

# Does voting on tax fund destination imply a direct democracy effect?

TARC DISCUSSION PAPER 034-21

NICOLAS JACQUEMET, ÉCOLE D'ÉCONOMIE DE PARIS  
STÉPHANE LUCHINI, AIX-MARSEILLE UNIVERSITÉ  
ANTOINE MALÉZIEUX, BURGUNDY SCHOOL OF BUSINESS

# Does Voting on Tax Fund Destination Imply a Direct Democracy Effect?\*

Nicolas Jacquemet<sup>†</sup>

Stéphane Luchini<sup>‡</sup>

Antoine Malézieux<sup>§</sup>

June 2021

## Abstract

Does giving taxpayers a voice over the destination of tax revenues lead to more honest income declarations? Previous experiments have shown that giving participants the opportunity to select the organization that receives their tax funds tends to increase tax compliance. The aim of this paper is to assess whether this increase in compliance is induced by the sole fact of giving subjects a choice — a “direct democracy effect”. To that aim, we ask participants to a tax evasion game to choose, in a collective or individual choice setting, between two very similar organizations which provide the same social (ecological) benefits. We elicit compliance for both organizations before the choice is made so as to control for the counter-factual compliance decision. We find that democracy does not increase compliance, and even observe a slight negative effect — in particular for women. Our results confirm the existence of a commitment effect of democracy, leading to favor more the selected organization when it was actively chosen. The commitment effect of democracy is however not enough to overcome the decrease in the level of compliance. Thanks to response times data, we show that prior choice on similar options as compared to a purely random selection weakens the preference for honesty. One important field application of our results is that democracy in tax spending must offer real choices to tax payers to improve compliance.

**Keywords:** Commitment; Direct democracy effect; Voting; Tax evasion game.

**JEL classification:** C9, D7, H2, H3.

---

\*We thank Niels Boissonnet, Béatrice Boulu-Reshef, Léa Bousquet, John D’Attoma, Pedro Dal Bó, Elisa Darriet, Miguel Fonseca, Duccio Gamannossi, Antoine Hémon, Nicolas Jacquemet, Matthew Johns, Erich Kirchler, Fabrice Le Lec, Tejas Nair, Olivier Simard-Casanova, Angela Sutan, Jean-Robert Tyran, Claire Vandendriessche, Marie-Claire Villeval, Ingrid Wahl for their helpful comments, as well as Kene Boun My for his assistance in running the experiments. We are also grateful to Wahl, Muehlbacher, and Kirchler; Brockmann, Genschel, and Seelkopf; Lamberton, De Neve, and Norton for sharing their data. Financial support from the *Conseil Régional de Lorraine*, the *Tax Administration Research Centre* and the National Research Agency (program Investissements d’Avenir, ANR-10-LABX-93-0 and ANR-17-EURE-0001), are gratefully acknowledged.

<sup>†</sup>Paris School of Economics and Université Paris 1 Panthéon–Sorbonne. Centre d’Economie de la Sorbonne, 106 Bd. de l’Hôpital, 75013 Paris. Nicolas.Jacquemet@univ-paris1.fr

<sup>‡</sup>Aix-Marseille Université. (Aix-Marseille School of Economics) and CNRS, 5-9 Boulevard Maurice Bourdet, 13001 Marseille, France. stephane.luchini@univmed.fr

<sup>§</sup>CEREN, EA 7477, Burgundy School of Business, Université Bourgogne Franche-Comté, 29 Rue Sambin, 21000 Dijon, France antoine.malezieux@bsb-education.com.

# 1 Introduction

It is now well-established that compliance with the law does not solely depend on the content of the law itself, but is shaped by the process leading to its implementation. According to a large strand of literature in behavioral law and economics, people tend to be more respectful of the same rules when they had a voice in their adoption as compared to when the rule is exogenously imposed on them. Such a direct democracy effect has been observed in a wide range of economically relevant situations like, e.g., when people take part in setting goals for themselves (Locke and Latham, 2002); when farmers decide on irrigation rules (Bardhan, 2000), or when workers have a voice in organizing their work (Ichniowski, Shaw, and Prennushi, 1997).

Most of the existing experimental literature focuses on the effect of direct democracy on cooperative behavior in social dilemma, and confirms drastic increases in compliance when the rules are democratically chosen before they are implemented. The empirical identification of such a spillover of democracy on compliance behavior faces the obvious challenge of self-selection into the policy when the choice of the rules is endogenous. Dal Bó, Foster, and Putterman (2010) are, to the best of our knowledge, the first to disentangle the selection effect and the endogenous effect of direct democracy in the context of fines imposed on deviations in a prisoner’s dilemma game. They find a modest selection effect (8% of the overall observed change in behavior when fines are enforced democratically rather than exogenously) and confirm that direct democracy induces a large increase in cooperation: the effectiveness of fines is 40% larger when chosen endogenously.

In this paper, we empirically investigate the direct democracy effect on tax compliance of giving taxpayers a voice over the use of collected taxes. While this question has received less attention in the literature, a few existing studies, starting with Alm, Jackson, and McKee (1993), tend to suggest that tax compliance is higher when expenditures are decided endogenously. This echoes recent trends in the policy sphere, where direct democracy becomes widespread as a tool to strengthen citizens commitment to the tax system.<sup>1</sup> Such an application of direct democracy is however likely to elicit strong selection effects, as citizens are offered the choice of which public good they want to fund. The aim of our study is to measure the direct democracy effect of such a policy while taking into account the change in preference-driven compliance. Our analysis proceeds in three step.

