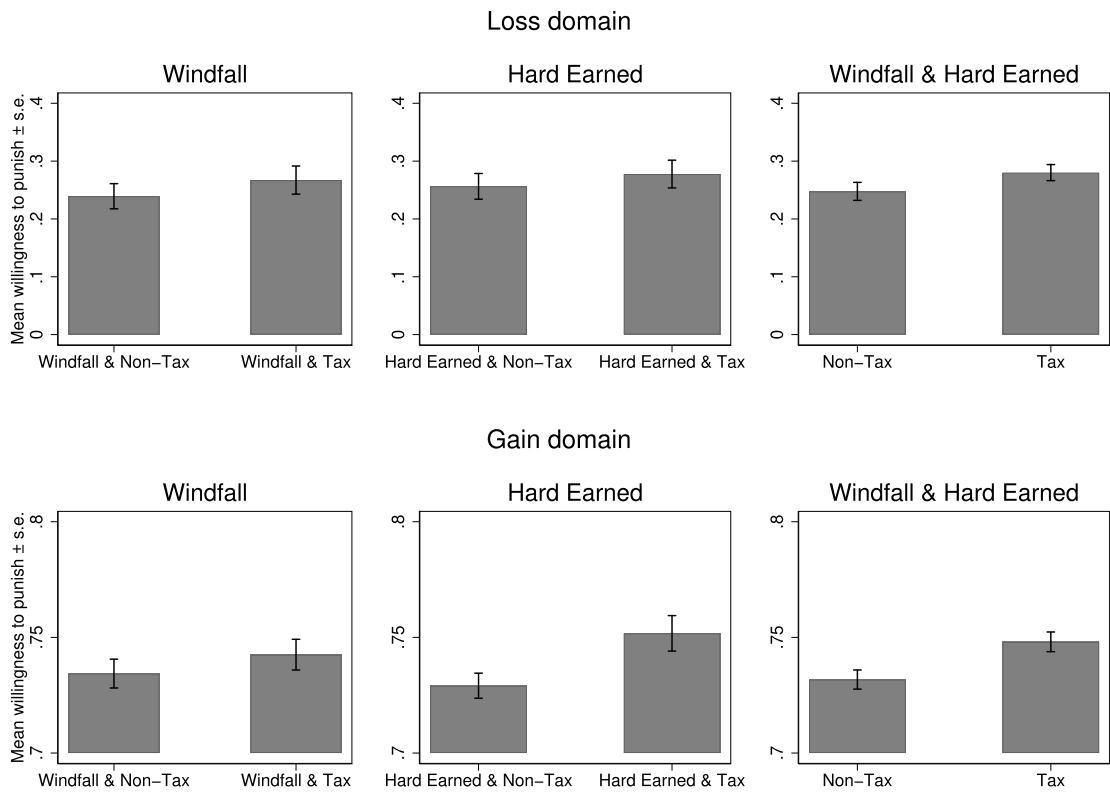
Loss domain

Gain domain

Fig. B.1. Effect of Tax treatments below and above reference point.

Table B.16
Effect of Tax treatments for investment levels below reference point.

	Windfall		Hard Earned		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Tax	0.099	0.094	0.075	0.071	0.115	0.114
	(0.116)	(0.116)	(0.116)	(0.115)	(0.074)	(0.074)
Male		−0.112		−0.298**		−0.124*
		(0.100)		(0.119)		(0.069)
Above median age		−0.004		−0.156		−0.020
		(0.117)		(0.117)		(0.074)
Above median education		−0.014		−0.032		−0.031
		(0.126)		(0.127)		(0.079)
Employed full-time		0.047		0.319**		0.187**
		(0.125)		(0.138)		(0.082)
Above median income		0.064		−0.021		−0.007
		(0.130)		(0.139)		(0.084)
Conservative		−0.053		0.026		0.023
		(0.122)		(0.119)		(0.077)
More politically engaged		0.199*		0.022		0.138*
		(0.120)		(0.117)		(0.075)
Constant	−0.095	−0.159	−0.034	0.002	−0.065	−0.156
	(0.077)	(0.173)	(0.079)	(0.155)	(0.055)	(0.106)
Observations	297	297	301	301	739	739
R^2	0.002	0.021	0.001	0.041	0.003	0.018

