

Accountability and taxation: Experimental evidence*

Ingrid Hoem Sjursen[†]

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). This paper presents evidence from a between-subject experiment that tests the effect of taxation on demand for accountability and the underlying mechanisms explaining this 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 being collected as tax. These features are theorized to increase the salience of fairness considerations in public service provision, and this increased salience of fairness is in turn hypothesized to increase demand for accountability. The main finding is that taxation causes a higher demand for accountability when both features of taxation are present. This result is evidence in support of the Rentier State Hypothesis.

Keywords: Taxation, experiment, fairness, behavioral economics, accountability
JEL Classification: H27, C91, D63, D90

*I am grateful to Bertil Tungodden for invaluable advice and inputs. I would also like to thank Ragnar Torvik, Alexander W. Cappelen, Erik Sørensen, Kjetil Bjorvatn, Ingvild Almås, Vincent Somville, Lars Ivar Oppedal Berge, Odd-Helge Fjeldstad, Ivar Kolstad, Arne Wiig, Björn Bartling, Uri Gneezy, Hallgeir Sjåstad, Ranveig Falch, Sandra Halvorsen and Charlotte Ringdal for valuable inputs. Thanks to participants at various seminars and conferences for useful feedback, and to Sebasitan Fest for letting me use an adapted version of his picture categorization task. I am thankful to the Center of Ethics and Economics and The Choice Lab at the Norwegian School of Economics, and Ingegerd and Arne Skaugs Forskningsfond for economic support.

[†]Affiliations: Centre for Applied Research (SNF) at NHH, FAIR - The Choice Lab, and Norwegian Center for Taxation (NoCeT), Norwegian School of Economics (NHH), email: Ingrid.Sjursen@snf.no.

1 Introduction

Corruption and mismanagement of public revenue are a widespread and serious obstacle to social and economic development in many poor countries (Ferraz and Finan, 2011; Lederman et al., 2005; UNDP, 2008). The so called *Rentier State Hypothesis*, coined by Borge et al. (2015), 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; Paler, 2013; Ross, 2001). Thus, the Rentier State Hypothesis suggests that financing public revenue through taxing citizens may promote political accountability and development.

Despite the popularity of the Rentier State Hypothesis, we lack causal evidence for the positive effect of taxation on accountability, and for the underlying mechanisms explaining this relationship. The focus of this paper is on the first part of the Rentier State Hypothesis: The effect of taxation on citizens' demand for accountability.¹ Accountability can be defined as “The quality of being accountable; liability to account for and answer for one’s conduct, performance of duties, etc. (...); responsibility.” (OED, 2017). Demand for accountability is typically thought to consist of willingness to i) monitor the behavior of the government and ii) impose a cost on the government when its behavior is not accountable, i.e., when the government’s actions are not responsible and in the best interest of the citizen, possibly at a cost. The focus of this paper is on the willingness to impose a cost on the government. As an example, citizens may sacrifice time and transportation costs to participate in a demonstration calling on politicians to be more accountable, imposing a political cost on the government as punishment for not being accountable.

The present paper reports results from a between-subjects experiment where 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 an endowment given to the group to invest in a common pool and how much to keep for him or herself. The citizen can costly punish the leader’s decision by reducing his or her payoff, but has no positive monetary incentives to do so. The citizen’s willingness to punish the leader thus captures the main features of, and is used as a proxy for, demand for accountability in the experiment. The experiment exogenously varies how the group endowment is financed in order to capture two major differences between tax and windfall revenue. First, the money on which citizens are taxed is *earned* by their work, whereas windfall revenue is not related to citizens’ effort and is *unearned*. Second, tax revenue initially accrues to citizens and is subsequently collected as tax, whereas windfall revenue accrues directly to the government. These two distinguishing features of tax revenue are theorized to increase the salience of fairness considerations to citizens, because the features entail that citizens actively contribute to the financing of revenue. Increased salience of fairness is, in turn, hypothesized to make citizens care more about how resources are distributed and increase their willingness to punish the government for unfair behavior.

The experiment has two treatments designed to capture these differences between tax and windfall revenue. In the “Rentier State” treatment, the group endowment is windfall and non-tax. In the “Tax State” treatment, the group endowment is produced by the work

¹From now on, the term “Rentier State Hypothesis” will be used to refer to the effect of taxation on demand for accountability.

of the citizen and the leader, and financed through a tax on earnings. The Rentier State treatment can be thought of as a stylization of the situation in a Rentier State, where government revenue is windfall revenue and taxation is low, and the Tax State treatment can be thought of as a stylization of a situation in a Tax State, where government revenue is produced by citizens' work and collected through taxation. Comparing punishment behavior in the two treatments provides a causal test of the effect of taxation on demand for accountability, and thus of the Rentier State Hypothesis.

The experiment also includes two treatments designed to shed light on the underlying mechanisms explaining the Rentier State Hypothesis. In the "Hard Earned & Non-Tax" treatment, the group endowment is produced by the work of the citizen and the leader, but is not collected as tax, and in the "Windfall & Tax" treatment, the group endowment is windfall, but is financed through taxation. Together with the main treatments, these additional treatments enable causal tests of the separate effect of the revenues i) being hard earned ("Hard Earned mechanism") and ii) having been in the possession of the citizen before being collected as tax ("Possession mechanism").

The paper offers three main findings. First, taxation causes a significant increase in citizens' demand for accountability. Demand for accountability is proxied by citizens' willingness to punish. It is defined as the highest investment for which the citizens punish the leader, and measured using the strategy method. Figure 1 shows the means and standard errors for citizens' willingness to punish in the two main treatments. It shows that when the group endowment is produced by the citizens' work and 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. This result provides evidence in support of the Rentier State Hypothesis. The finding is robust across different subgroups of the sample (gender, age, education, employment status, income, political view and political engagement).

[Figure 1 about here.]

Second, while the Hard Earned and the Possession mechanism increase citizens' willingness to punish when working in combination, neither has a separate significant effect. Thus, both the Hard Earned and the Possession mechanism are important in explaining the effect of taxation on willingness to punish. Third, the effect of taxation on willingness to punish can be partly explained by taxation causing citizens to have stronger negative emotions when the leader is unfair. Citizens in the Tax State treatment report that they would be more upset than citizens in the Rentier State treatment do if the leader invests less than what they perceive to be fair, and willingness to punish is strongly correlated with negative emotions.

The paper also investigates the effect of taxation on the leaders' decisions. The results show that leaders invest more in the common pool (i.e., are more accountable) when the group endowment is produced by work than when it is windfall revenue.

Overall, the results suggest that taxation is not only a means to generate government revenue, but may also increase citizens' demand for accountability, which is often assumed to enhance social and economic development through more accountable government spending. The results also imply 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, because this increases the salience of

fairness considerations.

The paper relates to the empirical literature investigating how the source of public revenue affects the political behavior of citizens and governments. A large strand of this literature uses cross- and within-country variation in observational data to investigate the relationship between reliance on tax or windfall revenue and governance indicators such as democratization, provision of public goods, and corruption (Andersen and Aslaksen, 2013; Aslaksen, 2007; Arezki and Brückner, 2011; Besley and Persson, 2013; 2014; Brollo et al., 2013; Caselli and Michaels, 2013; Gadenne, 2017; Leite and Weidmann, 2002; Martínez, 2016; Prichard et al., 2014; Ross, 2004; Sala-i-Martin and Subramanian, 2013). The studies largely show that good governance is positively associated with reliance on tax revenues and negatively associated with reliance on windfall revenues, with some notable exceptions (Haber and Menaldo, 2011; Herb, 2005).

However, most of these studies do not identify a causal relationship between tax or windfall revenue and accountability, and do not provide evidence for the underlying mechanisms explaining it. The present study contributes to this literature by providing a test of the first part of the suggested causal chain of the Rentier State Hypothesis, namely the effect of taxation on demand for accountability. Another contribution of the paper is to propose and test a new mechanism for the effect of taxation on demand for accountability, namely that it increases the salience of fairness considerations in government spending. This, in turn, is hypothesized to increase citizens' willingness to punish unfair spending.

The paper also relates to a handful of experimental studies that show mixed results for the effect of taxation on demand for accountability. First, in a lab experiment in Uganda, Martin (2016) finds that individuals have a higher willingness to punish a dictator when the group endowment is collected as tax from them, compared to when it is given directly to the dictator. Second, Bosman et al. (2005) report results from lab experiments in the Netherlands and Austria and find that individuals are *less* willing to punish a second party appropriating some of their endowment when this endowment is hard earned compared to when it is windfall. Third, Paler (2013) uses games to prime citizens to think about local government revenues as tax or windfall revenues in a field experiment in Indonesia. She finds that the tax priming leads to a higher stated willingness to monitor the government, but does not significantly affect participatory behavior. Finally, Weigel (2017) randomizes tax collection in a field experiment in D.R. Congo and finds that citizens residing in taxed areas have a higher demand for accountability, measured as attendance in town hall meetings about taxation and likelihood of submitting a suggestion card to the government. The present experiment contributes to this experimental literature in three ways. First, the present design rules out information as an alternative explanation, by holding it constant across treatments. Other papers have argued that taxation increases demand for accountability through *information* because the citizen is typically aware of how much he or she pays in tax, but not of how much the government earns from natural resource exports or receives in aid (see for instance Gadenne (2017), Paler (2013) and Sandbu (2006)). Second, the present study implements an experimental design that allows for the identification of the interaction effect between the Hard Earned and the Possession mechanism. Third, the experiment cleanly identifies the effect of taxation on how much citizens care about the leader investing a fair share by keeping the level of investment that is perceived as fair constant.

Finally, the paper relates to the large literature on social norms and human cooperation. Using public good and ultimatum games, this literature has established that individuals are willing to punish norm violators, even when such punishment is costly

and yields no material gains (de Quervain et al., 2004; Fehr and Gächter, 2000; 2002; Fehr et al., 2002; Fehr and Fischbacher, 2003; 2004a;b; Henrich et al., 2006; Henrich and Henrich, 2007; Jordan et al., 2016; Rand and Nowak, 2013; Xiao and Houser, 2005). The present paper contributes to this literature by investigating whether and why costly punishment of an investment decision depends on the source of income, and, thus, shedding more light on motivations behind punishment.

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

2 Experimental design and sample

This section explains the sequence of events, the structure of the experimental task, the treatments, and the setting and sample.²

2.1 Sequence of events

In the experiment, the participants play a two-player investment game where the citizen has the opportunity to costly punish the leader for his or her investment decision. The sequence of events is illustrated in Figure ?? and can be described as follows. In the first stage (uppermost panel, Figure ??), 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. Only pairs where both participants choose to work are included in the empirical analysis and the remaining explanation will focus on these. 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 2.2 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. 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

²The instructions for the main parts of the study are provided in Appendix G.

is his or her demand for accountability.

[Figure ?? about here.]

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.

2.2 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. 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 Figure 3 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 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 percent 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 percent 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.³ The four treatments can be summarized as follows:

Rentier State (T1): Citizen and leader each earn \$1. A windfall finances the \$2

³Please see Appendix A for a 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 A.1 show, manipulating duration does not yield results different from the main manipulation (productivity). The two treatments are therefore pooled in all the analyses presented in the paper unless otherwise specified.

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 percent. 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 percent. The \$1 tax collected from each finances the \$2 group endowment.

[Figure 3 about here.]

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 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. 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.

2.3 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 and were therefore dropped from the sample.⁴ Thus, the main analyses in this paper are based on a sample

⁴A 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

of 739 citizens. Descriptive statistics for these are reported in Table 1. Columns (1) - (5) show means for background variables for each of the four treatments and for the pooled sample, respectively. Column (5) 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.⁵ The sample is otherwise balanced.

[Table 1 about here.]

Tables B.5 and B.6 in Appendix B show that there are some significant differences in the share of citizens that choose not to work and the share of citizens that are inconsistent punishers between treatments. However, the balance regressions for all the 983 citizens recruited in Table B.2 in Appendix B show that the identified imbalances are not due to non-workers and inconsistent punishers being significantly different from workers and consistent punishers in terms of observable variables. This mitigates the concern that differential selection of consistent punishers between treatments is driving the treatment effects.

3 Theoretical framework

This section describes the theoretical framework guiding the experimental design. First, the citizen’s punishment decision is considered. His or her monetary payoff can be formulated as:

$$y_c = R + \frac{3}{4}I - (p \times c), \quad (1)$$

where 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. As Equation (1) shows, there are no monetary incentives to punish in the experiment. Furthermore, 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; Xiao and Houser, 2005). Fehr et al. (2002) refer to such behavior as *strong negative reciprocity*.⁶

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. Negative

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.

⁵Table B.1 in Appendix B provides a more detailed balance check.

⁶Evolutionary 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).

emotions are one possible mechanism through which this effect might work. Unfair investments may upset citizens, and punishing the leader is a way for them to express their anger to the leaders (Fehr and Gächter, 2002; Hopfensitz and Reuben, 2009; Nelissen and Zeelenberg, 2009; Xiao and Houser, 2005). Based on this, the citizen's punishment decision can be modeled as a tradeoff between monetary payoff and utility from punishing the leader for unfair investments.⁷ To formalize ideas, the following simple model of the citizen's utility, inspired by Cappelen et al. (2007), is introduced:

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

m_c is the citizen's fairness norm and specifies the investment share that he or she perceives to be fair. Thus, $m_c - I$ indicates how much the leader's investment deviates from the amount the citizen perceives as fair. γ_c is a parameter determining the citizen's non-monetary unconditional utility loss from unfair investments, and β is a parameter determining the citizen's utility from punishing unfair investments. The rest of the parameters are as defined for Equation (1) above. Equation (2) 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 , and negatively on the cost of punishment, c . It is assumed that the citizen derives negative unconditional utility from both positive and negative deviations from the fairness norm, but only derives utility from punishment of negative deviations. Both the unconditional disutility and the utility from punishment are assumed to be increasing in the size of the deviation from the fairness norm. The citizen's punishment behavior is characterized by:

$$c < \beta [\min\{0, m_c - I\}^2] \quad \text{Punish} \quad (3a)$$

$$c = \beta [\min\{0, m_c - I\}^2] \quad \text{Indifferent, randomize} \quad (3b)$$

$$c > \beta [\min\{0, m_c - I\}^2] \quad \text{Not punish} \quad (3c)$$

Next, the leader's investment decision is considered. It is modeled as a trade-off between expected monetary payoff and non-monetary disutility from deviations from fairness:

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

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 (1). $m_l - I$ 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:

$$I^* = m_l - \frac{1}{2\gamma_l} \left[\frac{1}{4} + \phi'(I) \times \theta \right]. \quad (5)$$

⁷An alternative theoretical approach is to model the citizen's non-monetary utility from punishing the leader's investment as *reference dependent*, i.e., determined by his or her reference point for investment. This approach is discussed in more detail in Subsection E.2 in Appendix E.

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, $\phi'(I)$.

The theoretical model can now be used to illustrate the effect of taxation on the citizen’s punishment behavior. As Equations (3a)-(3c) show, for any given investment level (I), two parameters influence the citizen’s punishment in the model; the level of investment the citizen perceives as fair (m_c), and the utility he or she derives from punishment of unfair investments (β). Since we assume that the fairness norms, m_c and m_l , are the same and equal to GE in all treatments, the effects of treatments must work through changes in the utility derived from punishing unfair investments, β .

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 have worked hard to earn the money that finances it, and they have had the tax money in their possession before it was 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), leading to higher willingness to punish.⁸ In the absence of taxation (Rentier State treatment), neither the citizen nor the leader actively contributes to the group endowment. Even though the citizen’s 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 citizens less willing to punish unfair investments. Accordingly, the theory predicts willingness to punish and investment levels to be higher in the Hard Earned & Non-Tax, Windfall & Tax and Tax State treatments than in the Rentier State treatment.

Next, the model is used to illustrate the effect of taxation on the leader’s investment decision. Equation (5) 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

⁸This 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 overweighed 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.

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.

To summarize, the theoretical framework predicts that taxation increases the citizen’s willingness to punish deviations from the fairness norm through increasing the salience of this norm, and thus increases the utility from punishment of deviations from it. Correspondingly, taxation is predicted to increase the leader’s investment in the common pool because it increases the salience of the fairness norm and thus increases the disutility derived from deviating.