The first step of our study, presented in Section 2.1, is based on a new analysis of the individual data from previous experiments. The estimation results on pooled data show an at most weak effect of direct democracy on compliance. We then move to an original experiment (presented in Section 3) designed so as to disentangle the causal effect of direct democracy from self-selection. Our baseline extends the design of Jacquemet, Luchini, Malézieux, and Shogren (2020) experiment

---

<sup>1</sup>As an example, the city of Paris asks its citizens to decide on the destination of 5% of its spendings (see <https://budgetparticipatif.paris.fr/bp/le-budget-participatif-.html>). Also see Matsusaka (2005) for examples in the US.

to tax compliance in favor of a randomly chosen organization. To account for preference-driven compliance, we purposefully chose two organizations that are as close as possible from the point of view of subjects' preferences, we elicit compliance for both organizations before the organization is selected, and we elicit subjects' preferences about the two organizations at the end of the baseline experiment. In our main treatment of interest, the organization is rather chosen by majority voting in groups of three subjects. The third step of our analysis builds on additional treatments aimed at clarifying the mechanisms that can possibly generate a direct democracy effect. First, we measure the effect of social coordination thanks to a variation of the vote treatment in which free-form communication is allowed between subjects before the voting stage. Second, at the other extreme, we eliminate the social context in a treatment in which the organization is chosen individually by subjects. Third, in all treatments, our design provides counter-factual observations of what compliance would have been would the non-selected organization had been selected. The resulting gap in compliance provides a measure of the commitment effect of direct democracy.

The results, described in Section 4, are threefold. First, we fail to replicate the increase in compliance previously observed when the use of collected taxes is decided by vote. We rather observe a decrease in the level of compliance for the preferred organization, which is mainly driven by female participants' response to the treatment. Second, based on the gap in compliance, our results confirm a commitment effect of democracy: the gap in compliance in favour of the preferred organization is larger when it has been actively chosen. This change is however dominated by the negative direct democracy effect, resulting in a lower level of compliance when democracy is implemented. More precisely, the main effect of direct democracy in our data is to sharply increase the share of partial compliers, at the expense of a lower share of full compliers. Third, as compared to a simple vote, our results confirm a slight increase in compliance when subjects are able to coordinate before deciding on their preferred organization, but we do not find any difference in compliance between a vote and a single decision-making situation. We discuss the implications of our results and the open avenues for further research in Section 5.

## 2 Existing evidence on the direct democracy effect

The behavioral consequences of offering decision-makers the possibility to choose the environment in which they behave has been studied in different settings in the empirical literature, ranging from the amount donated in a dictator game, to the allocation rule collectively implemented in a common pool resource game (see the Appendix, Section A for a summary of empirical studies measuring the DDE).<sup>2</sup> While a few studies introduce a vote mechanism into a dictator game or a trust game, most of the literature focuses on Public Good Games (PGG), and, to a lesser extent, on Common Pool Resource (CPR) games and Tax Evasion Games (TEG).

---

<sup>2</sup>We restrict to studies featuring a comparison between a vote and a baseline condition. This literature review is based on primary and secondary citations of the papers listed.

Interestingly, the feature of the game which is decided by vote is quite different across settings. Most of the literature using a PGG or a CPR investigates the consequences on contributions or extractions of allowing members of the group to choose an institution promoting cooperation — *e.g.*, fines and rewards, contribution rules, provision thresholds etc. When TEG are used the vote is rather implemented so as to choose how the amount collected is spent — and hence, what is the nature of the public good that is produced thanks to taxes. In contrast with the choice of institutions promoting cooperation in PGG (Dal Bó, Foster, and Putterman, 2010; Dal Bó, Foster, and Kamei, 2019), we are aware of no study measuring the DDE of choosing the destination of the funds collected in a TEG that takes into account the confounding effect of self-selection into the outcome of the vote — *i.e.*, subjects who are more likely to vote in favor of a particular use are also more likely to comply when the funds go to it, in such a way that the relation between the change in outcome behavior and the result of the vote is correlational rather than causal. This is the first aim of this study.

## 2.1 Why should we expect a direct democracy effect?

In spite of the large heterogeneity in the literature regarding the research question, the discussions about the mechanism that could possibly lie behind a Direct Democracy Effect (DDE) leads to a rather small set of arguments.

The first group of arguments is related to the behavioral consequences of the improvement in *legitimacy* that comes with a more democratic political process (see Grossman and Baldassarri, 2012, for a review of the literature in political sciences and psychology). An important driving force of such consequences is that citizens are more willing to defer to the power of authorities if this power is obtained through a fair procedure. Wahl, Muehlbacher, and Kirchler (2010), for instance, provide empirical evidence on this channel based on the observed correlation between compliance and the perceived procedural fairness of the allocation of collected funds when this decision is taken by vote. A second, related argument, that is more specific to tax compliance, is that direct democracy induces “*taxpayer agency*” (Lamberton, De Neve, and Norton, 2018). According to this mechanism, the lack of agency induces reactance (Brehm, 1966) which results in lower compliance. Compliance should thus raise when taxpayers are given a voice about the use of collected taxes and are allowed to express their preferences.