Robust standard errors in parentheses, $^* p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$

Note: The table reports regressions where the dependent variable is the standardized value of the willingness to punish for investment levels that leaves the citizen below his or her reference point in the Tax treatment. Columns (1)-(2) present results for a sample restricted to the Windfall treatments, Columns (3)-(4) present results for a sample restricted to the Hard Earned treatments and Columns (5)-(6) present results for the whole sample. "Tax" is an indicator variable taking the value of one for individuals in the tax treatments, Windfall & Tax and Hard Earned & Tax. See Table 3 for definitions of background variables.

Table B.17
Effect of Tax treatments for investment levels above reference point.

	Windfall		Hard Earned		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Tax	0.100	0.090	0.276**	0.276**	0.199***	0.197***
	(0.111)	(0.114)	(0.114)	(0.113)	(0.072)	(0.072)
Male		−0.093		−0.195*		−0.063
		(0.088)		(0.116)		(0.067)
Above median age		0.032		−0.134		0.026
		(0.110)		(0.114)		(0.074)
Above median education		0.103		−0.019		0.042
		(0.127)		(0.126)		(0.082)
Employed full-time		−0.035		0.050		0.004
		(0.117)		(0.121)		(0.079)
Above median income		0.017		0.034		0.031
		(0.136)		(0.140)		(0.086)
Conservative		−0.147		−0.000		−0.020
		(0.118)		(0.116)		(0.078)
More politically engaged		−0.049		0.039		0.032
		(0.115)		(0.114)		(0.076)
Constant	−0.081	−0.002	−0.145**	−0.043	−0.112**	−0.140
	(0.076)	(0.168)	(0.066)	(0.144)	(0.050)	(0.106)
Observations	297	297	301	301	739	739
R^2	0.003	0.015	0.020	0.033	0.010	0.012

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Note: The table reports regressions where the dependent variable is the standardized willingness to punish variable for investment levels that leaves the citizen above the reference point. Columns (1)-(2) present results for a sample restricted to the Windfall treatments, Columns (3)-(4) present results for a sample restricted to the Hard Earned treatments and Columns (5)-(6) present results for the whole sample. "Tax" is an indicator variable taking the value of one for individuals in the tax treatments, Windfall & Tax and Hard Earned & Tax. See Table 3 for definitions of background variables.

B.5. Leaders

Table B.18
Effect of treatment on background variables, leaders.

	Male	Older	Higher educ	Full time	Higher income	Conservative	Pol. engaged
Hard Earned & Non-Tax	0.065	0.013	0.052	0.077	0.008	0.049	0.018
	(0.053)	(0.051)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)
Windfall & Tax	0.091*	0.024	−0.055	−0.033	−0.016	−0.007	0.023
	(0.055)	(0.052)	(0.052)	(0.052)	(0.051)	(0.051)	(0.052)
Hard Earned & Tax	0.113**	0.017	−0.003	0.033	−0.005	0.008	−0.018
	(0.047)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)	(0.044)
Constant	0.428***	0.485***	0.546***	0.536***	0.464***	0.433***	0.469***
	(0.038)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
P-value of F-test	0.586	0.978	0.112	0.092	0.893	0.504	0.573

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Note: The table displays coefficients from estimated regressions of each of the background indicator variables as the dependent variable on indicator variables for the treatments for the leaders in the sample. The Rentier State treatment is the reference category. See Table 3 for definition of the variables. The reported p-values test the hypothesis that all the treatments have the same effect on the background indicator variables.

B.6. Comparison to previous studies

Table B.19
Comparison of treatment effects.