4 Empirical strategy

This section describes the empirical strategy for the analysis.⁹

4.1 Main analysis

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

$$y_i = \alpha + \beta^H H_i + \beta^T T_i + \theta^{H \times T} H_i \times T_i + \beta^X X_i + \beta^Z Z_i + \varepsilon_i. \quad (6)$$

y_i is the standardized willingness to punish for the citizen in pair i . α is a constant, H_i is an indicator variable taking the value of one for individuals in the Hard Earned treatments, T2 and T4. T_i is an indicator variable taking the value of one for individuals in the Tax treatments, T3 and T4, and $H_i \times T_i$ is an interaction term between H_i and T_i . $H_i \times T_i$ takes the value of one for individuals in the Tax State treatment (T4). 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 (6) are estimated; one including the treatment variables and their interaction term only; one including the treatment variables, their interaction term and X_i ; and one with treatment variables, their interaction term, X_i and Z_i , all OLS regressions with robust standard errors. The Rentier State treatment is the reference category, and the estimation of Equation (6) gives the following main parameters of interest:

$\beta_T + \beta_H + \theta^{H \times T}$: causal effect of going from the Rentier State treatment to the Tax State treatment. This tests the Rentier State Hypothesis.

⁹This, 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.socialscisearch.org/trials/2233>, registration number AEARCTR-0002233. The paper mainly follows the pre-analysis plan, with some minor deviations and a few additional specifications. Please see Appendix C for an overview.

β_T , $\beta_T + \theta^{H \times T}$: causal effect of going from the Rentier State to the Windfall & Tax treatment, and the causal effect of going from the Hard Earned & Non-Tax treatment to the Tax State treatment, respectively. These test the Possession mechanism.

β_H , $\beta_H + \theta^{H \times T}$: causal effect of going from the Rentier State treatment to the Hard Earned & Non-Tax treatment, and from the Windfall & Tax treatment to the Tax State treatment, respectively. These test the Hard Earned mechanism.

$\theta^{H \times T}$: the difference in causal effect of going from Windfall to Hard Earned treatments between Non-Tax and Tax treatments.

To investigate whether the treatments affect willingness to punish on the intensive or extensive margin, Equation (6) is also estimated for two indicator variables; one for high punishment, taking the value of one for citizens that always punish, or punish investment shares of 0.9, and zero otherwise (intensive margin); and one for positive punishment, taking the value of one for citizens that punish at least one investment share and zero for those who never punish (extensive margin).

To check the robustness of the main results Equation (6) is estimated for an alternative definition of the dependent variable, namely *the lowest share invested for which the citizen does not punish the leader*.

4.2 Heterogeneity analysis

To investigate whether different subgroups of the sample respond differently to the treatments, the following specification is estimated:

$$y_i = \alpha + \beta^H H_i + \beta^T T_i + \theta^{H \times T} H_i \times T_i + \beta^{\text{Var}} \text{Var}_i + \theta^H H_i \times \text{Var}_i + \theta^T T_i \times \text{Var}_i + \gamma H_i \times T_i \times \text{Var}_i + \beta^X X_i + \beta^Z Z_i + \varepsilon_i. \quad (7)$$

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). $H_i \times \text{Var}_i$ is an interaction term between H_i and Var_i , $T_i \times \text{Var}_i$ is an interaction term between T_i and Var_i , and $H_i \times T_i \times \text{Var}_i$ is a term for the triple interaction between H_i , T_i and Var_i .

The reference category is the subgroup for which Var_i takes the value of zero in the Rentier State treatment. To illustrate, if Var_i is the indicator variable for male, the reference category is female (male = 0) in the Rentier State treatment. Then β^H is the effect of going from the Rentier State to the Hard Earned & Non-Tax treatment for females, β^T is the effect of going from the Rentier State to the Windfall & Tax treatment for females, and so on. Estimating Equation (7) for each background variable gives the following parameters of interest:

$\beta_T + \beta_H + \theta^{H \times T}$: causal effect of going from the Rentier State to the Tax State treatment for subgroup $\text{Var}_i = 0$ (for instance females).

$\beta_T + \beta_H + \theta^{H \times T} + \theta^H + \theta^T + \gamma$: causal effect of going from the Rentier State to the Tax State treatment for subgroup $\text{Var}_i = 1$ (for instance females).

$\theta^H + \theta^T + \gamma$: difference in effect of going from the Rentier State to the Tax State treatment between $\text{Var}_i = 0$ and $\text{Var}_i = 1$ (for instance between females and males).

4.3 Leader decisions

To investigate the effect of treatments on leaders' investment decisions, Equation (6) is estimated using the standardized share invested in the common pool as the dependent variable.

5 Results

This section reports the results. The first part reports descriptive findings for the citizens' punishment behavior and for self-reported negative emotions. The second part presents the main analysis of the effect of taxation on the willingness to punish and on negative emotions. The third part presents the heterogeneity analysis. Finally, results from the leaders' investment behavior are described in the fourth part.

5.1 Descriptive statistics

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 0.1, an investment share of 0.2, an investment share of 0.3 and so on up until an investment share of 1. These decisions 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. The punishment decision that the citizen makes for the leader's actual investment decision is implemented. 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. 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.

Figure 4 divides citizens into four categories according to their punishment behavior. It shows that a significant share, about 45 percent, never punish the leader. The remaining 55 percent have a positive willingness to punish for at least one level of leader investment, most of whom (about 37 percent of the sample) are willing to punish high investments, meaning that the highest investment share they punish is between 60 and 90 percent. A small share (about 15 percent of the sample) are only willing to punish low investments, meaning that the highest investment share they punish is between 0 and 50 percent. A minority of three percent always punish the leader, even in the case when he or she invests 100 percent of the group endowment. Figure D.1 in Appendix D gives a more detailed description of punishment behavior by illustrating the entire distribution of willingness to punish.

[Figure 4 about here]

To shed more light on why citizens punish, they answer the unincentivized question "How upset would you be if the leader invests less than the fair share, when both of you completed the assignment?" Figure 5 illustrates their answers measured on an 11-point scale from "Not upset at all" to "Very upset". It shows that there is a large variation in

negative emotions associated with unfair leader investments. The figure also shows that the distribution of negative emotions is skewed to the right of the mean upsetness level, indicating that the majority of citizens report that they would be somewhat or more upset if the leader is unfair.

[Figure 5 about here.]

Figure 6 investigates the relationship between negative emotions and punishment behavior. The left panel divides the sample of citizens into two groups according to self-reported emotions; those who report to be less than or equal to the median level of upset (six), and those who report to be more than median upset if the leader is unfair. It illustrates the willingness to punish in these two groups and shows that the more upset punish significantly higher investment shares than the less upset. The magnitude of this difference is equivalent to 0.6 standard deviation and is highly significant ($p = 0.000$). The right panel of the figure illustrates the same relationship for all possible outcomes of the upsetness variable. Each bar illustrates the mean willingness to punish for that level of upsetness and shows that willingness to punish is linearly increasing in the strength of negative emotions reported. These results strongly suggest that negative emotions are an important driver of punishment behavior.

[Figure 6 about here.]

Table 2 reports regressions of willingness to punish on an indicator variable taking the value of one for citizens that are more than median upset if the leader is unfair in Columns (1)-(3) and on the 11-point scale measure of upsetness in columns (4-6). The table shows that the positive correlation between negative emotions and punishment replicates in a regression framework for both measures of upsetness.

[Table 2 about here.]

5.2 Main analysis

As illustrated in Figure 1, 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 citizen and the leader's 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 investigates how willingness to punish is affected by the Tax and Hard Earned manipulations, their interaction, and background variables, in a regression framework. Columns (1)-(3) report results for willingness to punish measured by a standardized version of the 11-point punishment scale. Column (1) reports estimates of Equation (6) 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. Going

from the Rentier State to the Tax State treatment significantly increases the willingness to punish (as indicated by the positive and significant coefficient of “Hard Earned + Tax + Hard Earned x Tax”, $p = 0.058$). The table also shows the estimated separate effects of the Hard Earned and Tax manipulations. The coefficients of “Hard Earned” and “Hard Earned + Hard Earned x Tax” test the separate effect of the Hard Earned manipulation when the group endowment is non-tax and when it is collected through taxation, respectively. They show that the Hard Earned effects are positive, but not statistically significant. Similarly, the coefficients of “Tax” and “Tax + Hard Earned + Tax” test the separate effect of the Tax manipulation when the group endowment is windfall and when it is produced by work, respectively. They show that the effects are positive, but not significant. The small and non-significant coefficient of “Hard Earned x Tax” suggests that there is no interaction effect between the two mechanisms on the willingness to punish.

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.

[Table 3 about here.]

Columns (4)-(6) report estimates of Equation (6) where the dependent variable is a standardized version of an indicator variable taking the value of one for individuals that punish very high shares invested (0.9 or 1): Column (4) reports estimates from the specification with treatment variables and the interaction only, and Columns (5) and (6) sequentially add the background political variables. The main result from Column (3) holds, and is slightly strengthened, when investigating the effect on the indicator for high willingness to punish in Column (6). For this specification, a significant effect of the Tax manipulation is also identified for the Hard Earned treatments; citizens are more likely to punish high investment shares when the group endowment is hard earned and collected through taxation than when it is hard earned and given directly to the leader.

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 ($\beta^H + \beta^T + \theta^{H \times T} = 0.186$, $p = 0.058$, Column (3)).*

Result 2 (Mechanisms): *The positive effect of the Tax State Treatment on the willingness to punish is explained by a combination of the Hard Earned and Possession mechanism. Both the Tax and the Hard Earned manipulations have a positive separate effect on the willingness to punish, but they are not statistically significant (Hard Earned mechanism: $\beta^H = 0.040$, $p = 0.715$, $\beta^H + \theta^{H \times T} = 0.057$, $p = 0.595$, Possession mechanism: $\beta^T = 0.129$, $p = 0.269$, $\beta^T + \theta^{H \times T} = 0.147$, $p = 0.134$, Column (3)). There is no interaction effect between the Hard Earned and the Possession mechanism ($\theta = 0.018$, $p = 0.907$).*

Result 1 provides evidence in support of the Rentier State Hypothesis. When the financing of the group fund 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.¹⁰

Result 2 shows that both the Hard Earned and the Possession mechanism are needed in order to generate the Rentier State effect. These mechanisms significantly affect the willingness to punish only when working in combination.

To further investigate the effect of taxation on the willingness to punish, Table D.3 in Appendix D reports regression results for an indicator variable taking the value of one if the citizen punishes some investment share. It shows that the treatments have no significant effects on the likelihood of a citizen punishing the leader. Thus, the results in Tables 3 and D.3 indicate that 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 D.3 also reports the results for a sample that includes the inconsistent punishers in Columns (4)-(6). These regressions show that the two samples give very similar results.

Finally, the treatment effect on negative emotions is investigated in Figure 7. It shows that, on average, citizens in the Tax State treatment report to be more upset than the citizens in the Rentier state treatment do (6.3 vs. 5.9 on a scale from 1 to 11), indicating that the Tax State treatment increases the citizens' negative emotions. The difference in negative emotions reported is equivalent of 0.15 standard deviation, but is not significant at conventional levels of significance ($p = 0.131$). Table 4 reports the corresponding regression analysis where negative emotions are regressed on indicator variables for treatments and their interaction. The table replicates the positive effect of the Tax State treatment on negative emotions. Additionally, it shows that the point estimates of the separate effects of the Hard Earned and the Tax manipulation on negative emotions go in opposite directions.

[Figure 7 about here.]

[Table 4 about here.]

The results show that willingness to punish is highly correlated with negative emotions. They also provide suggestive evidence in support of negative emotions being a mechanism for the effect of taxation on the willingness to punish: Taxation makes the citizen more upset by, and therefore more willing to punish, unfair leader investments.

5.3 Heterogeneity analysis

This section investigates whether subgroups of citizens respond differently to taxation. Figure 8 illustrates the difference in mean willingness to punish between the Rentier State and the Tax State treatment for the different subgroups of respondents. To illustrate, the left bars in the top left panel of the figure show the mean willingness to punish for women in the Rentier State treatment (dark gray) and women in the Tax State treatment (light gray), respectively. The two right bars of the panel show the mean willingness to punish for men in the Rentier State (dark gray) and Tax State (light gray) treatment. The panel illustrates that going from the Rentier State to the Tax State treatment causes men to have a significantly higher willingness to punish, but does not

¹⁰The estimated regressions for the robustness check of the Hard Earned mechanism are reported in Table A.1 in Appendix A. It shows that there is no significant difference in willingness to punish between citizens in the five- and 10-minute Hard Earned treatments.

affect the punishment behavior of women. Overall, the figure shows that Result 1 is robust across the different subgroups: Willingness to punish is consistently higher in the Tax State than in the Rentier State treatment in all subgroups except non-full-time employees (where the difference between treatments is not significant).

Table 5 reports the estimated regression coefficients for the effect on the willingness to punish of going from the Rentier State to the Tax State treatment for the different subgroups of the sample, as well as the difference in treatment effect between the groups. It shows that the Tax State treatment has a significant effect on the following subgroups: male, younger, more educated, full-time employees, richer, more politically conservative and the more politically engaged citizens, but that the difference in treatment effect is only significantly different between non-full-time and full-time employees. Tables D.1 and D.2 in Appendix D report the full regressions (see Tables D.5 and D.6 for specifications without controls.)

[Figure 8 about here.]

[Table 5 about here.]

Based on Table 5, the following main result for the heterogeneity analysis can be formulated:

Result 3 (Rentier State Hypothesis): *The positive effect of taxation on the willingness to punish is robust across all subgroups of respondents, with the exception of employment status.* Across gender, age, education, income, political view, political engagement and negative emotions, the Tax State treatment increases the willingness to punish compared to the Rentier State treatment ($\beta^T + \beta^H + \theta^{H \times T}$ and $\beta^T + \beta^H + \theta^{H \times T} + \theta^H + \theta^T + \gamma > 0$, Columns (1)-(3) and (5)-(7)), though not significantly in all groups.

Result 3 shows that Result 1, the Rentier State effect, is robust across the different subgroups. This indicates that the effect of taxation on willingness to punish is a general one and not driven by specific subgroups.

5.4 Leader decisions

This subsection investigates the leaders' decisions for investment in the common pool.¹¹ Figure 9 illustrates the distribution of leader investment behavior: each bar indicates the share of leaders that invested the given share in the common pool. It shows that the majority of leaders, 60 percent, invest everything. About 25 percent keep the entire group endowment to themselves and 4 percent invest half.

[Figure 9 about here.]

Next, the effect of taxation on leader investments is investigated. Figure 10 illustrates the effect of going from the Rentier State to the Tax State treatment on the share invested by the leader. It shows that the Hard Earned and the Possession mechanism do not significantly affect the leader's investment behavior when working in combination. Table 6 reports the estimated results of Equation (6) taking the standardized value of

¹¹Balance tables for leaders are reported in Tables F.1 - F.4 in Appendix F.

the leader’s investment as the dependent variable. The regressions replicate the descriptive result from Figure 10 for both the 11-point scale and the binary definition of leader investment (the binary outcome takes the value of one for leaders that invest everything in the common pool and zero otherwise). The regressions furthermore show that when the group endowment is collected as tax, going from the Windfall to the Hard Earned treatment significantly increases the investment share. This is indicative evidence of taxation positively affecting leader behavior through the Hard Earned mechanism.

[Figure 10 about here.]

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. Figure F.1 in Appendix F shows that the effect of the Hard Earned treatment on investment share is not driven by beliefs about citizens’ punishment, providing suggestive evidence that the effect of taxation on provision of accountability is not driven by beliefs about punishment.

[Table 6 about here.]

Based on the analysis of leader decisions, the following result can be formulated:

Result 4 (Leader investment): *Leaders invest more in the common pool when the group endowment is produced by work, but the effect is only statistically significant when the group endowment is also collected through taxation ($\beta^H + \theta = 0.167$, $p = 0.067$ and $\beta^H = 0.050$, $p = 0.616$, Column (3)).*

Result 4 provides suggestive evidence of taxation increasing the share invested by the leader.