These two mechanisms are expected to occur in the context of single decision-making about the use of the funds: they both predict higher compliance as soon as a taxpayer was able to first express preferences about the best way to spend the tax revenues. But when such decisions are made by vote, the process also delivers information about the prevalent preferences in the group. A third mechanism, *social coordination*, might then underlie a DDE because the democratic process works as a coordination device, provides a signal of one’s own type to other members of the voting group, or is used as a way to influence other voters’ behavior through peer pressure. As an

illustration, Feld and Tyran (2002) offered participants the choice to vote in groups of three for a penalty that can be levied on free-riders. Using the strategy method, participants decide how much they want to contribute to a public good in three different situations: in case either 0, 1, or 2 other subjects approved the penalty. Although two votes in favor of the penalty are enough for it to be implemented, the results show that the more other subjects in the group approved of the penalty (from 1 to 2), the higher the compliance rate amongst subjects who vote ‘yes’ (from 67% to 90%). This change, however, mixes a DDE induced by social coordination and self-selection since contributions are elicited conditional on a choice in favor or against the policy.

A fourth possible channel is *commitment*. According to social psychology (Kiesler and Sakumura, 1966; Kiesler, 1971; Joule and Beauvois, 1998) decisions taken freely commit people to behave consistently with this choice in subsequent decision-making situations. Such behavior arises due to either or both the willingness to avoid cognitive dissonance (the psychological tension induced by holding different beliefs or values; Festinger, 1957) or self-attribution (past actions serve a signal to oneself about one’s own, unknown, preferences; Bem, 1972; Bodner, Prelec, Brocas, and Carillo, 2003), both leading to a preference for self-consistency. In a classical experiment, Deutsch and Gerard (1955) for instance ask a pool of students to judge the length of lines that are shown to them. Once students are provided with additional information that their judgments were wrong, those who had to self-report their initial judgment are significantly less likely to revise their estimate than those who just kept their estimates for themselves. The strength of such commitment is larger when the decision is made publicly. In this same experiment, the share of students who decide not to revise their estimate is twice lower among those whose self-report is made public (see Cialdini, 2007, chapter 3, for a detailed review of behavioral changes induced by commitment). In the context of a policy implemented after an explicit choice from taxpayers, having expressed a preference for a particular policy would then make citizens feel committed to be consistent with their preferences, and hence comply more with the law.

The second aim of our study is to further investigate the mechanisms lying behind the DDE that (possibly) arises when the use of collected funds is decided by vote in a TEG. In contrast with legitimacy, that can be directly observed thanks to self-reported legitimacy and fairness questions, the other channels are rather hard to disentangle empirically. Our empirical strategy relies on comparing single-decision making, with no social dimension, to an otherwise identical group voting. We moreover provide evidence on the commitment effect of giving taxpayers a voice by measuring the gap in compliance that arises between the chosen and the unchosen option, for given preferences.

## 2.2 Democracy in tax evasion games

The existing literature which aims to measure a DDE in the context of a TEG is rather mixed. We are aware of four such studies, providing a between-subject comparison in the level of compliance

between a baseline condition in which the use of collected taxes is exogenously decided by design, and a treatment in which this same use is decided democratically. The Appendix, Section A.1, provides an empirical analysis based on existing evidence. Based on these results, accumulated evidence from previous experiments suggests that DDE in TEG is barely significant, and rather small in magnitude; suggesting that most of the effect documented in the empirical literature could be due to a self-selection of voters with particular preferences into particular policies, leading to more support for those policies.

### 3 Experimental design

The design of our experiment aims to provide evidence on the DDE while controlling for self-selection. The treatment variable is the choice offered to subjects about how collected taxes are used. To better understand how the mechanisms behind the effect of direct democracy work, we want to compare behavior after a vote within a group to behavior after an individual decision which does not induce any social effect. For that purpose, our target behavior is compliance in a TEG which can be easily adapted to either an individual or a group decision.

#### 3.1 Tax evasion game

The structure of the tax evasion game is common to all treatments, and proceeds in two steps. First, participants earn their income in a task borrowed from Alm, Cherry, Jones, and McKee (2008), in which they have to sort digits in ascending order as quickly as possible. The quicker a subject is the more money he earns, according to the following compensation rule (labeled in an Experimental Currency Unit):  $150 - (time \times 13)$ . Second, participants are asked to report the amount of income they earned in the first stage, knowing that it will be taxed at a 35% rate.

The amount of money collected through taxes in the experiment is donated to either of two organizations (between which subjects are asked to choose before the declaration stage in all treatments but the baseline, see Section 3.2 below): the WWF (*World Wide Fund for Nature*, [www.wwf.fr](http://www.wwf.fr)) or ASPAS (*Association pour la Protection des Animaux Sauvages*, [www.aspas-nature.org](http://www.aspas-nature.org)), a French organization for the protection of wildlife. The two organizations have been chosen according to two main criteria. First, both deliver donation certificates as part of their public campaign for funding. In the instructions, we inform participants that they will receive an individual certificate issued by the organization that received their donation, with an explicit mention of the identity of the donor.<sup>3</sup> This makes the donation credible to participants and avoids any risk that subjects believe there might be deception in the use of collected taxes. Second, and more importantly, the results obtained by Alm, Jackson, and McKee (1993) show

---

<sup>3</sup>The design of the experiment closely follows Jacquemet, Luchini, Malézieux, and Shogren (2020), who elicit compliance in favor of the WWF. An English translation of the original instructions in French is provided in the Appendix, Section B.1.

that compliance increases significantly when the use of collected taxes matches an option that is strictly preferred by subjects over another. This is the source of the main confounding factor in the identification of the effect of direct democracy (Dal Bó, 2014): if compliance is observed to increase when the organization is first chosen by subjects, it could be either because of a DDE; or because of self-selection, *i.e.*, subjects who vote for a particular organization are also more willing to donate to this organization — hence complying more. Our aim is to provide a causal measure of DDE — the change in compliance induced by the mere fact of selecting the organization receiving the money. We thus chose organizations that are as close as possible to one another so as to ensure as little difference in preferences among subjects between the two.<sup>4</sup>