Paper	Source, means	Source, coeff.	Range	Control mean	Treatment mean	Diff., means	Reg. coeff.	Diff. comp
de la Cuesta et al. (2022)	Table G.1, p. 39	Table 4, p. 312	[0,1,2,..,10]	4.49[*]	4.71	0.22	0.442	0.022
Martin (2014)	Text, p. 24	Table 2, p. 25	[0, 100, 200, ..., 1 000]	408	460	54	56.37	0.054
This paper								
R. State vs. WF & Tax	Table B.6, p. 45	Table B.7, p. 47	[0, 0.1, 0.2, ..., 1]	0.296	0.342	0.043	0.048	0.043
R. State vs. T. State	Table B.6, p. 45	Table B.7, p. 47	[0, 0.1, 0.2, ..., 1]	0.296	0.366	0.070	0.070	0.070

[*] Calculated average of Windfall Aid, Windfall Grant, and Windfall Oil.

Note: The table reports control mean, treatment mean, difference in means, and regression coefficient of treatment effect (with controls) for de la Cuesta et al. (2022), Martin (2014) and this paper. For this paper, the table reports these statistics for both the difference between Rentier State and Windfall & Tax, which is the closest comparison to de la Cuesta et al. (2022) and Martin (2014), and for the difference between Rentier State and Tax State. The "Diff. comp" column reports the difference in means converted to share of the group endowment transferred, to make the results comparable to those in the present paper.

Appendix C. Theoretical framework for leader decisions

The leader's investment decision can be modeled as a trade-off between expected monetary payoff and non-monetary disutility from deviations from fairness:

$$V_l = R + GE - \frac{1}{4}I - (\phi(I) \times \theta) - \gamma_l (m_l - I)^2, \tag{8}$$

where GE is the group endowment, $\phi(I)$ ($\phi'(I) < 0$) is the leader's subjective probability for being punished by the citizen as a function of investment, θ is the cost of being punished by the citizen, m_l is the leader's fairness norm for investment, γ_l is a parameter determining the leader's disutility from deviating from fairness, and the other parameters are as defined for Equation (6). $m_l - I$ indicates how much the leader's investment deviates from what he or she perceives as fair. It is assumed that the leader experiences a utility loss from both negative and positive deviations from the fairness norm, and that the disutility is increasing in the size of the deviation. For simplicity, it is also assumed that the leader is risk neutral. Given an interior solution, the leader's optimal investment in the common pool is given by:

$$I^* = m_l - \frac{1}{2\gamma_l}\left[\frac{1}{4} + \phi'(I) \times \theta\right]. \tag{9}$$

Thus, the leader's investment depends positively on his or her fairness norm for investment, m_l, the parameter determining disutility from unfair investment, γ_l, and the subjective belief about how sensitive the probability of punishment by the citizen is to a change in investment, $\phi'(I)$.

Next, the model is used to illustrate the effect of taxation on the leader's investment decision. Equation (9) shows that the investment decision is determined by the level of investment the leader perceives as fair (m_l), his or her disutility from unfair investments (γ_l) and the subjective belief about how sensitive the probability of punishment by the citizen is to a change in investment ($\phi'(I)$) for any given level of investment (I). As before, the fairness norm, m_l, is assumed to be constant across treatments, so the effect of treatments must go through disutility from unfair investments (γ_l) and the subjective belief about how sensitive the probability of punishment by the citizen is to a change in investment ($\phi'(I)$). Corresponding to the effect of taxation on the citizen's willingness to punish, the idea is that the leader's active contribution to the group endowment under taxation makes fairness considerations more salient and increases the disutility the leader derives from unfair investments, γ_l. This, in turn, is hypothesized to increase investments in the common pool. Taxation might also affect the leader's subjective belief about how sensitive the probability of the citizen punishing is to a change in investment, but it is difficult to formulate a theoretical prediction for this effect without further assumptions. Based on this, the theory predicts the leader's investments to be higher in the Hard Earned & Non-Tax, Windfall & Tax and Tax State treatments than in the Rentier State treatment.

Appendix D. Instructions

D.1. Introduction

NHH

The results from this experiment will be used in a research project at the Norwegian School of Economics. Participation in the study is completely voluntary.

You are free to decline to participate, or to end participation at any time and for any reason. You will receive a participation fee of $1 upon completion of the experiment. Depending on the actions you and others take, you may also earn additional money.