6 Discussion and conclusion

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, as underlying mechanisms explaining why it is perceived differently; tax revenue is produced by the citizens’ work and has been in their possession before being collected as tax. The paper offers three main findings. First, when revenue is tax revenue this causes a higher demand for accountability, measured as citizens’ willingness to costly punish the leader’s investment decision for the group endowment. Citizens have a significantly higher willingness to punish when the group endowment 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. Second, the two distinguishing features of tax revenue, that it is hard earned and has been in the citizen’s possession before being collected through taxation, do not have separate significant effects on the willingness to punish. Third, the effect of taxation on willingness to punish can be partly explained by negative emotions. Taxation causes citizens to have stronger negative emotions about unfair investments,

and citizens are more willing to punish the stronger their negative emotions are. This finding highlights the importance of emotions in decision-making.

The results provide important implications for our understanding of citizens' accountability behavior and for policy. First, 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. 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 Dhimi and al Nowaihi (2007) or Engström et al. (2015)). 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.

References

- Abeler, Johannes, Armin Falk, Lorenz Göette, and David Huffman (2011). “Reference points and effort provision,” *American Economic Review*, 101(2): 470–492.
- Andersen, Jørgen Juel and Silje Aslaksen (2013). “Oil and political survival,” *Journal of Development Economics*, 100(1): 89–106.
- Arezki, Rabah and Markus Brückner (2011). “Oil rents, corruption, and state stability: Evidence from panel data regressions,” *European Economic Review*, 55(7): 955 – 963.
- Aslaksen, Silje (2007). “Corruption and oil: Evidence from panel data,” Unpublished.
- Banerji, A. and Neha Gupta (2014). “Detection, identification, and estimation of loss aversion: Evidence from an auction experiment,” *American Economic Journal: Microeconomics*, 6(1): 91–133.
- Besley, Timothy and Torsten Persson (2013). “Chapter 2 - taxation and development,” in Martin Feldstein Alan J. Auerbach, Raj Chetty and Emmanuel Saez (eds.), “Handbook of Public Economics, vol. 5,” volume 5 of *Handbook of Public Economics*, Elsevier, pp. 51–110.
- Besley, Timothy and Torsten Persson (2014). “Why do developing countries tax so little?” *Journal of Economic Perspectives*, 28(4): 99–120.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2012). “Salience theory of choice under risk,” *Quarterly Journal of Economics*, 127(3): 1243–1285.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2013a). “Salience and asset prices,” *The American Economic Review*, 103(3): 623–628.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2013b). “Salience and consumer choice,” *Journal of Political Economy*, 121(5): 803–843.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2016). “Competition for attention,” *Review of Economic Studies*, 83(2): 481–513.
- Borge, Lars-Erik, Pernille Parmer, and Ragnar Torvik (2015). “Local natural resource curse?” *Journal of Public Economics*, 131: 101 – 114.
- Bosman, Ronald, Matthias Sutter, and Frans van Winden (2005). “The impact of real effort and emotions in the power-to-take game,” *Journal of Economic Psychology*, 26: 407–429.
- Bräutigam, Deborah, Odd-Helge Fjeldstad, and Mick Moore (2008). *Taxation and state-building in developing countries: Capacity and consent*, Cambridge University Press.
- Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini (2013). “The political resource curse,” *American Economic Review*, 103(5): 1759–96.
- Cappelen, Alexander W., Astri Drange Hole, Erik Ø. Sørensen, and Bertil Tungodden (2007). “The pluralism of fairness ideals: An experimental approach,” *American Economic Review*, 97(3): 818–827.

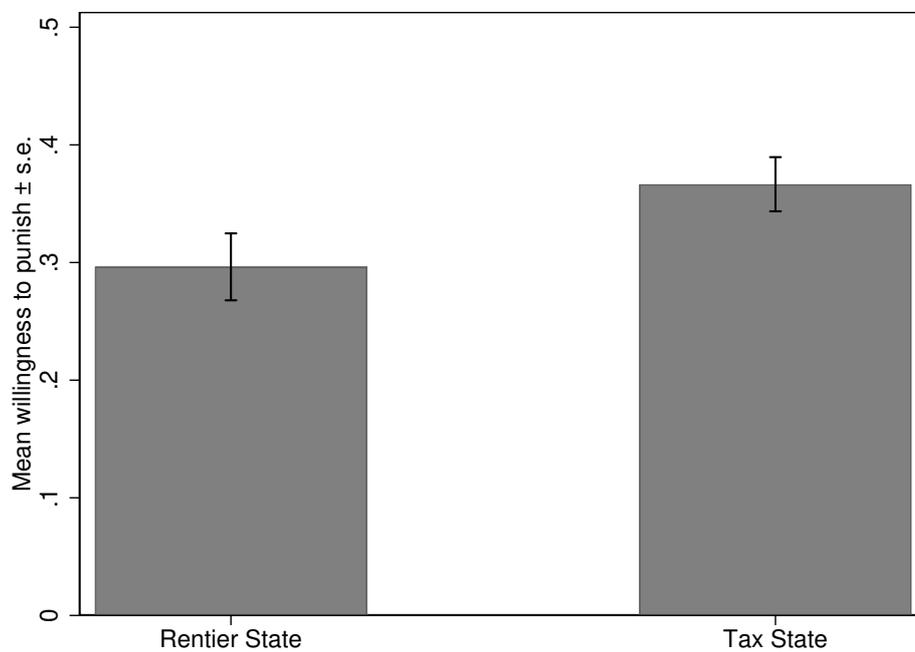
- Caselli, Francesco and Guy Michaels (2013). “Do oil windfalls improve living standards? Evidence from Brazil,” *American Economic Journal: Applied Economics*, 5(1): 208–238.
- Dharami, Sanjit and Ali al Nowaihi (2007). “Why do people pay taxes? Prospect theory versus expected utility theory,” *Journal of Economic Behavior & Organization*, 64(1): 171 – 192.
- Engström, Per, Katarina Nordblom, Henry Ohlsson, and Annika Persson (2015). “Tax compliance and loss aversion,” *American Economic Journal: Economic Policy*, 7(4): 132–64.
- Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde (2015). “The nature and predictive power of preferences: Global evidence,” IZA Discussion Paper. Bonn: Institute of Labor Economics.
- Fehr, Ernst and Urs Fischbacher (2003). “The nature of human altruism,” *Nature*, 425(6960): 785.
- Fehr, Ernst and Urs Fischbacher (2004a). “Social norms and human cooperation,” *Trends in Cognitive Sciences*, 8(4): 185–190.
- Fehr, Ernst and Urs Fischbacher (2004b). “Third-party punishment and social norms,” *Evolution and Human Behavior*, 25(2): 63–87.
- Fehr, Ernst, Urs Fischbacher, and Simon Gächter (2002). “Strong reciprocity, human cooperation, and the enforcement of social norms,” *Human Nature*, 13(1): 1–25.
- Fehr, Ernst and Simon Gächter (2000). “Fairness and retaliation: The economics of reciprocity,” *Journal of Economic Perspectives*, 14(3): 159–181.
- Fehr, Ernst and Simon Gächter (2002). “Altruistic punishment in humans,” *Nature*, 415: 137–140.
- Ferraz, Claudio and Frederico Finan (2011). “Electoral accountability and corruption: Evidence from the audits of local governments,” *American Economic Review*, 101(4): 1274–1311.
- Gadenne, Lucie (2017). “Tax me, but spend wisely? Sources of public finance and government accountability,” *American Economic Journal: Applied Economics*, 9(1): 274–314.
- Gill, David and Victoria Prowse (2012). “A structural analysis of disappointment aversion in a real effort competition,” *American Economic Review*, 102(1): 1–35.
- Haber, Stephen and Victor Menaldo (2011). “Do natural resources fuel authoritarianism? A reappraisal of the resource curse,” *The American Political Science Review*, 105(1): 1–26.
- Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, Frank Marlowe, David Tracer, and John Ziker (2006). “Costly punishment across human societies,” *Science*, 312(5781): 1767–1770.

- Henrich, Natalie and Joseph Patrick Henrich (2007). *Why humans cooperate: A cultural and evolutionary explanation*, Oxford University Press.
- Herb, Michael (2005). “No representation without taxation? rents, development, and democracy,” *Comparative Politics*, 37(3): 297–316.
- Hopfensitz, Astrid and Ernesto Reuben (2009). “The importance of emotions for the effectiveness of social punishment,” *The Economic Journal*, 119(540): 1534–1559.
- Jordan, Jillian J, Moshe Hoffman, Paul Bloom, and David G Rand (2016). “Third-party punishment as a costly signal of trustworthiness,” *Nature*, 530(7591): 473–476.
- Kahneman, Daniel and Amos Tversky (1979). “Prospect theory: An analysis of decision under risk,” *Econometrica*, 47(2): 263–291.
- Karl, Terry Lynn (2007). “Ensuring fairness: The case for a transparent fiscal social contract.” in Macartan Humphreys, Jeffrey Sachs, and Joseph E. Stiglitz (eds.), “Escaping the Resource Curse,” Columbia University Press, pp. 258 – 293.
- Köszegi, Botond and Matthew Rabin (2006). “A model of reference-dependent preferences,” *Quarterly Journal of Economics*, 121(4): 1133–1165.
- Köszegi, Botond and Matthew Rabin (2007). “Reference-dependent risk attitudes,” *American Economic Review*, 97(4): 1047–1073.
- Köszegi, Botond and Matthew Rabin (2009). “Reference-dependent consumption plans,” *The American Economic Review*, 99(3): 909–936.
- Köszegi, Botond and Adam Szeidl (2013). “A model of focusing in economic choice,” *Quarterly Journal of Economics*, 128(1): 53 – 104.
- Lederman, Daniel, Norman V. Loayza, and Rodrigo R. Soares (2005). “Accountability and corruption: Political institutions matter,” *Economics & Politics*, 17(1): 1 – 35.
- Leite, Carlos and Jens Weidmann (2002). “Does mother nature corrupt? Natural resources, corruption, and economic growth,” in George Abed and Sanjeev Gupta (eds.), “Governance, corruption and economic performance,” International Monetary Fund.
- Mahdavy, Hossein (1970). “The patterns and problems of economic development in rentier states: The case of Iran.” in M. Cook (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, Lucy (2016). “Taxation, loss aversion and accountability: Theory and experiment evidence for taxation’s effect on citizen behavior,” mimeo, University of North Carolina – Chapel Hill.
- Martínez, Luis R. (2016). “Sources of revenue and government performance: evidence from Colombia,” mimeo, University of Chicago.
- Marzilli Ericson, Keith M. and Andreas Fuster (2011). “Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments,” *Quarterly Journal of Economics*, 126(4): 1879 – 1907.

- Nelissen, Rob M. A. and Marcel Zeelenberg (2009). "Moral emotions as determinants of third-party punishment: Anger, guilt and the functions of altruistic sanctions," *Judgment and Decision making*, 4(7): 543.
- OED (2017). "accountability, n." in "Oxford English Dictionary Online," Oxford University Press.
- Paler, Laura (2013). "Keeping the public purse: An experiment in windfalls, taxes, and the incentives to restrain government," *American Political Science Review*, 107(4): 706–725.
- Prichard, Wilson, Paola Salardi, and Paul Segal (2014). "Taxation, non-tax revenue and democracy: New evidence using new cross-country data," ICTD Working Paper.
- de Quervain, D.J., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr (2004). "The neural basis of altruistic punishment," *Science*, 305(5688): 1254–1258.
- Rand, David G and Martin A Nowak (2013). "Human cooperation," *Trends in Cognitive Sciences*, 17(8): 413–425.
- Ross, Michael (2004). "Does taxation lead to representation?" *British Journal of Political Science*, 34: 229–249.
- Ross, Michael L. (2001). "Does oil hinder democracy?" *World Politics*, 53(3): 325–361.
- Sala-i-Martin, Xavier and Arvind Subramanian (2013). "Addressing the natural resource curse: An illustration from Nigeria," *Journal of African Economies*, 22(4): 570–615.
- Sandbu, M. E. (2006). "Natural wealth accounts: A proposal for alleviating the natural resource curse." *World Development*, 34(7): 1153–1170.
- UNDP (2008). *Primer on corruption and development*, United Nations Development Programme.
- Weigel, Jonathan L. (2017). "Building state and citizen: How tax collection in Congo engenders citizen engagement with the state," mimeo, Harvard University.
- Xiao, Erte and Daniel Houser (2005). "Emotion expression in human punishment behavior," *Proceedings of the National Academy of Sciences of the United States of America*, 102(20): 7398–7401.

Figures and tables

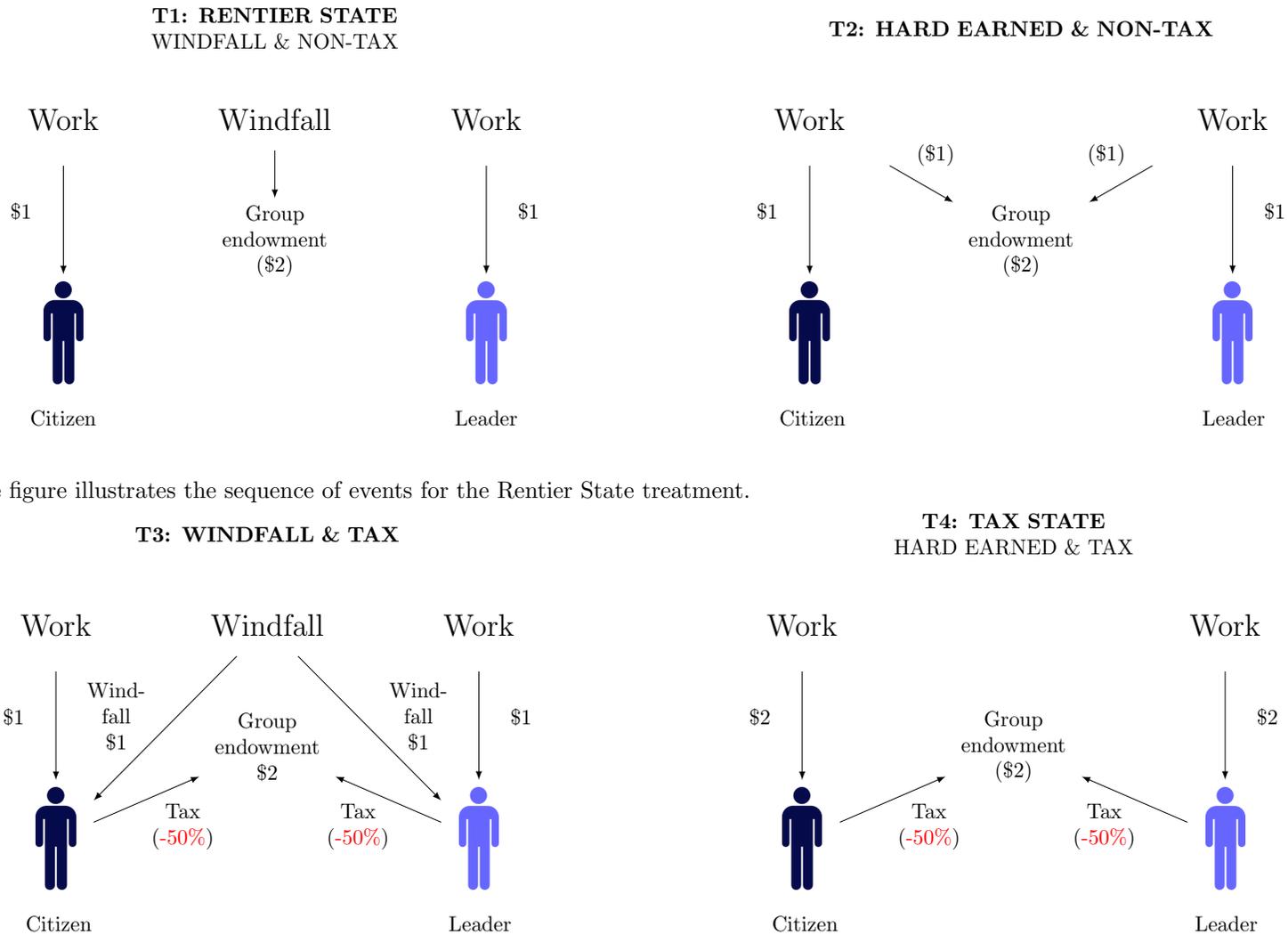
Figure 1: Effect of taxation on the willingness to punish



The figure shows the mean willingness to punish for citizens in the Rentier State (group endowment is windfall and non-tax) and Tax State (group endowment is produced by work and financed through taxation) treatments. The estimated standard errors are also indicated.