Although this design choice will likely undermine self-selection, we further neutralize its confounding effect by applying the strategy method to the income reporting decision: subjects are asked how much income they want to declare in both the situation where WWF turns-out to be selected and the situation where it is ASPAS. We thus elicit two compliance decisions per subject. Only one of the two is binding, depending on the organization that happens to be selected. Participants are informed about the selected organization, and their resulting payments, once all decisions have been made.<sup>5</sup>

At the end of the experiment, subjects are asked to answer a questionnaire aimed at collecting subjective information about the experiment. Specifically, we ask them “*What was the importance of choosing the organization receiving the tax?*” (5 points Likert scale: Extremely important, very important, slightly important, not important, not at all important); “*How would you judge the legitimacy of the process you went through to designate the organization getting the tax?*” (5 points Likert scale: Very illegitimate, illegitimate, neither illegitimate or legitimate, legitimate, very legitimate); “*How would you judge the fairness of the process you went through to designate the organization getting the tax?*” (5 points Likert scale: Very unfair, unfair, neither unfair or fair, fair, very fair); “*Did you know WWF/ ASPAS before taking part to this experiment?*” (yes/no), “*What is your opinion on WWF/ASPAS’s actions?*” and “*What is your opinion about WWF/ASPAS?*” (7 points Likert scale: totally anti, anti, rather anti, indifferent, rather pro, pro, totally pro).

### 3.2 Experimental Treatments and Hypotheses

Our between-subject treatments implement changes in the way the organization is chosen. This choice adds an additional step in the experiment, which happens just before the income reporting

---

<sup>4</sup>A downside of this criterion is that there are only very few alternatives to the WWF, leading us to choose an organization which is by far less famous. As shown below, Section 3.3, this does not impact neither subjects’ perception about the organizations, nor their level of compliance.

<sup>5</sup>Dal Bó discusses two additional identification issues: “endogenous democracy”, *i.e.*, that groups in which the choice is available might differ from groups in which it is not and “policy choice”, the idea that the set of policies available reflects individual preferences. Both are easily tackled in the context of a controlled experiment, thanks to randomization of participants to the treatments and because the set of alternatives is pre-determined by the experimenter. Self-selection, by contrast, contributes to the observed change in behavior happening in line with a vote as soon as subjects have strong preferences over the available options.

part of the TEG. There is no such step in the BASELINE treatment, in which a random draw decides (with equal probability) whether WWF or ASPAS is selected. As an additional control variable, we ask participants of the BASELINE which organization they would have chosen if they had the opportunity at the end of the experiment.

In our main treatment of interest, VOTE, participants are randomly and anonymously matched into groups of three participants. After having earned their income, they learn about the tax simulation that is about to come but are asked, before the declaration stage, to vote in favor of the organization (either WWF or ASPAS) that will receive the taxes collected from all members of the group. At the end of the experiment, the organization is selected based on majority voting.<sup>6</sup> This treatment replicates the typical design used to measure the effect of a vote on compliance. As compared to compliance for the preferred organization in the baseline condition, a DDE thus predicts higher compliance when the same organization has been chosen by vote.

**Hypothesis 1** (*DDE*) *As compared to BASELINE, VOTE elicits an higher level of compliance for the preferred organization.*

In the VOTE+ treatment, our aim is to reinforce the social effect of the voting stage. To that end, we add cheap talk between participants, in the sense that participants enter a computerized and anonymous group chat with the members of their trio for up to three minutes, before participating to the VOTE treatment. Subjects are explicitly told that the aim of this step is for them to collectively decide on the organization they would like to vote for. In both the VOTE and the VOTE+ treatments, the uncertainty that subjects actually face regarding the selected organization depends on their beliefs about the vote of other subjects. To control for such heterogeneity, we add an item to the final questionnaire in both treatments, asking subjects “*Did you think at least another participant would vote like you did?*” (yes/no/don’t know).

**Hypothesis 2** (*Social coordination*) *Thanks to the opportunity to coordinate, participants in VOTE+ comply more for the preferred organization than participants in both VOTE and BASELINE.*

In both VOTE and VOTE+, the decision about the use of collected taxes is made within a group. To isolate the possible effect of giving taxpayers some agency over the use of the taxes, our last treatment replicates the endogenous choice of the organization but in the context of a single decision making rather than inside a group. We design this treatment as a probabilistic choice of the organization so as to mimic the uncertainty that subjects face in the VOTE treatment. Participants are asked to choose between two possible options: in option 1, WWF is selected with probability 2/3, while ASPAS will receive the funds with a 1/3 probability; option 2 maintains the same probability distribution but favors ASPAS (selected with a 2/3 probability) over WWF

---

<sup>6</sup>An English translation of the original treatment-specific instructions in French is provided in the Appendix, Section B.2.

(1/3 probability). We can moreover use the answers to the legitimacy question included in the post-experiment questionnaire to investigate the role played by the perceived legitimacy of the selection process in the observed changes in compliance.

**Hypothesis 3** (*Taxpayer agency*) *Compliance for the preferred organization is higher in CHOICE than in BASELINE, but lower than in both VOTE and VOTE+.*

**Hypothesis 4** (*Legitimacy*) *The difference in compliance for the preferred organization between BASELINE and all other treatments is increasing in the perceived legitimacy of the selection process.*

A common feature in all treatments as compared to BASELINE is that participants actively choose one organization over the other. If such a choice induces self-commitment, we should observe an higher level of compliance for the selected organization as compared to the non-selected organization (Rauchdobler, Sausgruber, and Tyran, 2010). Since commitment is stronger when it is made publicly, a corollary of this hypothesis is that commitment should be stronger when the decision is made in group, and even stronger when the group decision happens after the ability to coordinate.