Your will remain anonymous throughout the experiment. None of the information collected can be traced back to individual participants. We will only use your Worker ID to assign payments and to check that you have not participated in this experiment before.

Please read the instructions carefully. The duration of the experiment is approximately 15-20 minutes.

If you have any questions regarding this experiment, please contact thechoicelab@nhh.no.

I have read and understood the above information, and agree to participate in this study.

Yes

No

`>>`

Fig. D.1. Introduction (all participants).

D.2. Role assignment

NHH

You have been randomly assigned to a group consisting of you and another participant who is also an Amazon Mechanical Turk worker. You have been randomly assigned to the role as citizen and the person you have been grouped with has been randomly assigned to the role as leader in the experiment. You will not get to know anything about the person you are paired with, and this person will not get to know anything about you.

>>

Fig. D.2. Role assignment (all citizens).

NHH

You have been randomly assigned to a group consisting of you and another participant who is also an Amazon Mechanical Turk worker. You have been randomly assigned to the role as leader and the person you have been grouped with has been randomly assigned to the role as citizen in the experiment. You will not get to know anything about the person you are paired with, and this person will not get to know anything about you.

>>

Fig. D.3. Role assignment (all leaders).

D.3. Task description (citizens)

NHH

You now get the opportunity to earn $1 by completing a 5-minute picture categorization assignment.

The leader you have been grouped with got the opportunity to complete the same assignment as you. **He or she chose to do so, and earned $1.**

Regardless of your and the leader's choice to complete the assignment or not, your group receives $2 in group money.

It is the task of the leader to decide how much of the group money to invest in a common pool. The amount invested in the common pool will be multiplied by 1.5 and shared equally between you and the leader. The remaining group money will be given to the leader.

Your total payment will consist of two parts:

1. Earnings ($1 if you choose to complete the assignment, $0 otherwise)
2. Half of the common pool ($0 - $1.5, depending on the leader's decision)

The leader's total payment will consist of three parts:

1. Earnings ($1 because he or she chose to complete the assignment)
2. Half of the common pool ($0 - $1.5, depending on the leader's decision)
3. The group money not invested in the common pool ($0 - $2, depending on the leader's decision)

Fig. D.4. Task description, Rentier State.

NHH

You now get the opportunity to earn $1 by completing a 5-minute picture categorization assignment. If you complete the assignment, your group receives $1 in group money.

The leader you have been grouped with got the opportunity to complete the same assignment as you. **He or she chose to do so, and earned $1. Because of the leader's completion of the assignment, your group receives $1 in group money.**

It is the task of the leader to decide how much of the group money to invest in a common pool. The amount invested in the common pool will be multiplied by 1.5 and shared equally between you and the leader. The remaining group money will be given to the leader.

Your total payment will consist of two parts:

1. Earnings ($1 if you choose to complete the assignment, $0 otherwise)
2. Half of the common pool ($0 - $1.5, depending on the leader's decision and the group money available)

The leader's total payment will consist of three parts:

1. Earnings ($1 because he or she chose to complete the assignment)
2. Half of the common pool ($0 - $1.5, depending on the leader's decision and the group money available)
3. The group money not invested in the common pool ($0 - $2, depending on the leader's decision and the group money available)

Fig. D.5. Task description, Hard Earned & Non-Tax.

NHH

You now get the opportunity to earn $1 by completing a 5-minute picture categorization assignment. Regardless of whether you chose to complete the assignment, you additionally receive $1.

The leader you have been grouped with got the opportunity to complete the same assignment as you. **He or she chose to do so, and earned $1. In addition, the leader received $1, independent of his or her choice to complete the assignment.**

Your and the leader's total earnings (earnings from assignment + additional dollar received) will be taxed at 50%. It is the task of the leader to decide how much of the tax revenues collected to invest in a common pool. The amount invested in the common pool will be multiplied by 1.5 and shared equally between you and the leader. The remaining tax revenues will be given to the leader.