Figure 2: Sequence of events

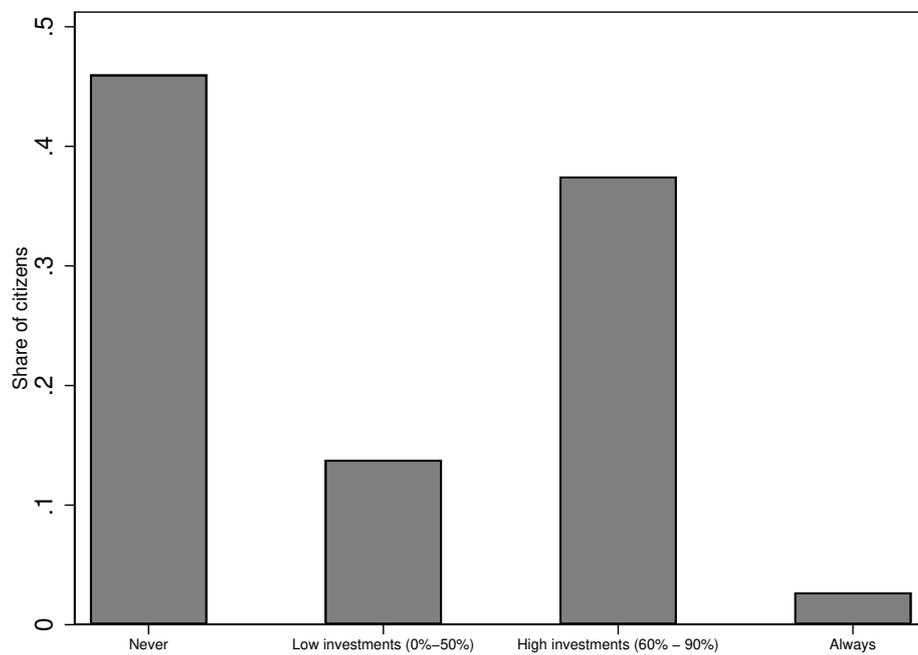
Figure 3: Treatments: Generation and collection of group endowment



The figure illustrates the sequence of events for the Rentier State treatment.

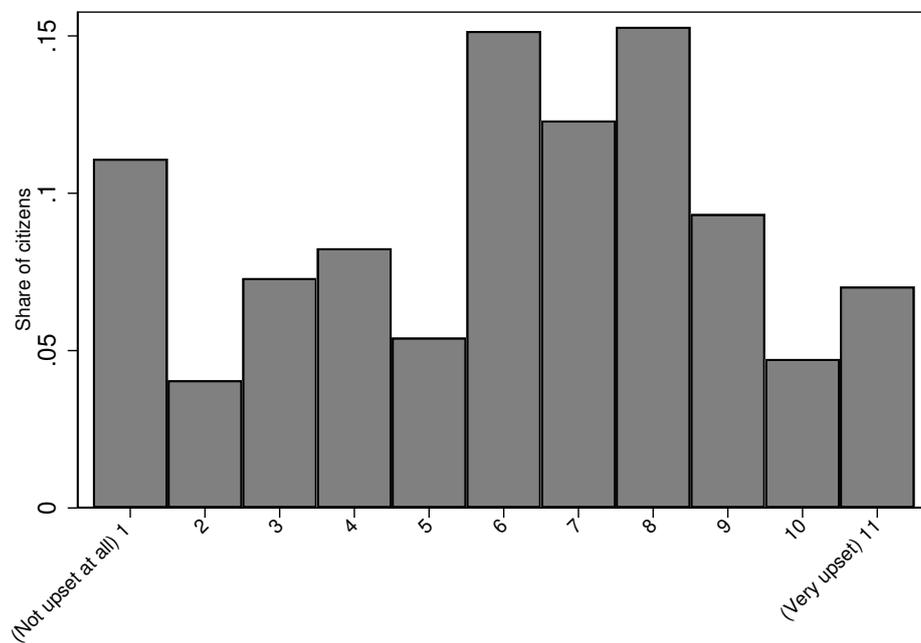
The figure illustrates the production and financing of the group endowment for the four treatments.

Figure 4: Types of punishment behavior



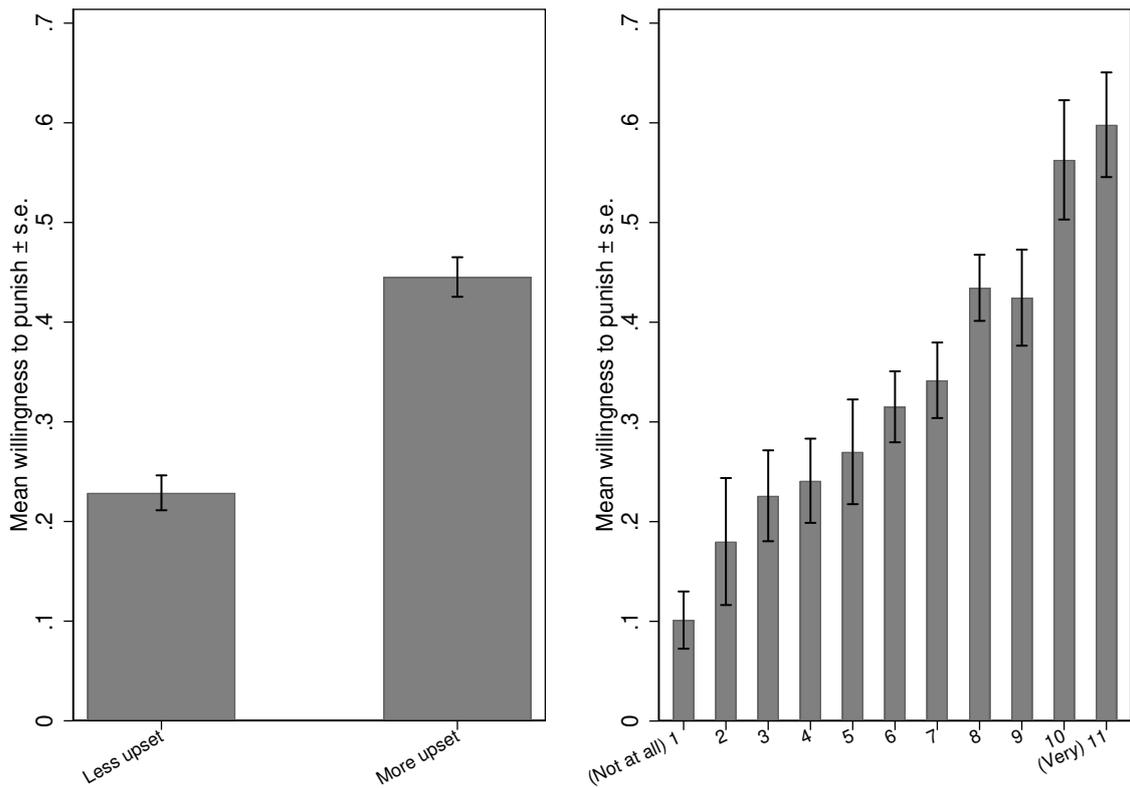
Note: The figure shows the share of citizens characterized by four different types of punishment behavior. “Never”: citizens that never punish the leader. “Low investments (0%-50%)”: citizens that punish an investment share of between 0% and 50% as the highest. “High investments (60%-90%)”: citizens that punish an investment share between 60% and 90% as the highest. “Always”: citizens that punish all investment decisions of the leader, even when he or she invests 100%.

Figure 5: How upset would you be if the leader invests less than the fair share?



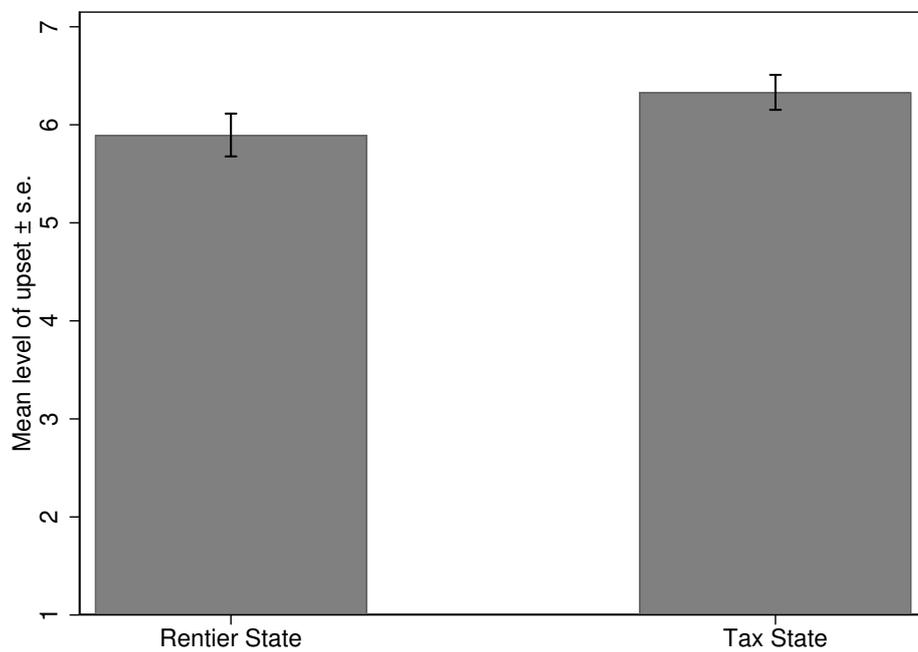
Note: The figure illustrates how citizens answered the question “How upset would you be if the leader invests less than the fair share, when both of you completed the assignment?” on a 11-point scale from “Not upset at all” (1) to “Very upset” (11). Each bar indicates the share of citizens that answered each of the numbers on the scale.

Figure 6: The willingness to punish and negative emotions



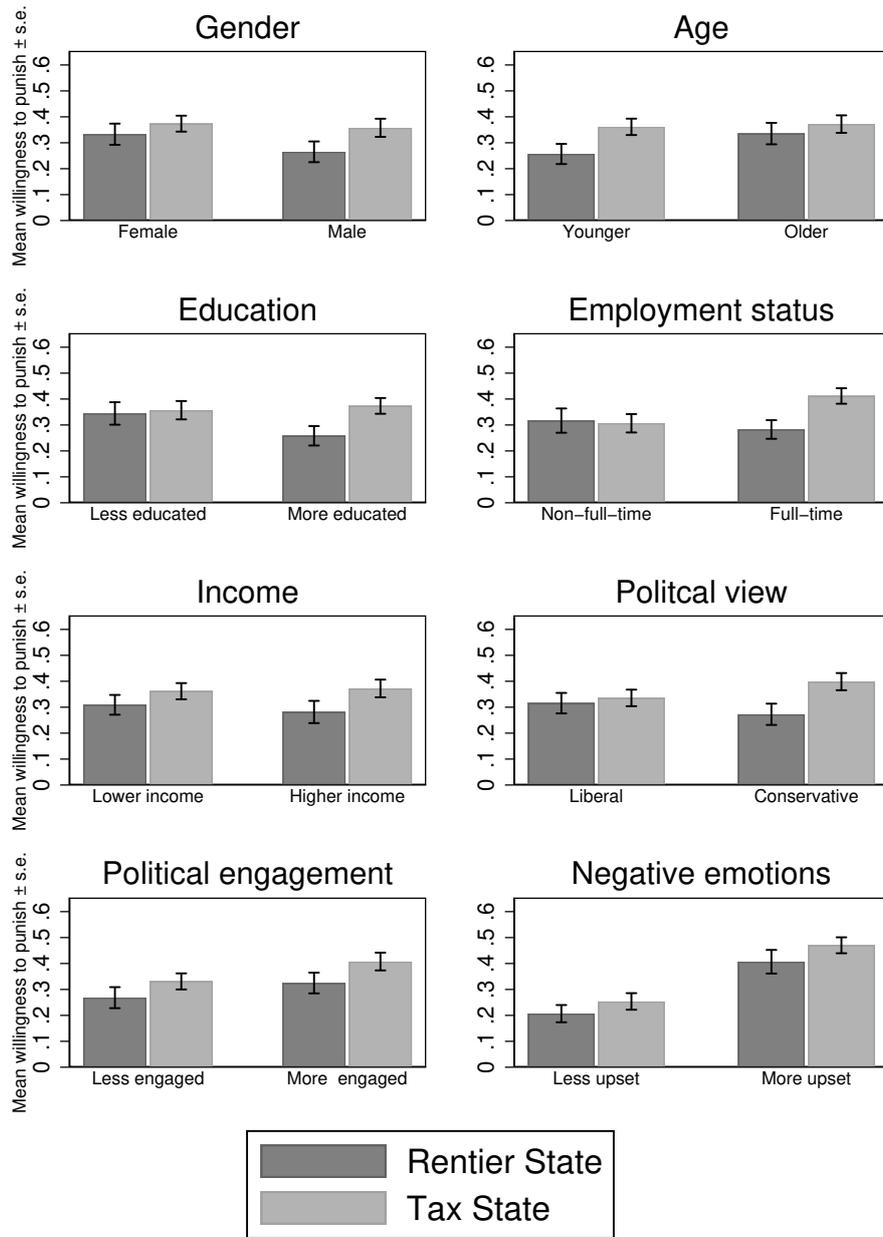
Note: **Left panel:** illustrates the mean willingness to punish and estimated standard error for citizens that would be less than median and citizens that would be more than median upset if the leader invests less than the fair share, respectively. **Right panel** illustrates the mean willingness to punish and estimated standard error for citizens according to their answer to the question “How upset would you be if the leader invests less than the fair share”. Each bar illustrates the mean willingness to punish for that level of negative emotions.

Figure 7: Effect of taxation on negative emotions



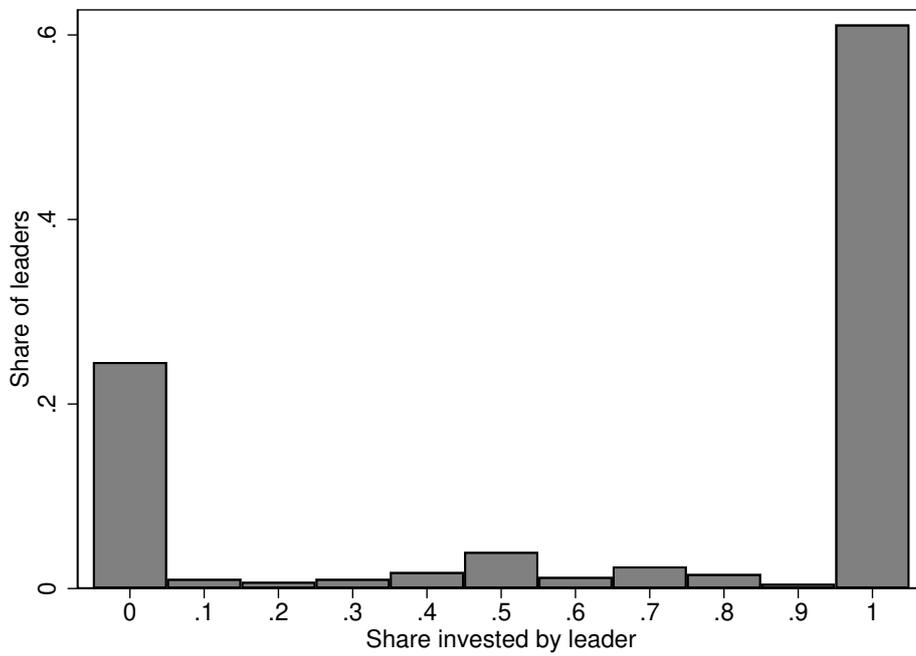
Note: The figure shows the mean values for how upset the citizen would be if the leader invest less than the fair share for the Rentier State and Tax State treatments, respectively. The variable is measured on a scale from 1 (Not upset at all) to 11 (Very upset). The figure also indicates the estimated standard errors.