**Hypothesis 5** (*Commitment*) *The gap in compliance between the selected and the non-selected organization is higher in CHOICE than in BASELINE; higher in VOTE than in CHOICE and higher in VOTE+ than in VOTE.*

### 3.3 Procedures

The data is made of ten experimental sessions, two for BASELINE and CHOICE and three for VOTE and VOTE+, each including between 24 and 27 subjects (variations are due to the show-up rate among invited participants). We over-sampled the VOTE and VOTE+ conditions because the voting groups induce within subjects correlations that lowers the power of the statistical analysis. All sessions took place in the Strasbourg University lab (LEES) from January 2016 to January 2018. The recruitment of subjects was carried out from LEES database among people who had successfully completed their registration on the ORSEE recruitment system (Greiner, 2015). The experiment was computerized on the web platform EconPlay ([www.econplay.fr](http://www.econplay.fr)). Each session lasted for about an hour, with an average payoff of 20 euro (17 euro directly given to the participants and 3 euro given to a charity), including a 5 euro show-up fee.

Overall, the experiment involved a total of 250 subject — 97% of whom are students. Table 1 reports summary descriptive statistics on the pool of subjects, along with the results from F-tests on the joint significance of the differences between treatments. A few of these comparisons turn-out significant. This is mainly driven by the higher share of women in BASELINE — as women in our sample tend to be older than men, less likely to study economics, to perceive organizations better, and have a better opinion of their actions. We will thus condition all statistical analysis

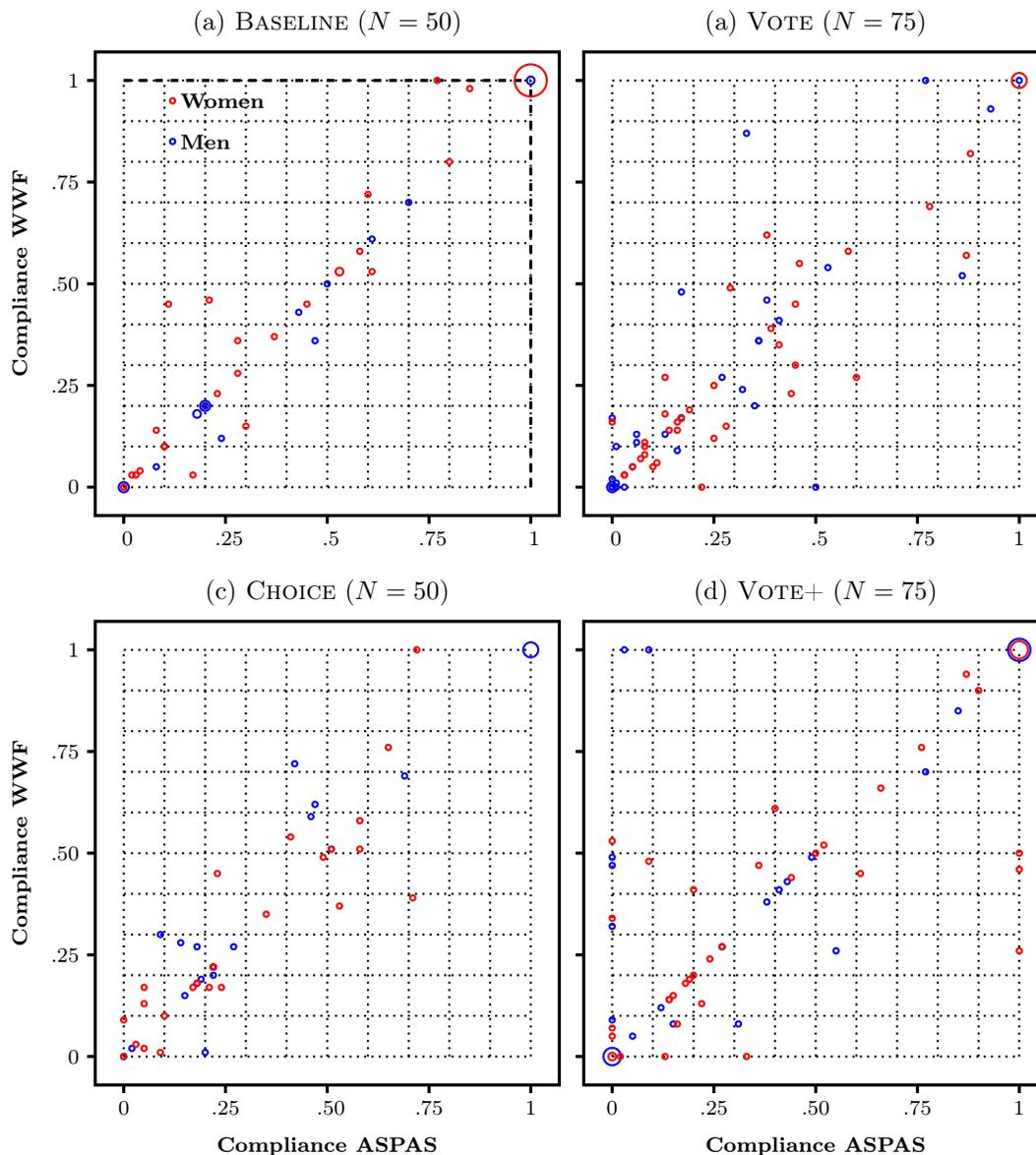
Table 1: Summary statistics on individual covariates and compliance measures