Your total payment will consist of two parts:

1. Post-tax earnings ($1 if you choose to complete the assignment, $0 otherwise)
2. Half of the common pool ($0 - $1.5, depending on the leader's decision and the tax revenues available)

The leader's total payment will consist of three parts:

1. Post-tax earnings ($1 because he or she chose to complete the assignment
2. Half of the common pool ($0 - $1.5, depending on the leader's decision and the tax revenues available)
3. The tax revenues not invested in the common pool ($0 - $2, depending on the leader's decision and the tax revenues available)

Fig. D.6. Task description, Windfall & Tax.

NHH

You now get the opportunity to earn $2 by completing a 5-minute picture categorization assignment.

The leader you have been grouped with got the opportunity to complete the same assignment as you. **He or she chose to do so, and earned $2.**

Your and the leader's earnings will be taxed at 50%. It is the task of the leader to decide how much of the tax revenues collected to invest in a common pool. The amount invested in the common pool will be multiplied by 1.5 and shared equally between you and the leader. The remaining tax revenues will be given to the leader.

Your total payment will consist of two parts:

1. Post-tax earnings ($1 if you choose to complete the assignment, $0 otherwise)
2. Half of the common pool ($0 - $1.5, depending on the leader's decision and the tax revenues available)

The leader's total payment will consist of three parts:

1. Post-tax earnings ($1 because he or she chose to complete the assignment)
2. Half of the common pool ($0 - $1.5, depending on the leader's decision and the tax revenues available)
3. The tax revenues not invested in the common pool ($0 - $2, depending on the leader's decision and the tax revenues available)

Fig. D.7. Task description, Tax State.

D.4. Control questions

T1: RENTIER STATE TREATMENT

If I complete the assignment, I earn

    ⃝ $0

    ⃝ $1

    ⃝ $2

The leader I have been grouped with

    ⃝ Chose to complete the assignment and earned $1

    ⃝ Chose not to complete the assignment and earned $0

My group receives the $2 in group money regardless of whether I complete the assignment

    ⃝ True

    ⃝ False

What is the task of the leader?

    ⃝ No task

    ⃝ Decide how much of the group money to invest in the common pool

Fig. D.8. Control questions 1, Windfall & Non-Tax.

T2: HARD EARNED & NON-TAX

If I complete the assignment, I earn

○ $0

○ $1

○ $2

The leader I have been grouped with

○ Chose to complete the assignment and earned $1. This additionally generated $1 for the group.

○ Chose not to complete the assignment and earned $0.

The total group money generated amounts to

○ $0

○ $1 if I complete the assignment ($0 if I do not complete the assignment)

○ $2 if I complete the assignment ($1 if I do not complete the assignment)

○ $4

What is the task of the leader?

○ No task

○ Decide how much of the group money to invest in the common pool

---

Fig. D.9. Control questions 1, Hard Earned & Non-Tax.

T3: WINDFALL & NON-TAX

If I complete the assignment, I earn

○ $0, pre-tax

○ $1, pre-tax

○ $2, pre-tax

I additionally receive $1, regardless of whether I complete the assignment or not

○ True

○ False

The 50% tax applies to

○ None of my earnings

○ The $1 I earn if I complete the assignment

○ The additional $1 I receive

○ My total earnings ($1 from assignment + additional $1)

The leader I have been grouped with

○ Chose to complete the assignment and earned $1. He or she additionally received $1.

○ Chose not to complete the assignment and earned $0. He or she additionally received $1.

The tax revenues for the group amount to

○ $0

○ $1 if I complete the assignment ($0.5 if I do not complete the assignment)

○ $2 if I complete the assignment ($1.5 if I do not complete the assignment)

○ $4

What is the task of the leader?

○ No task

○ Decide how much of the tax revenues to invest in the common pool

Fig. D.10. Control questions 1, Windfall & Tax.

T4: HARD EARNED & TAX

If I complete the assignment, I earn

   ◯ $0, pre-tax

   ◯ $1, pre-tax

   ◯ $2, pre-tax

My earnings are taxed at 50%

   ◯ True

   ◯ False

The leader I have been grouped with

   ◯ Chose to complete the assignment earned $2. His or her earnings are taxed at 50%.