Figure 8: Effect of taxation on the willingness to punish by subgroup



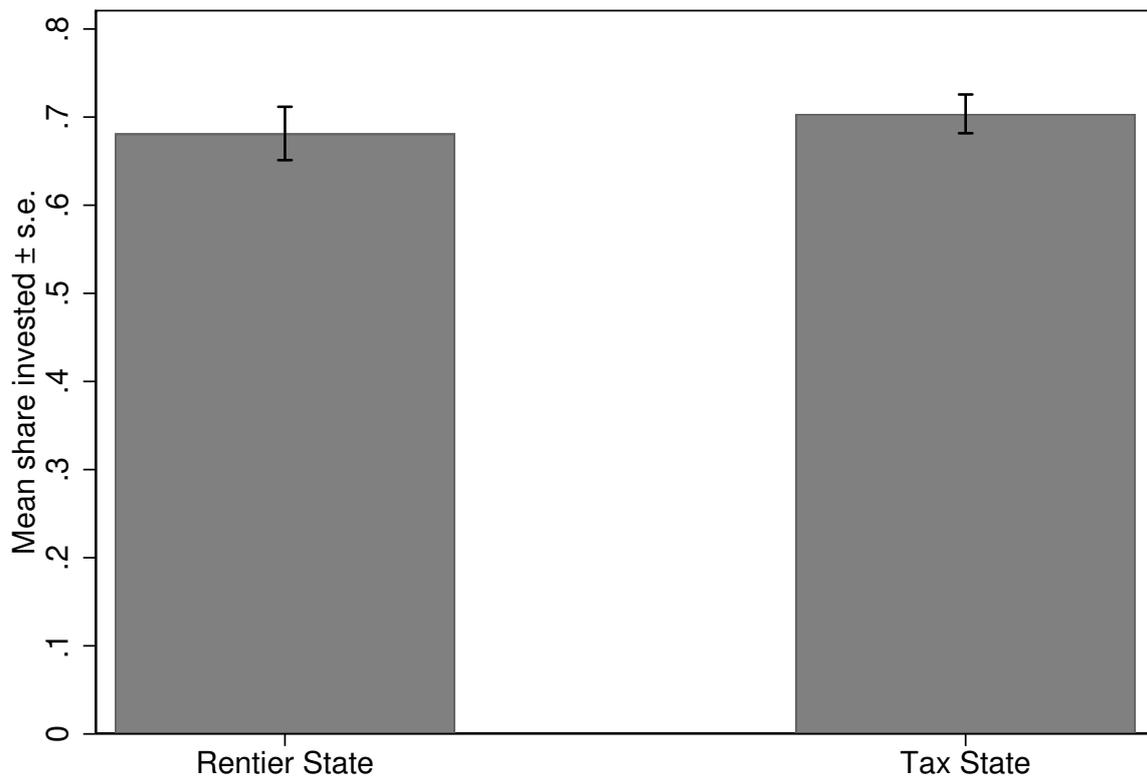
Note: The figure shows the mean willingness to punish and standard error for the Rentier State and Tax Treatments, by subgroups. “Male”: indicator variable taking the value of 1 for males, “Above median age” indicator variable taking the value of 1 if individual is older than the median age of the sample (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), “More politically engaged”: 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, and “More upset”: indicator variable taking the value of 1 for individuals who report that they would be higher than median upset if the leader invests less than the fair share.

Figure 9: Distribution of share invested by leaders



Note: The figure shows the distribution of share the leaders invest in the common pool for the total sample. Each bar illustrates the share of leaders that made given investment in the common pool.

Figure 10: Effect of taxation on share invested by leaders



Note: The figure shows the mean share invested by leader and estimated standard errors by Rentier State and Tax State treatments.

Table 1: 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	

mean coefficients; semean in parentheses

Note: The table provides mean values 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.

Table 2: Effect of negative emotions on the willingness to punish

	(1)	(2)	(3)	(4)	(5)	(6)
Above median upset	0.578*** (0.071)	0.572*** (0.072)	0.565*** (0.072)			
Upset				0.122*** (0.011)	0.121*** (0.011)	0.120*** (0.011)
Male		-0.061 (0.066)	-0.059 (0.066)		-0.047 (0.065)	-0.046 (0.065)
Above median age		0.031 (0.072)	0.028 (0.072)		0.056 (0.070)	0.052 (0.070)
Above median education		-0.034 (0.076)	-0.037 (0.077)		-0.018 (0.074)	-0.021 (0.075)
Employed full-time		0.114 (0.078)	0.116 (0.078)		0.111 (0.076)	0.113 (0.076)
Above median income		-0.015 (0.080)	-0.015 (0.080)		-0.007 (0.077)	-0.008 (0.077)
Conservative			0.011 (0.074)			0.014 (0.072)
More politically engaged			0.066 (0.073)			0.059 (0.071)
Constant	-0.282*** (0.047)	-0.303*** (0.087)	-0.334*** (0.099)	-0.743*** (0.071)	-0.793*** (0.105)	-0.820*** (0.114)
Observations	739	739	739	739	739	739
R^2	0.084	0.087	0.088	0.127	0.131	0.132

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

Note: The table reports regressions of the standardized willingness to punish on negative emotions and a set of explanatory variables. Columns (1)-(3) include an indicator variable for above median negative emotions, “Above median upset” and Columns (4)-(6) includes “Upset”, measured on a scale from 1 (not upset at all) to 11 (very upset). See Table 3 for definitions of background variables.

Table 3: Effect of taxation on the willingness to punish

	Willingness to punish			High punishment dummy		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned	0.034 (0.107)	0.031 (0.107)	0.040 (0.108)	-0.006 (0.102)	-0.006 (0.102)	-0.010 (0.102)
Tax	0.121 (0.116)	0.133 (0.116)	0.129 (0.117)	0.076 (0.112)	0.077 (0.112)	0.071 (0.113)
Hard Earned x Tax	0.033 (0.151)	0.019 (0.152)	0.018 (0.153)	0.126 (0.148)	0.118 (0.149)	0.131 (0.150)
Male		-0.113 (0.069)	-0.108 (0.069)		-0.046 (0.068)	-0.043 (0.069)
Above median age		0.001 (0.074)	-0.003 (0.074)		0.007 (0.075)	0.013 (0.075)
Above median education		-0.004 (0.080)	-0.013 (0.080)		0.077 (0.081)	0.070 (0.083)
Employed full-time		0.137* (0.081)	0.139* (0.081)		-0.023 (0.080)	-0.027 (0.079)
Above median income		0.006 (0.084)	0.007 (0.084)		-0.019 (0.085)	-0.014 (0.085)
Conservative			0.006 (0.077)			-0.065 (0.077)
More politically engaged			0.120 (0.076)			-0.013 (0.075)
Constant	-0.101 (0.076)	-0.125 (0.109)	-0.185 (0.122)	-0.088 (0.072)	-0.087 (0.108)	-0.052 (0.122)
Hard Earned + Tax + Hard Earned x Tax	0.188* (0.098)	0.182* (0.098)	0.186* (0.098)	0.195** (0.097)	0.189* (0.097)	0.192** (0.097)
Hard Earned + Hard Earned x Tax	0.066 (0.107)	0.050 (0.107)	0.057 (0.108)	0.120 (0.107)	0.112 (0.108)	0.121 (0.109)
Tax + Hard Earned x Tax	0.154 (0.097)	0.151 (0.097)	0.147 (0.098)	0.202** (0.097)	0.195** (0.097)	0.202** (0.097)
Observations	739	739	739	739	739	739
R^2	0.006	0.013	0.017	0.008	0.010	0.011

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

Note: The table reports regressions of the standardized value of willingness to punish on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4), “Tax” (indicator variable taking the value of 1 for individuals in the Tax treatments (T3 and T4)) and “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”) and a set of explanatory 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 always punish, or 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 variables 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. “Hard Earned + Tax + Hard Earned x Tax” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”. “Tax + Hard Earned x Tax” is the sum of the estimated parameters for “Tax” and “Hard Earned”. “Hard Earned + Hard Earned x Tax” is the sum of estimated parameters for “Hard Earned” and “Hard Earned x Tax”.

Table 4: Effect of taxation on negative emotions

	(1)	(2)
Hard Earned	-0.077 (0.109)	-0.082 (0.108)
Tax	0.150 (0.113)	0.148 (0.114)
Hard Earned x Tax	0.075 (0.151)	0.071 (0.152)
Male		-0.193*** (0.069)
Above median age		-0.162** (0.074)
Above median education		0.032 (0.080)
Employed full-time		0.070 (0.080)
Above median income		0.041 (0.083)
Conservative		-0.009 (0.077)
More politically engaged		0.168** (0.074)
Constant	-0.068 (0.075)	-0.043 (0.116)
Hard Earned + Tax + Hard Earned x Tax	0.149 (0.096)	0.137 (0.096)
Hard Earned + Hard Earned x Tax	-0.002 (0.105)	-0.011 (0.106)
Tax + Hard Earned x Tax	0.226** (0.100)	0.219** (0.100)
Observations	739	739
R^2	0.009	0.034

Robust standard errors in parentheses

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

Note: The table reports regressions of negative emotions (“How upset would you be if the leader invests less than the fair share?”), measured on a standardized 11-point scale, on the treatment variables “Hard Earned & Non-Tax” (indicator variable taking the value of one for individuals in the Hard Earned & Non-Tax treatment (T2), “Windfall & Tax” (indicator variable taking the value of one for individuals in the Windfall & Tax treatment (T3) and “Tax State” (indicator variable taking the value of one for citizens in the Tax State treatment) and a set of explanatory variables (see Table 3 for definitions).

Table 5: Effect of going from Rentier State to Tax State treatment on the willingness to punish for subgroups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male	Older	Higher educ	Full time	Income	Conserv.	Pol. eng.	Upset
Total, var = 0	0.127 (0.133)	0.256* (0.134)	0.024 (0.149)	-0.039 (0.157)	0.136 (0.133)	0.036 (0.135)	0.145 (0.138)	0.126 (0.125)
Total, var = 1	0.245* (0.136)	0.119 (0.144)	0.311** (0.129)	0.343*** (0.124)	0.243* (0.145)	0.352** (0.141)	0.233* (0.140)	0.174 (0.147)
Difference	0.117 (0.183)	-0.137 (0.197)	0.287 (0.197)	0.382* (0.200)	0.107 (0.197)	0.316 (0.195)	0.088 (0.197)	0.048 (0.194)
Background vars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Political vars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	739	739	739	739	739	739	739	739
R^2	0.018	0.021	0.028	0.022	0.019	0.024	0.021	0.092

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

Note: The table reports the estimated regression coefficients for the effect of going from the Rentier State to the Tax State treatment on the standardized value of willingness to punish for different subgroups of the sample. “Total, Var = 0” is the estimated effect on subgroups for which the respective variables in the column headers take the value of zero (the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax” in Table D.1 in Appendix D), “Total, Var = 1” is the estimated effect on subgroups for which the respective variables in the column headers take the value of one (the sum of the estimated parameters for “Hard Earned”, “Tax”, “Hard Earned x Tax”, “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var” in Table D.1 in Appendix D) and “Difference” is the difference in estimated effect of the Tax State treatment between the subgroup for which “Var” takes the value of zero and subgroups for which “Var” takes the value of one (the sum of the estimated parameters for “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var” in Table D.1 in Appendix D). “Upset”: indicator taking the value of 1 for individuals who state they will be more upset than the median of the sample if the leader invests less than the fair share. See table 3 for explanation of the rest of the variables.

Table 6: Effect of taxation on share invested in the common pool (leaders)

	Share invested			High investment dummy		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned	0.033 (0.100)	0.043 (0.099)	0.050 (0.099)	0.114 (0.101)	0.119 (0.100)	0.124 (0.100)
Tax	-0.094 (0.103)	-0.099 (0.105)	-0.104 (0.105)	-0.040 (0.105)	-0.053 (0.106)	-0.059 (0.106)
Hard Earned x Tax	0.113 (0.136)	0.114 (0.135)	0.117 (0.135)	0.034 (0.136)	0.040 (0.135)	0.045 (0.135)
Male		-0.028 (0.065)	-0.035 (0.064)		0.044 (0.064)	0.037 (0.063)
Above median age		0.227*** (0.065)	0.241*** (0.065)		0.275*** (0.065)	0.287*** (0.065)
Above median education		0.100 (0.068)	0.074 (0.068)		0.088 (0.067)	0.060 (0.067)
Employed full-time		-0.213*** (0.070)	-0.223*** (0.070)		-0.192*** (0.070)	-0.203*** (0.070)
Above median income		-0.061 (0.073)	-0.053 (0.072)		-0.071 (0.072)	-0.066 (0.071)
Conservative			-0.133** (0.067)			-0.110* (0.066)
More politically engaged			0.131** (0.066)			0.161** (0.066)
Constant	-0.009 (0.070)	-0.019 (0.094)	-0.012 (0.102)	-0.059 (0.073)	-0.123 (0.096)	-0.136 (0.104)
Hard Earned + Tax + Hard Earned x Tax	0.052 (0.087)	0.058 (0.086)	0.062 (0.086)	0.108 (0.089)	0.105 (0.088)	0.109 (0.088)
Hard Earned + Hard Earned x Tax	0.145 (0.092)	0.157* (0.092)	0.167* (0.092)	0.148 (0.091)	0.158* (0.091)	0.169* (0.090)
Tax + Hard Earned x Tax	0.019 (0.088)	0.015 (0.086)	0.013 (0.085)	-0.006 (0.086)	-0.014 (0.085)	-0.014 (0.084)
Observations	957	957	957	957	957	957
R^2	0.003	0.032	0.042	0.004	0.035	0.046

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

Note: The table reports regressions of the standardized value of share invested by the leader on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4)), “Tax” (indicator variable taking the value of one for individuals in the Tax treatments (T3 and T4)) and “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”) 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 citizen invests everything in the common pool. “Hard Earned + Tax + Hard Earned x Tax” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”. “Tax + Hard Earned x Tax” is the sum of the estimated parameters for “Tax” and “Hard Earned”. “Hard Earned + Hard Earned x Tax” is the sum of estimated parameters for “Hard Earned” and “Hard Earned x Tax”.

Appendix A 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 be 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 Earned mechanism is manipulated, a robustness treatment, "Tax State Extra Hard Earned" is implemented. The treatment is identical to the Renter State treatment in all respects apart from the real effort task being ten minutes long instead 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.

A.1 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:

$$y_i = \alpha + \beta^{T4b} T4b_i + \beta^X X_i + \beta^Z Z_i + \varepsilon_i, \quad (8)$$

where $T4b_i$ is 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.

A.2 Results

Table A.1: Robustness check for Hard Earned mechanism

	(1)	(2)	(3)
Extra Hard Earned	0.038 (0.123)	0.038 (0.123)	0.068 (0.122)
Male		-0.081 (0.132)	-0.106 (0.131)
Above median age		0.076 (0.125)	0.050 (0.124)
Above median education		0.007 (0.130)	0.023 (0.130)
Employed full-time		0.354** (0.141)	0.353** (0.142)
Above median income		-0.105 (0.142)	-0.100 (0.142)
Conservative			0.263** (0.127)
More politically engaged			0.264** (0.127)
Constant	0.067 (0.088)	-0.092 (0.154)	-0.346** (0.174)
Observations	284	284	284
R^2	0.000	0.024	0.047

Robust standard errors in parentheses

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

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 1.