	BASELINE ( $N = 50$ )	CHOICE ( $N = 50$ )	VOTE ( $N = 75$ )	VOTE+ ( $N = 75$ )	F-test $p$ -values
Age	22.2	20.8	21.1	22.3	.0000
Men	30%	50%	45%	44%	.2015
Not speaking French at home	38%	30%	33%	28%	.6788
Students in economics	34%	34%	67%	41%	.0002
Monthly income (in euros)	540	410	473	540	.1860
French nationality	76%	78%	83%	87%	.4249
Believing in God	32%	44%	47%	40%	.4146
Parents' financial help	76%	78%	73%	64%	.2943
Self honesty	5.14	5.00	4.92	5.61	.1613
Happiness	4.84	5.12	4.93	5.05	.6427
Experimental Income	347	348	343	348	.9717
Knew WWF	82%	82%	81%	87%	.8198
Knew ASPAS	12%	14%	5%	12%	.3794
Perception of WWF's actions	6.06	5.72	5.52	5.84	.0291
Perception of ASPAS's actions	5.66	5.54	5.37	5.56	.5441
Opinion on WWF	5.94	5.64	5.50	5.69	.1353
Opinion on ASPAS	5.72	5.54	5.24	5.57	.0865
Importance of the stake	3.14	3.24	3.29	3.17	.8465
Legitimacy of selection process	3.16	2.76	3.24	3.42	.0008
Fairness of selection process	3.24	2.76	3.18	3.21	.0141

**Note.** Summary statistics on individual covariates and answers to the post-experiment questionnaire, by treatment. The last column provides the  $p$ -value from F-test on the joint significance of the differences between treatments. Significance levels: \*: 10%, \*\*: 5%, \*\*\*: 1%

on this variable. Importantly, the income that subjects earn and will be asked to report in the experiment does not vary on average across conditions (the distributions of earned income by treatment and gender are provided in the Appendix, Section C.1). In all treatments, participants have an excellent perception of both organizations and their actions, with a slight preference for WWF over ASPAS. We do not observe any bias towards a particular organization in subjects choices — which confirms the high level of similarity of both organizations in subjects' view. In both BASELINE and CHOICE, 58% of subjects select WWF (29/50 in both conditions) and 42% select ASPAS (21/50), which is statistically no different from a 50% split ( $p = .322$  and  $p = .322$ , proportion tests). Similarly, 55% of subjects select the WWF (41/75) in the VOTE treatment, and 57% (43/75) of subjects do so in VOTE+. Both are statistically equivalent to a perfect split of subjects between organizations ( $p = .488$  and  $p = .248$ ). We also do not reject the hypothesis that the distribution of subjects across organizations does not vary over treatments ( $p = .977$ , Chi-square test).

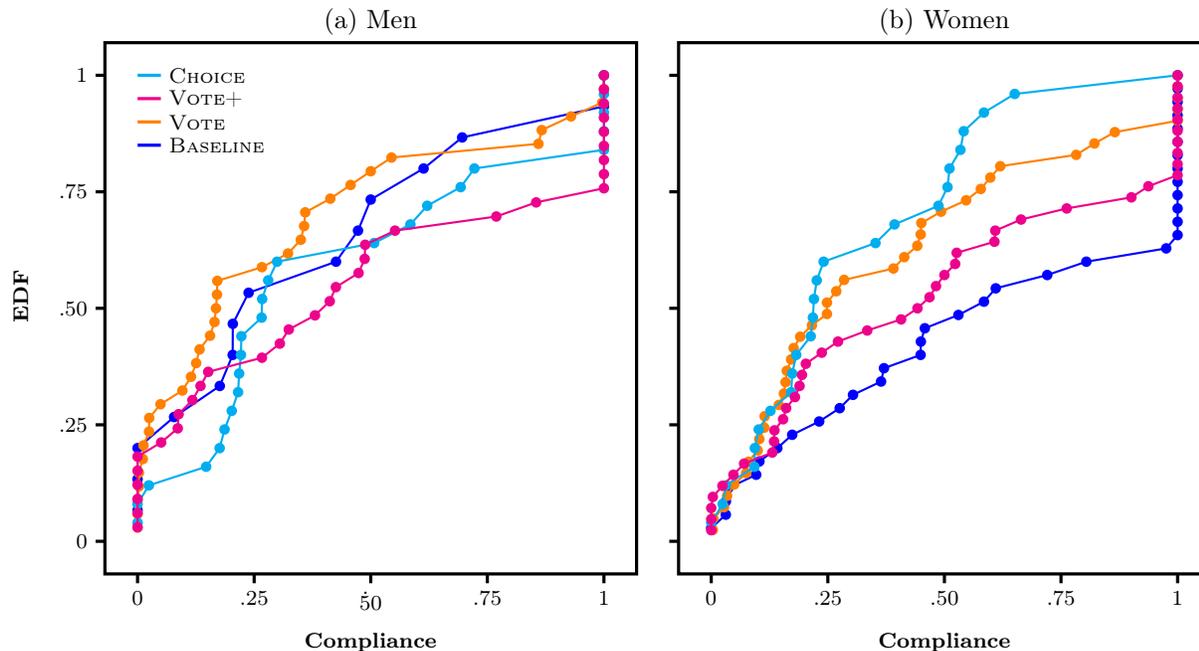
Figure 1: Joint distribution of compliance for ASPAS and WWF by treatment and gender



## 4 Results

Figure 1 provides descriptive statistics on the main outcomes from the experiment. We present the joint distribution of compliance for ASPAS and WWF by treatment and gender (to account for the imbalance between treatment on this dimension, see Section 3.3). The size of the circles are proportional to the number of subjects. The figure indicates that, in all treatments, most subjects are located around the  $45^\circ$  line, with a mass point at  $(1, 1)$ , for women in BASELINE. The correlation between the two is equal to  $.977$  in BASELINE. This also holds for CHOICE, in which

Figure 2: Empirical distribution functions (EDF) of compliance for the selected organization



the correlation coefficient is .940, and as well, although to a lesser extent, in VOTE (.842) and in VOTE+ (.752). This evidence confirms that compliance towards the two associations are very much alike.