   ◯ Chose not to complete the assignment earned $0. His or her earnings are taxed at 50%.

The tax revenues for the group amount to

   ◯ $0

   ◯ $1 if I complete the assignment ($0.5 if I do not complete the assignment)

   ◯ $2 if I complete the assignment ($1 if I do not complete the assignment)

   ◯ $4

What is the task of the leader?

   ◯ No task

   ◯ Decide how much of the tax revenues to invest in the common pool

----

Fig. D.11. Control questions 1, Hard Earned & Tax.

D.5. Punishment decision

NHH

For completing the picture categorization assignment, you earn $1.

To confirm that you have understood how much you earn from completing the assignment, please enter the amount (without the dollar sign) in the box below.

Fig. D.12. Punishment decision, Rentier State.

NHH

We now ask you to make your decisions for each of the possible decisions the leader can make. You can do one of the following:

1. Reduce leader's payment by $0.5 by paying $0.05.
2. Do nothing.

We will implement your choice for the leader's actual decision. The table below illustrates the payoff you and the leader will receive from the common pool for all possible decisions the leader can make. In the first row, you indicate what you want to do if the leader invests nothing in the common pool. In the intermediate rows, you indicate what you want to do when the leader invests intermediate amounts in the common pool. In the last row, you indicate what you want to do if the leader invests everything in the common pool.

Please indicate your decision for **all** the possible leader decisions.

	Do you want to reduce the leader's payment by $0.5 by paying $0.05?	
	Yes	No
$0.0 (leader: $2.00, YOU: $0.00)	○	○
$0.2 (leader: $1.95, YOU: $0.15)	○	○
$0.4 (leader: $1.90, YOU: $0.30)	○	○
$0.6 (leader: $1.85, YOU: $0.45)	○	○
$0.8 (leader: $1.80, YOU: $0.60)	○	○
$1.0 (leader: $1.75, YOU: $0.75)	○	○
$1.2 (leader: $1.70, YOU: $0.90)	○	○
$1.4 (leader: $1.65, YOU: $1.05)	○	○
$1.6 (leader: $1.60, YOU: $1.20)	○	○
$1.8 (leader: $1.55, YOU: $1.35)	○	○
$2.0 (leader: $1.50, YOU: $1.50)	○	○

Why did you make the decisions the way you did? Please explain briefly in the box below.

Fig. D.12. (*continued*)

References

Ahmadov, A.K., 2014. Oil, democracy, and context: a meta-analysis. Comp. Polit. Stud. 47 (9), 1238–1267. https://doi.org/10.1177/0010414013495358.

Bagchi, S., Duöek, L., 2021. The effects of introducing withholding and third-party reporting on tax collections: evidence from the U.S. state personal income tax. J. Public Econ. 204. https://doi.org/10.1016/j.jpubeco.2021.104537. https://www.sciencedirect.com/science/article/pii/S0047272721001730.

Barr, A., Packard, T., Serra, D., 2014. Participatory accountability and collective action: experimental evidence from Albania. Eur. Econ. Rev. 68, 250–269. https://doi.org/10.1016/j.euroecorev.2014.01.010. https://www.sciencedirect.com/science/article/pii/S0014292114000129.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc., Ser. B, Methodol. 57 (1), 289–300.

Besley, T., 2020. State capacity, reciprocity, and the social contract. Econometrica 88 (4), 1307–1335.

Besley, T., Persson, T., 2014. Why do developing countries tax so little? J. Econ. Perspect. 28 (4), 99–120. https://doi.org/10.1257/jep.28.4.99. http://www.aeaweb. org/articles?id=10.1257/jep.28.4.99.

Bordalo, P., Gennaioli, N., Shleifer, A., 2012. Salience theory of choice under risk. Q. J. Econ. 127 (3), 1243–1285.

Bordalo, P., Gennaioli, N., Shleifer, A., 2013a. Salience and asset prices. Am. Econ. Rev. 103 (3), 623–628. http://www.jstor.org/stable/23469804.