Appendix B Balance tables

Table B.1: Effect of treatments on background variables, citizens

	Male	Older	Higher educ	Full time	Higher income	Conservative	Pol. engaged
Hard Earned & Non-Tax	-0.053 (0.058)	-0.047 (0.056)	-0.027 (0.056)	-0.025 (0.055)	0.040 (0.056)	-0.043 (0.055)	-0.073 (0.056)
Windfall & Tax	0.047 (0.068)	-0.048 (0.058)	0.001 (0.058)	-0.042 (0.058)	-0.004 (0.058)	-0.105* (0.057)	0.033 (0.058)
Tax State	-0.087* (0.051)	-0.014 (0.049)	0.016 (0.049)	-0.031 (0.048)	-0.009 (0.049)	0.042 (0.049)	-0.032 (0.049)
Constant	0.528*** (0.042)	0.503*** (0.039)	0.558*** (0.039)	0.601*** (0.038)	0.466*** (0.039)	0.448*** (0.039)	0.497*** (0.039)
P-value of F-test	0.088*	0.718	0.694	0.961	0.602	0.012**	0.193

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). The Rentier State treatment is the reference category. The reported p-values tests the hypothesis that all the treatments have the same effect on the background indicator variables.

Table B.2: Effect of treatments on background variables, citizens, worked and not worked

	Male	Older	Higher educ	Full time	Higher income	Conservative	Pol. engaged
Hard Earned & Non-Tax	-0.040 (0.051)	-0.025 (0.050)	-0.058 (0.050)	-0.022 (0.049)	0.054 (0.050)	-0.027 (0.050)	-0.091* (0.050)
Windfall & Tax	0.067 (0.056)	-0.071 (0.050)	-0.023 (0.050)	0.005 (0.050)	-0.001 (0.050)	-0.086* (0.050)	-0.045 (0.050)
Tax State	-0.089** (0.045)	-0.026 (0.043)	-0.017 (0.043)	-0.027 (0.043)	0.003 (0.043)	0.047 (0.043)	-0.051 (0.043)
Constant	0.515*** (0.037)	0.500*** (0.035)	0.574*** (0.035)	0.598*** (0.034)	0.456*** (0.035)	0.456*** (0.035)	0.505*** (0.035)
P-value of F-test	0.007***	0.542	0.632	0.758	0.445	0.008***	0.583

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, on the total sample of citizens, i.e.. including citizens that chose not to work or who are identified as inconsistent switchers. 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.

Table B.3: Preferences by treatments, citizens

	Windfall & Non-Tax	Hard Earned & Non-Tax	Windfall & Tax	Hard Earned & Tax	Total	P-value, F-test
Risk averse	0.52 (0.04)	0.47 (0.04)	0.53 (0.04)	0.55 (0.03)	0.52 (0.02)	0.35
More altruistic	0.38 (0.04)	0.34 (0.04)	0.33 (0.04)	0.28 (0.03)	0.32 (0.02)	0.37
High positive reciprocity	0.48 (0.04)	0.49 (0.04)	0.57 (0.04)	0.47 (0.03)	0.50 (0.02)	0.14
High negative reciprocity	0.44 (0.04)	0.46 (0.04)	0.47 (0.04)	0.55 (0.03)	0.49 (0.02)	0.11
Loss averse	0.61 (0.04)	0.50 (0.04)	0.59 (0.04)	0.51 (0.03)	0.54 (0.02)	0.23
Above median upset	0.45 (0.04)	0.43 (0.04)	0.53 (0.04)	0.52 (0.03)	0.49 (0.02)	0.13
Observations	163	158	134	284	739	

mean coefficients; semean in parentheses

Note: The table provides mean values for preferences 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. “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. (2015), 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 = $0.2631 \times$ Willingness to punish if oneself is treated unfairly + $0.2631 \times$ Willingness punish if other is treated unfairly + $0.3738 \times$ Willingness to take revenge. These weights were obtained by Falk et al. (2015) 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). The reported p-values test the hypothesis that all the treatments have the same effect on the preference indicator variables.

Table B.4: Effect of treatments on preferences, citizens

	Risk averse	More altruistic	High pos. reci.	High neg. reci	Loss averse	More upset
Hard Earned & Non-Tax	-0.047 (0.056)	-0.045 (0.054)	0.009 (0.056)	0.020 (0.056)	-0.107* (0.055)	-0.017 (0.056)
Windfall & Tax	0.008 (0.058)	-0.052 (0.056)	0.090 (0.058)	0.035 (0.058)	-0.018 (0.057)	0.082 (0.058)
Tax State	0.024 (0.049)	-0.102** (0.047)	-0.013 (0.049)	0.114** (0.049)	-0.097** (0.049)	0.073 (0.049)
Constant	0.521*** (0.039)	0.380*** (0.038)	0.485*** (0.039)	0.436*** (0.039)	0.607*** (0.038)	0.448*** (0.039)
P-value of F-test	0.353	0.369	0.139	0.110	0.233	0.131

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 preference indicator variables as the dependent variable on indicator variables for the treatments. The Rentier State treatment is the reference category. See Table B.3 for definition of the variables. The reported p-value tests the hypothesis that all the treatments have the same effect on the preference indicator variables.

Table B.5: Effect of treatments on decision to work

	(1)	(2)
Tax State	0.001 (0.019)	0.003 (0.019)
Hard Earned & Non-Tax	0.029 (0.018)	0.032* (0.018)
Windfall & Tax	-0.062** (0.027)	-0.059** (0.027)
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*** (0.015)	0.932*** (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.6: Effect of treatments on inconsistent punishment behavior

	(1)	(2)
Tax State	0.079** (0.034)	0.073** (0.035)
Hard Earned & Non-Tax	0.026 (0.038)	0.024 (0.039)
Windfall & Tax	0.043 (0.041)	0.047 (0.041)
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*** (0.026)	0.170*** (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 (dependent variable is indicator variable taking the value of one for citizens who switch the wrong way) 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 & No-Tax” = “Windfall & Tax”. The background variables are as defined in Table 3.

Appendix C Deviations from Pre-Analysis Plan

C.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 change the results. The results for the original definition of the dependent variable is reported in Table D.4.

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.

In addition to socioeconomic background characteristics and negative emotions, the pre-analysis plan specified risk preferences, altruism, positive reciprocity, negative reciprocity and loss aversion as dimensions for the heterogeneity analysis. Table B.4 in Appendix B tests whether the citizens’ preferences and negative emotions are affected by treatments, by regressing the respective preference and emotion measures on indicator variables for the Hard Earned & Non-Tax, the Windfall & Non-Tax and the Tax State treatment (Rentier State is the reference category). It shows that the Tax State treatment significantly reduces the citizens’ self-reported altruism and loss aversion, and significantly increases negative reciprocity and there is therefore no heterogeneity analysis for these dimensions in the paper. There are no good theoretical reasons to expect heterogeneity in effects of taxation according to positive reciprocity and altruism and the analysis is not presented in the main paper, but can be found in Table D.7.

C.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.

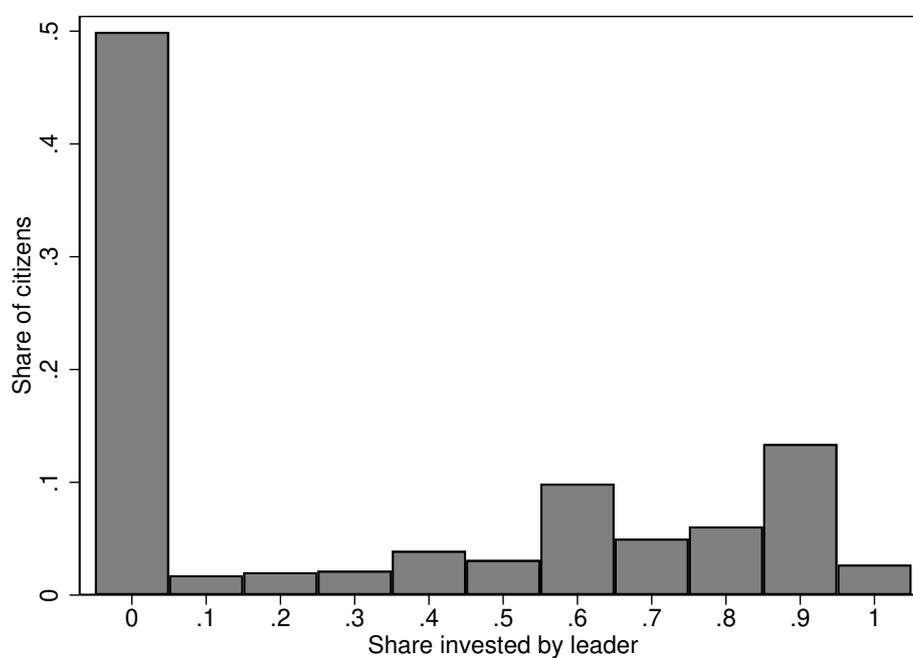
- Regressions with binary dependent variable reported in Columns (4)-(6) in Table 3, and in Columns (1)-(3) in Table D.3

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 D Additional figures and tables

Figure D.1: Distribution of the willingness to punish



Note: The figure illustrates the share of citizens that for each of the 11 possible investment levels punish that level as the highest.

Table D.1: Heterogeneity analysis

	(1) Male	(2) Older	(3) Higher educ	(4) Full-time	(5) Income	(6) Conservative	(7) Pol. engaged
Hard Earned	0.017 (0.150)	0.195 (0.148)	-0.045 (0.160)	-0.036 (0.173)	0.007 (0.149)	0.078 (0.146)	0.135 (0.150)
Tax	0.052 (0.158)	0.289* (0.158)	-0.243 (0.169)	-0.025 (0.180)	-0.004 (0.160)	0.105 (0.152)	0.078 (0.168)
Var	-0.190 (0.135)	-0.005 (0.075)	-0.273* (0.156)	-0.081 (0.161)	-0.101 (0.158)	-0.099 (0.153)	0.117 (0.154)
HE x Tax	0.058 (0.208)	-0.228 (0.208)	0.312 (0.223)	0.022 (0.237)	0.133 (0.211)	-0.148 (0.201)	-0.068 (0.214)
HE x Var	0.038 (0.203)	-0.323 (0.216)	0.144 (0.216)	0.122 (0.221)	0.072 (0.217)	-0.105 (0.216)	-0.227 (0.217)
Tax x Var	0.140 (0.187)	-0.333 (0.233)	0.668*** (0.231)	0.260 (0.237)	0.287 (0.233)	0.035 (0.237)	0.099 (0.234)
HE x Tax x Var	-0.061 (0.272)	0.519* (0.304)	-0.526* (0.304)	0.000 (0.310)	-0.253 (0.306)	0.386 (0.307)	0.217 (0.307)
Constant	-0.139 (0.140)	-0.278** (0.138)	-0.029 (0.147)	-0.058 (0.147)	-0.131 (0.139)	-0.138 (0.137)	-0.189 (0.139)
Total, var = 0	0.127 (0.133)	0.256* (0.134)	0.024 (0.149)	-0.039 (0.157)	0.136 (0.133)	0.036 (0.135)	0.145 (0.138)
Total, var = 1	0.245* (0.136)	0.119 (0.144)	0.311** (0.129)	0.343*** (0.124)	0.243* (0.145)	0.352** (0.141)	0.233* (0.140)
Difference	0.117 (0.183)	-0.137 (0.197)	0.287 (0.197)	0.382* (0.200)	0.107 (0.197)	0.316 (0.195)	0.088 (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

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

Note: The table reports regressions of standardized value of the willingness to punish on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4), “Tax” (indicator variable taking the value of one for individuals in the Tax treatments (T3 and T4)), “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”), “Hard Earned x Var” (interaction term between “Hard Earned” and “Var”), “Tax x Var” (interaction term between “Tax” and “Var” and “Hard Earned x Tax x Var” (interaction term between “Hard Earned”, “Tax” and “Var”) and a set of explanatory variables (defined in Table 3). “Total, Var = 0” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”, “Total, Var = 1” is the sum of the estimated parameters for “Hard Earned”, “Tax”, “Hard Earned x Tax”, “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”, “Difference” is the sum of the estimated parameters for “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”.

Table D.2: Heterogeneity, preferences and emotion

	(1)
Hard Earned	0.043 (0.129)
Tax	0.025 (0.145)
Above median upset	0.516*** (0.152)
HE x Tax	0.058 (0.192)
HE x Above median upset	0.012 (0.213)
Tax x Above median upset	0.115 (0.227)
HE x Tax x Above median upset	-0.079 (0.296)
Constant	-0.392*** (0.131)
Total, Var = 0	0.126 (0.125)
Total, Var = 1	0.174 (0.147)
Difference	0.048 (0.194)
Background vars	Yes
Political vars	Yes
Observations	739
R^2	0.092

Robust standard errors in parentheses

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

Note: The table reports regressions of standardized value of the willingness to punish on the treatment variables (see Table D.1 for definitions of these and the interaction terms) and a set of explanatory variables (see Table 3 for definitions). “Upset”: indicator taking the value of 1 for individuals who state they will be more upset than the median of the sample if the leader invests less than the fair share. “Total, Var = 0” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”, “Total, Var = 1” is the sum of the estimated parameters for “Hard Earned”, “Tax”, “Hard Earned x Tax”, “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”, “Difference” is the sum of the estimated parameters for “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”.

Table D.3: Treatment effects on the willingness to punish: dummy for positive punishment

	Only consistent punishers			Total sample of citizens who worked		
	(1)	(2)	(3)	(4)	(5)	(6)
Hard Earned	0.070 (0.112)	0.058 (0.112)	0.072 (0.113)	0.086 (0.103)	0.071 (0.103)	0.085 (0.103)
Tax	0.086 (0.117)	0.098 (0.117)	0.098 (0.116)	0.102 (0.106)	0.106 (0.106)	0.117 (0.106)
Hard Earned x Tax	-0.076 (0.153)	-0.086 (0.153)	-0.097 (0.154)	-0.044 (0.138)	-0.053 (0.138)	-0.075 (0.139)
Male		-0.175** (0.069)	-0.172** (0.069)		-0.188*** (0.063)	-0.189*** (0.063)
Above median age		-0.092 (0.075)	-0.102 (0.075)		-0.083 (0.066)	-0.095 (0.067)
Above median education		-0.038 (0.079)	-0.043 (0.079)		-0.061 (0.070)	-0.062 (0.070)
Employed full-time		0.176** (0.082)	0.182** (0.082)		0.166** (0.072)	0.172** (0.072)
Above median income		0.048 (0.084)	0.045 (0.083)		0.041 (0.072)	0.038 (0.072)
Conservative			0.056 (0.077)			0.084 (0.068)
More politically engaged			0.159** (0.075)			0.122* (0.067)
Constant	-0.062 (0.079)	-0.030 (0.110)	-0.130 (0.122)	-0.094 (0.073)	-0.036 (0.099)	-0.132 (0.111)
Hard Earned + Tax + Hard Earned x Tax	0.080 (0.099)	0.070 (0.098)	0.073 (0.098)	0.144 (0.089)	0.124 (0.089)	0.127 (0.089)
Hard Earned + Hard Earned x Tax	-0.006 (0.105)	-0.027 (0.105)	-0.024 (0.105)	0.041 (0.092)	0.018 (0.093)	0.010 (0.093)
Tax + Hard Earned x Tax	0.010 (0.099)	0.012 (0.099)	0.001 (0.100)	0.058 (0.088)	0.053 (0.088)	0.042 (0.089)
Observations	739	739	739	929	929	929
R^2	0.001	0.018	0.024	0.003	0.019	0.023

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

Note: The table reports regressions of a dummy taking the value of one for citizens that punish some investment and zero for citizens that never punish on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4)), “Tax” (indicator variable taking the value of one for individuals in the Tax treatments (T3 and T4)) and “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”) and a set of explanatory variables (see Table 3 for definitions). “Hard Earned + Tax + Hard Earned x Tax” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”. “Tax + Hard Earned x Tax” is the sum of the estimated parameters for “Tax” and “Hard Earned”. “Hard Earned + Hard Earned x Tax” is the sum of estimated parameters for “Hard Earned” and “Hard Earned x Tax”. Column (1) - (3) shows the regressions on the sample without the inconsistent punishers, and Column (4) - (6) shows the regressions on the sample including the inconsistent punishers.