#### 4.1 Causal evidence on the (lack of) DDE

In order to document the DDE in our study, we first focus on the selected organization — based on the post experiment questionnaire in BASELINE, the individual decision in CHOICE and the vote in VOTE and VOTE+. Figure 2 provides an overview of the compliance observed in each treatment, based on the Empirical Distribution Functions (EDF) of compliance for the selected organization by treatment and gender. Each dot in the figure represents a subject. According to bootstrap Kolmogorov-Smirnov (KS) tests on the EDF of men, displayed in Figure 2.a, men exhibit no increase in compliance in either of the treatments as compared to BASELINE ( $p = .241$  in VOTE,  $p = .623$  in VOTE+ and  $p = .774$  in CHOICE).<sup>7</sup> This is also reflected in the small changes in average compliance (see the Appendix, Section C.2, for a summary of average compliance by gender and treatment), which is equal to .374 in BASELINE, .313 in VOTE, .465 in VOTE+ and .434 in CHOICE, on which the lack of statistically significant treatment effects is also confirmed

<sup>7</sup>This result comes from a bootstrap version of the univariate Kolmogorov-Smirnov test. This modified test provides correct coverage even when the distributions being compared are not entirely continuous and, unlike the traditional Kolmogorov-Smirnov test, allows for ties (see Abadie, 2002; Sekhon, 2011).

Table 2: Share of partial compliers

		BASELINE	VOTE	VOTE+	CHOICE
Men	Share (%)	66.7	88.2	57.6	72.0
	Mean compliance	0.35	0.27	0.33	0.29
Women	Share (%)	62.9	87.8	81.0	96.0
	Mean compliance	0.36	0.26	0.35	0.31

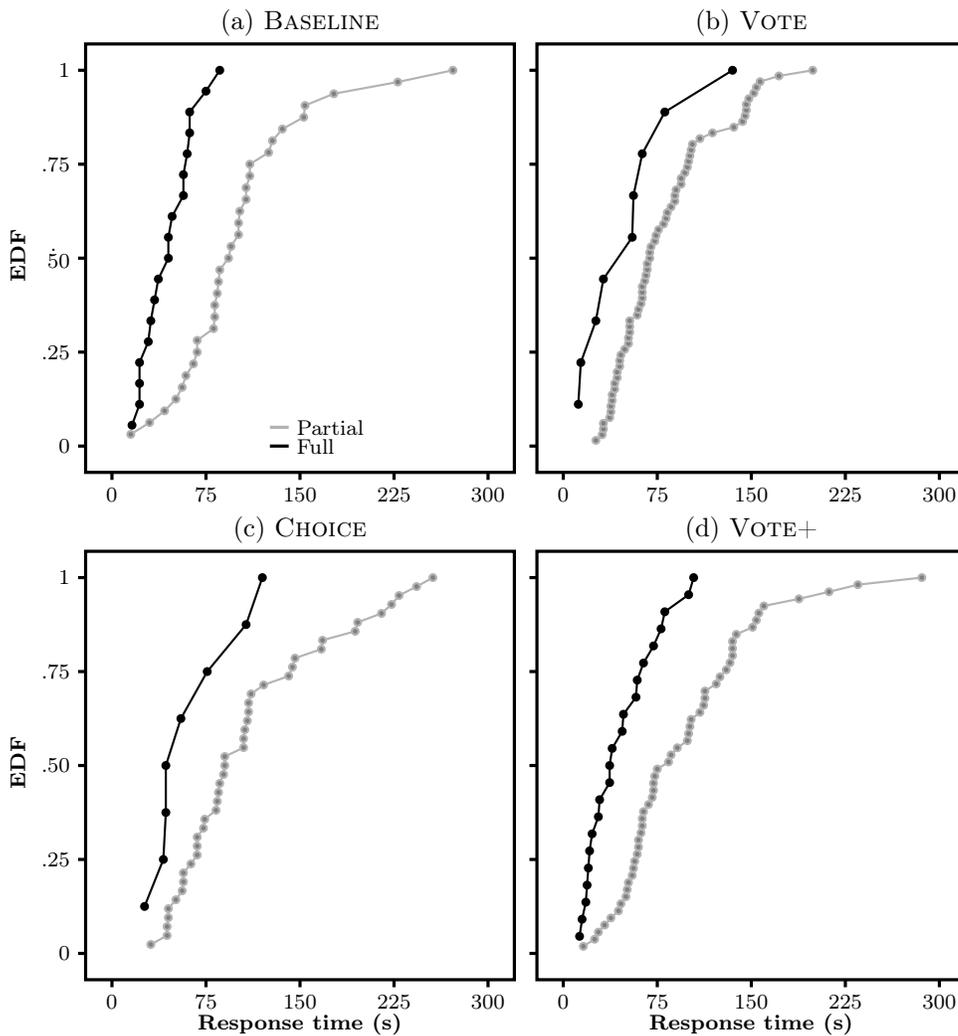
**Note.** For each treatment in *column*, and each gender *in row*, the table reports the share of subjects whose compliance for the preferred organization is partial (strictly between 0 and 1) and the average level of compliance within this subsample.