Bordalo, P., Gennaioli, N., Shleifer, A., 2013b. Salience and consumer choice. J. Polit. Econ. 121 (5), 803–843.

Bordalo, P., Gennaioli, N., Shleifer, A., 2016. Competition for attention. Rev. Econ. Stud. 83 (2), 481–513.

Borge, L.E., Parmer, P., Torvik, R., 2015. Local natural resource curse? J. Public Econ. 131, 101–114. https://doi.org/10.1016/j.jpubeco.2015.09.002. http://www. sciencedirect.com/science/article/pii/S0047272715001504.

Bräutigam, D., Fjeldstad, O.H., Moore, M., 2008. Taxation and State-Building in Developing Countries: Capacity and Consent. Cambridge University Press.

Cohn, A., Maréchal, M.A., 2018. Laboratory measure of cheating predicts school misconduct. Econ. J. 128 (615), 2743–2754. https://doi.org/10.1111/ecoj.12572. https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12572.

de la Cuesta, B., Martin, L., Milner, H.V., Nielson, D.L., 2022. Owning it: accountability and citizens' ownership over aid, oil, and taxes. J. Polit. https://www. brandondelacuesta.com/research/#book.

Dai, Z., Galetto, F., Villeval, M.C., 2018. Cheating in the lab predicts fraud in the field: an experiment in public transportation. Manag. Sci. 64 (3), 1081–1100.

de la Cuesta, B., Martin, L., Milner, H.V., Nielson, D.L., 2021. Do indirect taxes bite? How hiding taxes erases accountability demands from citizens. Mimeo.

Dhami, S., al Nowaihi, A., 2007. Why do people pay taxes? Prospect theory versus expected utility theory. J. Econ. Behav. Organ. 64 (1), 171–192. https:// doi.org/10.1016/j.jebo.2006.08.006. http://www.sciencedirect.com/science/article/pii/S0167268107000765.

Engelmann, D., Nikiforakis, N., 2015. In the long-run we are all dead: on the benefits of peer punishment in rich environments. Soc. Choice Welf. 45 (3), 561–577.

Engström, P., Nordblom, K., Ohlsson, H., Persson, A., 2015. Tax compliance and loss aversion. Am. Econ. J.: Econ. Policy 7 (4), 132–164. https://doi.org/10.1257/ pol.20130134. http://www.aeaweb.org/articles?id=10.1257/pol.20130134.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences. Q. J. Econ. 133 (4), 1645–1692.

Fehr, E., Fischbacher, U., 2003. The nature of human altruism. Nature 425 (6960), 785–791.

Fehr, E., Fischbacher, U., 2004a. Social norms and human cooperation. Trends Cogn. Sci. 8 (4), 185–190.

Fehr, E., Fischbacher, U., 2004b. Third-party punishment and social norms. Evol. Hum. Behav. 25 (2), 63–87.

Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. Nature 415 (6868), 137–140.

Fehr, E., Krajbich, I., 2014. Social preferences and the brain. In: Neuroeconomics. Elsevier, pp. 193–218.

Fehr, E., Fischbacher, U., Gächter, S., 2002. Strong reciprocity, human cooperation, and the enforcement of social norms. Hum. Nat. 13 (1), 1–25.

Gadenne, L., 2017. Tax me, but spend wisely? Sources of public finance and government accountability. Am. Econ. J. Appl. Econ. 9 (1), 274–314. https://doi.org/10. 1257/app.20150509. http://www.aeaweb.org/articles?id=10.1257/app.20150509.

Hanna, R., Wang, S.Y., 2017. Dishonesty and selection into public service: evidence from India. Am. Econ. J.: Econ. Policy 9 (3), 262–290.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J., 2006. Costly punishment across human societies. Science 312 (5781), 1767–1770. https://doi.org/10.1126/science.1127333. http://science.sciencemag. org/content/312/5781/1767. http://science.sciencemag.org/content/312/5781/1767.full.pdf.

Henrich, N., Henrich, J.P., 2007. Why Humans Cooperate: A Cultural and Evolutionary Explanation. Oxford University Press.