Table D.4: Treatment effects on lowest investment share not punished

	(1)	(2)	(3)
Hard Earned	0.048 (0.107)	0.044 (0.108)	0.054 (0.108)
Tax	0.128 (0.116)	0.139 (0.116)	0.136 (0.117)
Hard Earned x Tax	0.002 (0.152)	-0.012 (0.152)	-0.014 (0.153)
Male		-0.124* (0.069)	-0.119* (0.069)
Above median age		-0.010 (0.074)	-0.016 (0.074)
Above median education		-0.007 (0.080)	-0.015 (0.080)
Employed full-time		0.143* (0.081)	0.146* (0.081)
Above median income		0.006 (0.084)	0.006 (0.084)
Conservative			0.011 (0.077)
More politically engaged			0.130* (0.075)
Constant	-0.102 (0.076)	-0.116 (0.108)	-0.183 (0.122)
Hard Earned + Tax + Hard Earned x Tax	0.178* (0.097)	0.172* (0.098)	0.176* (0.097)
Hard Earned + Hard Earned x Tax	0.050 (0.107)	0.032 (0.108)	0.040 (0.108)
Tax + Hard Earned x Tax	0.130 (0.098)	0.128 (0.098)	0.122 (0.098)
Observations	739	739	739
R^2	0.005	0.013	0.017

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

Note: The table reports regressions of standardized value of willingness to punish, measured as the lowest share invested for which the citizen does not punish, on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4), “Tax” (indicator variable taking the value of one for individuals in the Tax treatments (T3 and T4)) and “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”) and a set of explanatory variables (see Table 3 for definitions). “Hard Earned + Tax + Hard Earned x Tax” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”. “Tax + Hard Earned x Tax” is the sum of the estimated parameters for “Tax” and “Hard Earned”. “Hard Earned + Hard Earned x Tax” is the sum of estimated parameters for “Hard Earned” and “Hard Earned x Tax”.

Table D.5: Heterogeneity regressions without controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Male	Older	Higher educ	Full time	Income	Conservative	Pol. engaged
Hard Earned	-0.018 (0.148)	0.199 (0.147)	-0.047 (0.159)	-0.025 (0.172)	-0.014 (0.147)	0.078 (0.144)	0.158 (0.148)
Tax	0.044 (0.157)	0.298* (0.157)	-0.241 (0.168)	-0.027 (0.180)	-0.002 (0.158)	0.110 (0.152)	0.097 (0.168)
Var	-0.200 (0.133)	0.210 (0.151)	-0.230 (0.153)	-0.092 (0.157)	-0.074 (0.153)	-0.115 (0.152)	0.151 (0.152)
HE x Tax	0.075 (0.206)	-0.218 (0.206)	0.321 (0.222)	0.025 (0.235)	0.157 (0.208)	-0.134 (0.200)	-0.088 (0.213)
HE x Var	0.087 (0.201)	-0.342 (0.213)	0.139 (0.215)	0.098 (0.220)	0.100 (0.215)	-0.121 (0.214)	-0.268 (0.215)
Tax x Var	0.151 (0.188)	-0.367 (0.232)	0.648*** (0.229)	0.259 (0.236)	0.267 (0.232)	-0.002 (0.237)	0.036 (0.232)
HE x Tax x Var	-0.081 (0.271)	0.527* (0.303)	-0.513* (0.302)	0.016 (0.307)	-0.265 (0.303)	0.406 (0.307)	0.286 (0.303)
Constant	0.004 (0.105)	-0.207** (0.104)	0.027 (0.116)	-0.046 (0.125)	-0.067 (0.102)	-0.050 (0.105)	-0.176 (0.108)
HE + HE x Var	0.068 (0.145)	-0.143 (0.154)	0.092 (0.144)	0.073 (0.137)	0.086 (0.157)	-0.044 (0.159)	-0.110 (0.156)
Tax + Tax x Var	0.195 (0.142)	-0.068 (0.171)	0.407*** (0.155)	0.232 (0.152)	0.265 (0.170)	0.108 (0.183)	0.133 (0.160)
Total, var = 0	0.101 (0.133)	0.279** (0.134)	0.034 (0.149)	-0.028 (0.157)	0.140 (0.132)	0.054 (0.136)	0.167 (0.136)
Total, var = 1	0.258* (0.135)	0.098 (0.143)	0.308** (0.129)	0.345*** (0.125)	0.242* (0.146)	0.337** (0.141)	0.221 (0.141)
Difference	0.157 (0.182)	-0.182 (0.195)	0.275 (0.197)	0.373* (0.200)	0.102 (0.197)	0.282 (0.196)	0.054 (0.196)
Observations	739	739	739	739	739	739	739
R ²	0.010	0.011	0.018	0.015	0.008	0.013	0.014

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

Note: The table reports regressions of standardized value of the willingness to punish on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4), “Tax” (indicator variable taking the value of one for individuals in the Tax treatments (T3 and T4)), “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”), “Hard Earned x Var” (interaction term between “Hard Earned” and “Var”), “Tax x Var” (interaction term between “Tax” and “Var” and “Hard Earned x Tax x Var” (interaction term between “Hard Earned”, “Tax” and “Var”). “HE + HE x Var” is the sum of the estimated parameters for “Hard Earned” and “Hard Earned x Var”, “Tax + Tax x Var” is the sum of the estimated parameters for “tax” and “Tax x Var”, “Total, Var = 0” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”, “Total, Var = 1” is the sum of the estimated parameters for “Hard Earned”, “Tax”, “Hard Earned x Tax”, “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”, “Difference” is the sum of the estimated parameters for “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”.

Table D.6: Heterogeneity regressions for negative emotions, without controls

	(1)
Hard Earned	0.047 (0.128)
Tax	0.016 (0.141)
Above median upset	0.535*** (0.151)
HE x Tax	0.062 (0.189)
HE x Above median upset	-0.011 (0.211)
Tax x Above median upset	0.116 (0.223)
HE x Tax x Above median upset	-0.061 (0.293)
Constant	-0.341*** (0.089)
HE + HE x Above median upset	0.037 0.168
Tax + Tax x Above median upset	0.132 0.172
Total, var = 0	0.126 0.123
Total, var = 1	0.169 0.147
Difference	0.044 0.192
Observations	739
R^2	0.088
Robust standard errors in parentheses	
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$	

Note: The table reports regressions of standardized value of the willingness to punish on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4)), “Tax” (indicator variable taking the value of one for individuals in the Tax treatments (T3 and T4)), “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”), “Hard Earned x Var” (interaction term between “Hard Earned” and “Var”), “Tax x Var” (interaction term between “Tax” and “Var”) and “Hard Earned x Tax x Var” (interaction term between “Hard Earned”, “Tax” and “Var”). “HE + HE x Var” is the sum of the estimated parameters for “Hard Earned” and “Hard Earned x Var”, “Tax + Tax x Var” is the sum of the estimated parameters for “tax” and “Tax x Var”, “Total, Var = 0” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”, “Total, Var = 1” is the sum of the estimated parameters for “Hard Earned”, “Tax”, “Hard Earned x Tax”, “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”, “Difference” is the sum of the estimated parameters for “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”.

Table D.7: Heterogeneity, preferences

	(1) Risk averse	(2) Positive reciprocity
Hard Earned	-0.037 (0.160)	0.255* (0.150)
Tax	0.207 (0.176)	0.309* (0.163)
Var	-0.135 (0.153)	0.237 (0.155)
HE x Tax	-0.003 (0.226)	-0.297 (0.215)
HE x Var	0.149 (0.216)	-0.442** (0.217)
Tax x Var	-0.143 (0.233)	-0.349 (0.232)
HE x Tax x Var	0.032 (0.305)	0.625** (0.304)
Constant	-0.110 (0.152)	-0.297** (0.142)
HE + HE x Var	0.111 (0.146)	-0.187 (0.157)
Tax + Tax x Var	0.064 (0.154)	-0.040 (0.166)
Total, Var = 0	0.167 (0.147)	0.267** (0.133)
Total, Var = 1	0.206 (0.131)	0.102 (0.144)
Difference	0.038 (0.197)	-0.165 (0.197)
Background vars	Yes	Yes
Political vars	Yes	Yes
Observations	739	739
R^2	0.022	0.023

Robust standard errors in parentheses

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

Note: The table reports regressions of the standardized value of the willingness to punish on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4), “Tax” (indicator variable taking the value of one for individuals in the Tax treatments (T3 and T4)), “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”), “Hard Earned x Var” (interaction term between “Hard Earned” and “Var”), “Tax x Var” (interaction term between “Tax” and “Var”) and “Hard Earned x Tax x Var” (interaction term between “Hard Earned”, “Tax” and “Var”) and a set of explanatory variables. “Risk averse”: indicator variable taking the value of 1 for individuals that are more than median risk averse, “Positive reciprocity” indicator variable taking the value of 1 if individual is more than median risk averse, “Upset”: indicator taking the value of 1 for individuals who state they will be more upset than the median of the sample if the leader invests less than the fair share. “Total, Var = 0” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”, “Total, Var = 1” is the sum of the estimated parameters for “Hard Earned”, “Tax”, “Hard Earned x Tax”, “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”, “Difference” is the sum of the estimated parameters for “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”.

Table D.8: Heterogeneity regressions with dummy as dependent variable with controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Male	Older	Higher educ	Full time	Income	Conservative	Pol. engaged
Hard Earned	-0.019 (0.145)	0.072 (0.141)	0.048 (0.156)	-0.091 (0.171)	0.002 (0.144)	-0.042 (0.143)	0.047 (0.146)
Tax	0.110 (0.155)	0.178 (0.156)	-0.025 (0.157)	-0.135 (0.174)	0.142 (0.157)	0.046 (0.151)	0.112 (0.166)
Male	-0.049 (0.125)	0.097 (0.143)	0.019 (0.150)	-0.187 (0.156)	0.006 (0.150)	-0.151 (0.143)	-0.047 (0.142)
HE x Tax	0.074 (0.206)	-0.049 (0.209)	0.106 (0.215)	0.318 (0.233)	0.032 (0.210)	0.110 (0.203)	-0.061 (0.213)
HE x Var	0.019 (0.191)	-0.170 (0.204)	-0.112 (0.205)	0.136 (0.213)	-0.025 (0.207)	0.071 (0.201)	-0.139 (0.202)
Tax x Var	-0.068 (0.169)	-0.226 (0.223)	0.173 (0.223)	0.356 (0.229)	-0.155 (0.224)	0.044 (0.221)	-0.075 (0.223)
HE x Tax x Var	0.111 (0.258)	0.382 (0.296)	0.050 (0.297)	-0.325 (0.304)	0.215 (0.301)	0.050 (0.294)	0.415 (0.298)
Constant	-0.045 (0.138)	-0.095 (0.135)	-0.027 (0.140)	0.033 (0.150)	-0.062 (0.138)	-0.013 (0.137)	-0.040 (0.136)
Total, var = 0	0.164 (0.136)	0.201 (0.130)	0.129 (0.144)	0.092 (0.163)	0.176 (0.133)	0.115 (0.139)	0.097 (0.132)
Total, var = 1	0.226* (0.132)	0.186 (0.144)	0.240* (0.132)	0.259** (0.120)	0.211 (0.144)	0.280** (0.133)	0.298** (0.142)
Difference	0.062 (0.183)	-0.015 (0.193)	0.111 (0.195)	0.167 (0.202)	0.035 (0.196)	0.165 (0.191)	0.201 (0.193)
Background vars	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	739.	739	739	739	739	739	739
R ²	0.012	0.013	0.014	0.014	0.012	0.012	0.016

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

Note: The table reports regressions of standardized value of the high punishment indicator variable (takes the value of one for individuals that punish shares invested of 0.9 or 1) on the treatment variables “Hard Earned” (indicator variable taking the value of one for individuals in the Hard Earned treatments (T2 and T4), “Tax” (indicator variable taking the value of one for individuals in the Tax treatments (T3 and T4)), “Hard Earned x Tax” (interaction term between “Hard Earned” and “Tax”), “Hard Earned x Var” (interaction term between “Hard Earned” and “Var”), “Tax x Var” (interaction term between “Tax” and “Var” and “Hard Earned x Tax x Var” (interaction term between “Hard Earned”, “Tax” and “Var”) and a set of explanatory variables (see Table 3 for definitions). “Total, Var = 0” is the sum of the estimated parameters for “Hard Earned”, “Tax” and “Hard Earned x Tax”, “Total, Var = 1” is the sum of the estimated parameters for “Hard Earned”, “Tax”, “Hard Earned x Tax”, “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”, “Difference” is the sum of the estimated parameters for “Hard Earned x Var”, “Tax x Var” and “Hard Earned x Tax x Var”.

Appendix E Testing the theoretical framework

E.1 Empirical specification

To further investigate the Possession mechanism, the effect of citizens having had the money in their possession is scrutinized in the loss and in the gain domain for leader investments separately. Two dependent variables are defined. First, $y_{i,\text{loss}}$ is the highest level of investments in the loss domain (a share of 0 - 0.6 of group endowment is invested) that the citizen punishes. This variable takes the value of 0.6 for both individuals that punish investments of 0.6 and higher and individuals that punish 0.6, but not higher. Second, $y_{i,\text{gain}}$ is the lowest investment in the gain domain (0.7 - 1 of group endowment invested) for which the citizen does not punish higher investments. This variable takes the value of 0.7 for both individuals that do not punish investments of 0.7 and individuals that punish investments of 0.7, but do not punish higher investments.

Different versions of the following equation are estimated:

$$y_i = \alpha + \beta^T \text{Treatment}_i + \beta^X X_i + \beta^Z Z_i + \varepsilon_i. \quad (9)$$

The first four take $y_{i,\text{loss}}$ as the dependent variable. The first estimates Equation (9) for a sample limited to T1 and T3, where Treatment_i is an indicator variable taking the value of one for individuals in T3. The second estimates the equation for a sample limited to T2 and T4, where Treatment_i is an indicator variable taking the value of one for individuals in T4. The third estimates the equation for the total sample, where Treatment_i is an indicator variable taking the value of one for individuals in the two tax treatments, T3 and T4. The fourth estimates the equation for a sample restricted to T1 and T4, where Treatment_i is an indicator variable taking the value of one for individuals in T4. The last four versions take $y_{i,\text{gain}}$ as the dependent variable, but are otherwise similar to the first four.

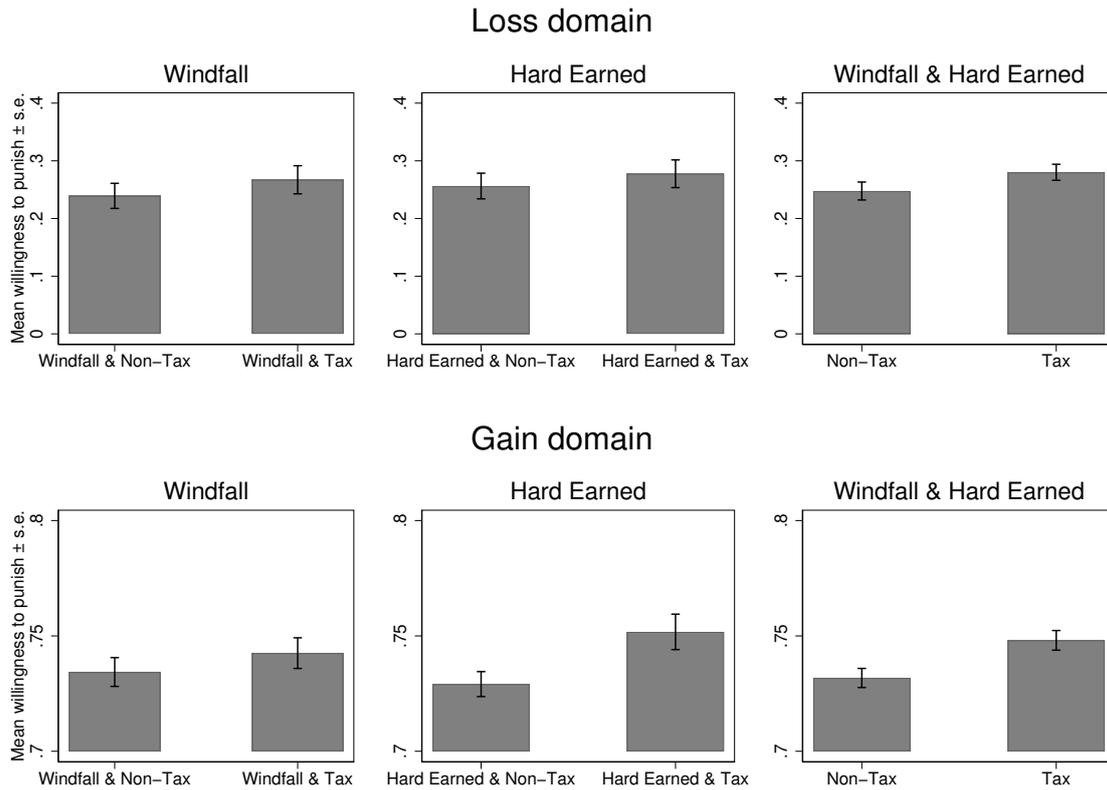
E.2 Results

The paper theorizes that the effect of taxation on demand for accountability can be explained by taxation causing fairness considerations to be more salient, which in turn makes citizens care more about, and more willing to punish, unfair government behavior. An alternative explanation is that citizens' willingness to punish the government is *reference dependent*, meaning that it is determined by their reference point for the government's behavior (Martin, 2016; Paler, 2013; Sandbu, 2006). More specifically, it could be theorized that citizens get a higher utility from punishing negative deviations from their reference point than from punishing positive deviations from their reference point, because negative deviations are perceived as losses (Kahneman and Tversky, 1979). Recent research argues that reference points are expectations-based (Köszegi and Rabin, 2006; 2007; 2009; Abeler et al., 2011; Marzilli Ericson and Fuster, 2011; Gill and Prowse, 2012; Banerji and Gupta, 2014), and the citizens' reference point can thus be modeled as expectations about what the government will do. Martin (2016) argues that the natural reference point for the citizens' expectations is their pre-tax income. Simplifying this theory, taxed citizens can be assumed to expect to receive benefits from the government equal to the tax they pay, whereas non-taxed citizens can be assumed to expect nothing. Thus, taxation causes a higher demand for accountability through increasing the citizens' expectations about the benefits the government will provide.

The present data can shed light on how well this simplified theory predicts willingness to punish by comparing behavior in Non-Tax to Tax-treatments. The theory predicts that citizens in the Tax treatments expect to get a monetary payoff equal to their pre-tax income, \$2. Their post-tax income is \$1, and the leader therefore has to invest a share of at least 0.7 of the \$2 group endowment in order for the citizen to get his pre-tax reference payoff. Citizens in the Non-Tax treatments, on the other hand, do not pay any tax, and their “pre-tax” income is \$1, which they get regardless of the leaders’ investments. Thus, the leaders do not have to invest anything in order for the citizens to get their reference payoffs in the Non-Tax treatments. If the citizens’ utility from punishment is higher for investments below the reference point, citizens in the Tax treatments should have a higher willingness to punish investment shares between 0 and 0.6 than citizens in the Non-Tax treatments, but there should be no difference in punishment behavior for investment shares of 0.7 and higher. This is because investment shares between 0 and 0.6 are below the reference point for citizens in the Tax treatments but above the reference point for citizens in the Non-tax treatments, whereas investments shares of 0.7 or higher are above the reference point for citizens in both Non-Tax and Tax treatments.

Figure E.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, for Hard Earned treatments 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. 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. Tables E.1 and E.2 show that these results replicate in a regression framework. 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. 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.

Figure E.1: Effect of Tax treatments below and above reference point



Note: **Upper panel:** The figure shows the means and estimated standard errors for the willingness to punish investments where citizens in the tax treatments are below their reference point. The left panel illustrates the effect of taxation when the group endowment is windfall (p-value of t-test = 0.39), the middle panel illustrates the effect of taxation when the group endowment is hard earned (p-value of t-test = 0.51) and the right panel illustrates the effect of taxation for windfall and hard earned group endowments combined (p-value of t-test = 0.12). **Lower panel:** The figure shows the mean and estimated standard errors for the willingness to punish for levels of investments where citizens in all treatments are above their reference point. The left panel illustrates the effect of taxation when the group endowment is windfall (p-value of t-test = 0.37), the middle panel illustrates the effect of taxation when the group endowment is hard earned (p-value of t-test = 0.015), and the right panel illustrates the effect of taxation for windfall and hard earned group endowments combined.

E.3 Main analysis

Table E.1: Effect of Tax treatments for investment levels below reference point

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

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)-(6) present results for the discrete punishment variable. 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 E.2: Effect of Tax treatments for investment levels above reference point

	Willingness to punish						High punishment dummy					
	Windfall		Hard Earned		All		Windfall		Hard Earned		All	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Tax	0.100 (0.111)	0.090 (0.114)	0.276** (0.114)	0.276** (0.113)	0.199*** (0.072)	0.197*** (0.072)	0.076 (0.112)	0.068 (0.114)	0.188 (0.116)	0.199* (0.115)	0.160** (0.073)	0.158** (0.072)
Male		-0.093 (0.088)		-0.195* (0.116)		-0.063 (0.067)		-0.088 (0.089)		-0.140 (0.120)		-0.050 (0.068)
Above median age		0.032 (0.110)		-0.134 (0.114)		0.026 (0.074)		0.003 (0.112)		-0.122 (0.115)		0.014 (0.075)
Above median education		0.103 (0.127)		-0.019 (0.126)		0.042 (0.082)		0.115 (0.127)		-0.031 (0.129)		0.072 (0.083)
Employed full-time		-0.035 (0.117)		0.050 (0.121)		0.004 (0.079)		-0.015 (0.120)		-0.038 (0.124)		-0.025 (0.079)
Above median income		0.017 (0.136)		0.034 (0.140)		0.031 (0.086)		-0.080 (0.135)		0.080 (0.136)		-0.015 (0.085)
Conservative		-0.147 (0.118)		-0.000 (0.116)		-0.020 (0.078)		-0.143 (0.112)		-0.045 (0.118)		-0.056 (0.076)
More politically engaged		-0.049 (0.115)		0.039 (0.114)		0.032 (0.076)		-0.106 (0.111)		0.006 (0.115)		-0.016 (0.075)
Constant	-0.081 (0.076)	-0.002 (0.168)	-0.145** (0.066)	-0.043 (0.144)	-0.112** (0.050)	-0.140 (0.106)	-0.088 (0.072)	0.056 (0.170)	-0.094 (0.072)	0.042 (0.154)	-0.091* (0.051)	-0.059 (0.109)
Observations	297	297	301	301	739	739	297	297	301	301	739	739
R ²	0.003	0.015	0.020	0.033	0.010	0.012	0.002	0.016	0.009	0.018	0.006	0.009

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)-(6) present results for the discrete punishment variable. Column (7)-(12) presents results for the high punishment indicator variable. 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.

Appendix F Leaders

Table F.1: Background by treatment, leaders

	Windfall & Non-Tax	Hard Earned & Non-Tax	Windfall & Tax	Hard Earned & Tax	Total	F-test
Male	0.43 (0.04)	0.49 (0.04)	0.52 (0.04)	0.54 (0.03)	0.50 (0.02)	0.59
Above median age	0.48 (0.04)	0.50 (0.04)	0.51 (0.04)	0.50 (0.03)	0.50 (0.02)	0.98
Above median education	0.55 (0.04)	0.60 (0.03)	0.49 (0.04)	0.54 (0.03)	0.55 (0.02)	0.11
Employed full time	0.54 (0.04)	0.61 (0.03)	0.50 (0.04)	0.57 (0.03)	0.56 (0.02)	0.09
Above median income	0.46 (0.04)	0.47 (0.04)	0.45 (0.04)	0.46 (0.03)	0.46 (0.02)	0.89
Conservative	0.43 (0.04)	0.48 (0.04)	0.43 (0.04)	0.44 (0.03)	0.45 (0.02)	0.50
More politically engaged	0.47 (0.04)	0.49 (0.04)	0.49 (0.04)	0.45 (0.03)	0.47 (0.02)	0.57
Observations	194	199	183	381	957	

mean coefficients; semean in parentheses

Note: The table provides mean values 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. See Table 3 for definitions of background variables.

Table F.2: Effect of treatment on background variables, leaders

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

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.

Table F.3: Preferences by treatment, leaders

	Windfall & Non-Tax	Hard Earned & Non-Tax	Windfall & Tax	Hard Earned & Tax	Total	F-test
Risk averse	0.47 (0.04)	0.46 (0.04)	0.53 (0.04)	0.45 (0.03)	0.47 (0.02)	0.21
More altruistic	0.42 (0.04)	0.32 (0.03)	0.34 (0.04)	0.31 (0.02)	0.34 (0.02)	0.72
High positive reciprocity	0.45 (0.04)	0.48 (0.04)	0.45 (0.04)	0.45 (0.03)	0.46 (0.02)	0.80
High negative reciprocity	0.54 (0.04)	0.48 (0.04)	0.56 (0.04)	0.49 (0.03)	0.51 (0.02)	0.10
Loss averse	0.54 (0.04)	0.45 (0.04)	0.42 (0.04)	0.48 (0.03)	0.48 (0.02)	0.42
Observations	194	199	183	381	957	

mean coefficients; semean in parentheses

Note: The table provides mean values for preferences in the four respective treatments and for the whole sample of leaders. The F-test column provides the p-value for an f-test of no difference in means between the four treatments. See Table B.3 for definitions of variables.

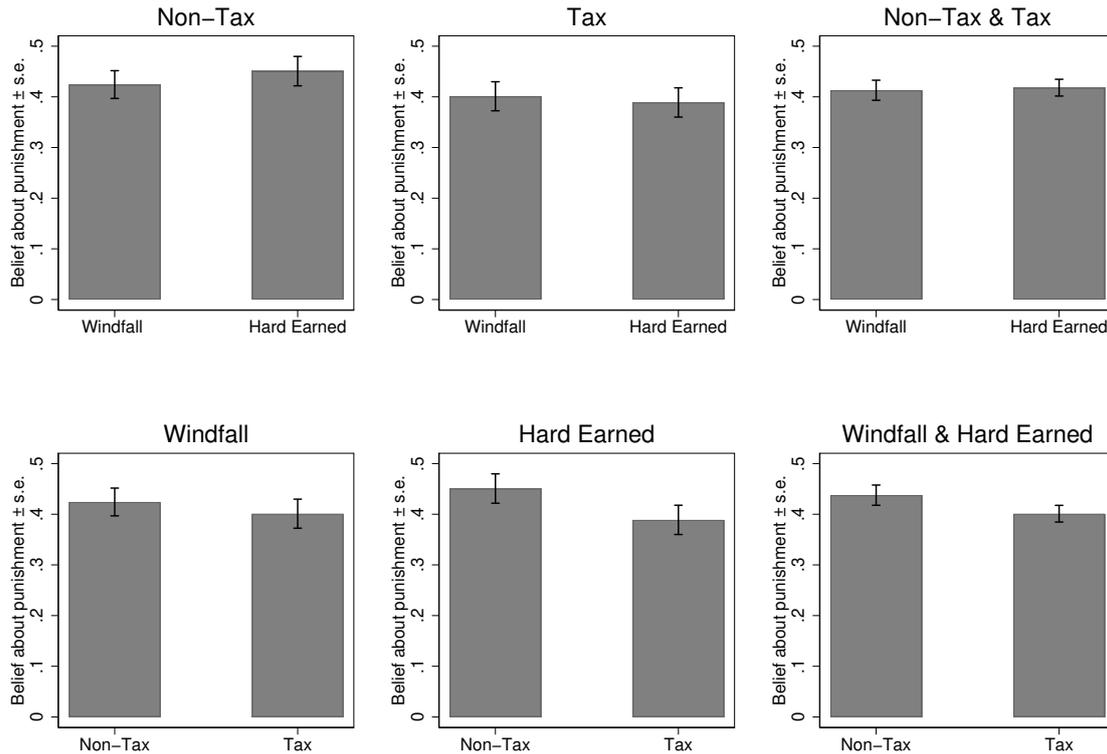
Table F.4: Effect of treatment on preferences, leaders

	Risk averse	More altruistic	High pos. reci.	High neg. reci.	Loss averse
Hard Earned & Non-Tax	-0.037 (0.050)	-0.106** (0.049)	0.014 (0.050)	-0.049 (0.050)	-0.079 (0.050)
Windfall & Tax	0.035 (0.052)	-0.078 (0.050)	-0.011 (0.051)	0.016 (0.051)	-0.099* (0.051)
Tax State	-0.041 (0.044)	-0.113*** (0.043)	-0.015 (0.044)	-0.079* (0.044)	-0.043 (0.044)
Constant	0.490*** (0.036)	0.423*** (0.036)	0.454*** (0.036)	0.546*** (0.036)	0.536*** (0.036)
P-value of F-test	0.213	0.714	0.795	0.103	0.415

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 preference indicator variables for leaders as the dependent variable on indicator variables for the treatments. The Rentier State treatment is the reference category. See Table B.3 for definition of the variables. The reported p-value tests the hypothesis that all the treatments have the same effect on the preference indicator variables.

Figure F.1: Effect of treatments on leaders' beliefs about punishment



Note: The figure shows the mean and error bars for the lowest share invested the leader thinks the citizen does not punish. **Upper panel:** The left panel illustrates the effect of the Hard Earned treatment when the group endowment is given directly to the leader, the middle panel illustrates the effect of Hard Earned treatment when the group endowment is collected through tax, and the right panel illustrates the pooled effect of the Hard Earned treatment. **Lower panel:** The left panel illustrates the effect of the tax treatment when the group endowment is windfall, the middle panel illustrates the effect of the tax treatment when the group endowment is earned, and the right panel illustrates the pooled effect of the Tax treatment.

Appendix G Instructions

G.1 Introduction

Figure G.1: Introduction (all participants)



Introduction

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

G.2 Role assignment

Figure G.2: Role assignment (all citizens)

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.

Figure G.3: Role assignment (all leaders)

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.

G.3 Task description (citizens)

Figure G.4: Task description, Rentier State

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)

Figure G.5: Task description, Hard Earned & Non-Tax

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)

Figure G.6: Task description, Windfall & Tax

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)

Figure G.7: Task description, Tax State

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)

[Control questions]

G.4 Punishment decision

Figure G.8: Punishment decision, Rentier State

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.

The leader you have been grouped with also completed the picture categorization assignment and earned \$1.

Your group receives \$2 in group money unconditional on your or the leader's completion of the assignment.

The leader has been given the task of deciding how much of the group money to invest in the common pool. The leader gets the amount of group money not invested in the common pool.

You will be given the opportunity to reduce the leader's payment at a small cost. You can choose between:

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

You will make this choice for each possible decision the leader can make about how much of the group money to invest in the common pool. We will implement your choice for the leader's actual decision.

You will not get to know the leader's decision until we pay out the additional payments from the experiment. This will be done as soon as possible and within 3 weeks of submission of the HIT.

We now ask you to make your decisions about what you would like to do 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)	<input type="radio"/>	<input type="radio"/>
\$0.2 (leader: \$1.95, YOU: \$0.15)	<input type="radio"/>	<input type="radio"/>
\$0.4 (leader: \$1.90, YOU: \$0.30)	<input type="radio"/>	<input type="radio"/>
\$0.6 (leader: \$1.85, YOU: \$0.45)	<input type="radio"/>	<input type="radio"/>
\$0.8 (leader: \$1.80, YOU: \$0.60)	<input type="radio"/>	<input type="radio"/>
\$1.0 (leader: \$1.75, YOU: \$0.75)	<input type="radio"/>	<input type="radio"/>
\$1.2 (leader: \$1.70, YOU: \$0.90)	<input type="radio"/>	<input type="radio"/>
\$1.4 (leader: \$1.65, YOU: \$1.05)	<input type="radio"/>	<input type="radio"/>
\$1.6 (leader: \$1.60, YOU: \$1.20)	<input type="radio"/>	<input type="radio"/>
\$1.8 (leader: \$1.55, YOU: \$1.35)	<input type="radio"/>	<input type="radio"/>
\$2.0 (leader: \$1.50, YOU: \$1.50)	<input type="radio"/>	<input type="radio"/>