Hvide, H.K., Lee, J., Odean, T., 2019. Easy Money, Cheap Talk, or Spuds: Inducing Risk Aversion in Economics Experiments. SSRN.

Jordan, J.J., Hoffman, M., Bloom, P., Rand, D.G., 2016. Third-party punishment as a costly signal of trustworthiness. Nature 530 (7591), 473–476.

Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–291. http://www.jstor.org/stable/1914185.

Karl, T.L., 2007. Ensuring fairness: the case for a transparent fiscal social contract. In: Humphreys, M., Sachs, J., Stiglitz, J.E. (Eds.), Escaping the Resource Curse. Columbia University Press, pp. 258–293.

Köszegi, B., Szeidl, A., 2013. A model of focusing in economic choice. Q. J. Econ. 128 (1), 53–104. http://search.ebscohost.com/login.aspx?direct=true&db=bth& AN=85819361&site=ehost-live.

Levi, M., 1988. Of Rule and Revenue. University of California Press. http://www.jstor.org/stable/10.1525/j.ctt1pngtk.

Mahdavy, H., 1970. The patterns and problems of economic development in Rentier States: the case of Iran. In: Cook, M. (Ed.), Studies in the History of the Middle East from the Rise of Islam to the Present Day. Oxford University Press, pp. 428–467.

Martin, L., 2014. Taxation, loss aversion and accountability: Theory and experiment evidence for taxation's effect on citizen behavior. Working Paper. Yale University.

Martínez, L.R., 2016. Sources of revenue and government performance: evidence from Colombia. University of Chicago. Mimeo.

Mehlum, H., Moene, K.O., Torvik, R., 2006. Institutions and the resource curse. Econ. J. 116 (508), 1–20.

Moore, M., Prichard, W., Fjeldstad, O.H., 2018. Taxing Africa: Coercion, Reform and Development. Zed Books Ltd.

Paler, L., 2013. Keeping the public purse: an experiment in windfalls, taxes, and the incentives to restrain government. Am. Polit. Sci. Rev. 107 (4), 706–725.

Prichard, W., 2015. Taxation, Responsiveness, and Accountability in Sub-Saharan Africa: The Dynamics of Tax Bargaining. Cambridge University Press.

Prichard, W., Salardi, P., Segal, P., 2018. Taxation, non-tax revenue and democracy: new evidence using new cross-country data. World Dev. 109, 295–312. https:// doi.org/10.1016/j.worlddev.2018.05.014. http://www.sciencedirect.com/science/article/pii/S0305750X18301621.

Rand, D.G., Nowak, M.A., 2013. Human cooperation. Trends Cogn. Sci. 17 (8), 413–425.

Ross, M.L., 2001. Does oil hinder democracy? World Polit. 52 (3), 325–361. http://www.jstor.org/stable/25054153.

Ross, M.L., 2015. What have we learned about the resource curse? Annu. Rev. Pol. Sci. 18, 239–259.

Sandbu, M.E., 2006. Natural wealth accounts: a proposal for alleviating the natural resource curse. World Dev. 34 (7), 1153–1170.

State of Alaska, 2023. Permanent Fund Dividend. Website. https://pfd.alaska.gov/.

Villeval, M.C., 2020. Public goods, norms and cooperation. In: Capra, M.C., Croson, R., Rigdon, M., Rosenblat, T. (Eds.), Handbook of Experimental Game Theory. Edward Elgar Publishing.

Vossler, C.A., McKee, M., Bruner, D.M., 2021. Behavioral effects of tax withholding on tax compliance: implications for information initiatives. J. Econ. Behav. Organ. 183, 301–319. https://doi.org/10.1016/j.jebo.2020.12.030. https://www.sciencedirect.com/science/article/pii/S0167268121000019.

Weigel, J.L., 2020. The participation dividend of taxation: how citizens in Congo engage more with the state when it tries to tax them. Q. J. Econ., 1–55. https:// academic.oup.com/qje/article-pdf/doi/10.1093/qje/qjaa019/33498167/qjaa019.pdf.