( $p = .565$ ,  $p = .592$ ,  $p = .315$ , t-tests). Last, the EDF shows that the main difference between the distributions is driven by the share of full compliers (who appear at the north-east in the graphs), with an increase in full compliance in the VOTE+ treatment to reach 27.2% from 13.3% in BASELINE — this change, however, is statistically non-significant ( $p = .244$ , one-sided proportion test).

The EDF of compliance among women, presented in Figure 2.b, suggests a different pattern. The EDF of compliance in BASELINE first order dominates the EDF of compliance in CHOICE ( $p = .002$ , KS bootstrap test) and VOTE ( $p = .036$ , KS bootstrap test). This is also true graphically about the EDF in VOTE+, but the difference is non significant ( $p = .241$ ). While we do not observe any change due to DDE among men, we thus even observe a decrease in compliance among women when they are given the choice, either individually or collectively, to favour their preferred option. This change results in large variations in mean compliance, from .592 in BASELINE, .380 in VOTE, .483 in VOTE+ to .307 in CHOICE. Classic t-tests are in line with KS bootstrap tests and indicate significant mean differences between BASELINE and both VOTE ( $p = .012$ ) and CHOICE ( $p < .001$ ). The mean compliance in BASELINE and VOTE+ are not significantly different ( $p = .212$ ). This decrease in compliance is again mainly driven by the share of full compliers, which decreases from 37.1% in BASELINE to 4% in CHOICE. Further KS bootstrap tests indicate that the EDF are not significantly different between men and women in CHOICE, VOTE and VOTE+ ( $p = .244$ ,  $p = .262$  and  $p = .741$ ), while the EDF of women first order dominates that of men in BASELINE. Women are therefore more compliant than men in BASELINE ( $p = .076$ , KS bootstrap test), due to a greater number of full compliers among women (37.1%) than men (13.3%,  $p = .089$ ).

To further investigate the reasons why we observe such a counter-veiling effect of direct democracy on compliance, we disaggregate compliance patterns in Table 2 based on the share of partial compliers (subjects who neither report their full income, nor fully evade) by gender and treatment. The second row in each panel reports the average compliance observed in this subsample. In all treatments and for both gender, the average compliance among partial compliers is fairly stable. The main change that occurs in the data is in terms of the share of subjects who fall in this category: with respect to BASELINE, partial lying is more widespread in both samples in VOTE ( $p = .081$  for men and  $p = .011$  for women, one-sided proportion test). The change is statistically

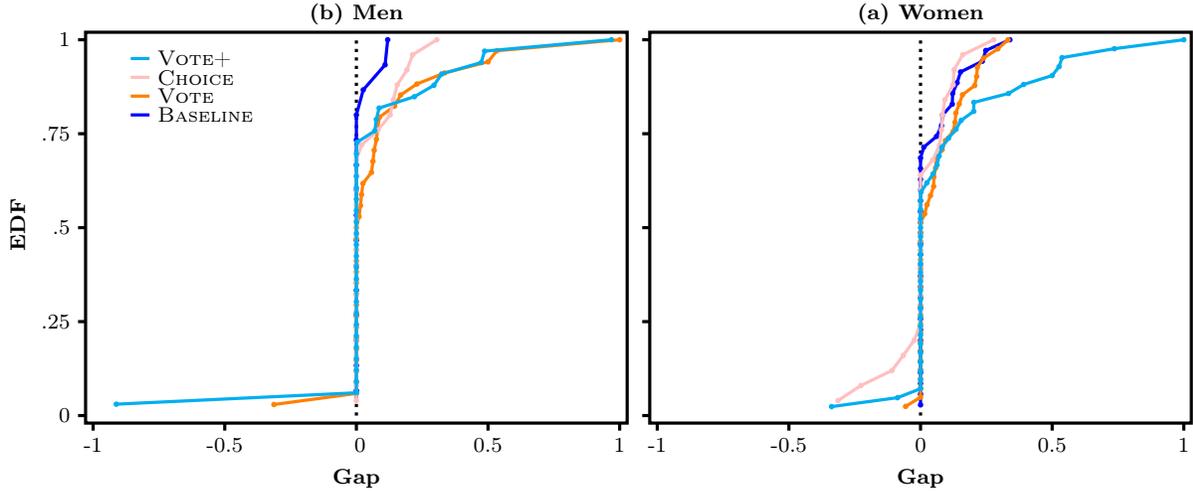
Figure 3: Empirical distribution functions (EDF) of response time by treatment



significant only for women in CHOICE ( $p = .004$ ) and VOTE+ ( $p = .064$ ;  $p = .500$  and  $p = .609$  for men). In a recent study, Jacquemet, Luchini, Malézieux, and Shogren (2020) put forward the idea that partial lying arises as the result of weak preferences towards honesty — subjects who partially comply do so because they struggle to decide between honesty and evasion. They show in particular that this interpretation is supported by the pattern of response times in the data, higher response times being clear indicators of the weakness of preferences.

Looking at response times, the pattern we observe in the data is very similar to the one documented in Jacquemet, Luchini, Malézieux, and Shogren (2020): partial compliers take much more time to make their decisions in BASELINE (the median response time is 84 seconds; see the Appendix, Section C.3, for a summary of median response times by treatment) than both full evaders

Figure 4: Compliance gap between most preferred and least preferred options



(33 seconds,  $p = .021$ , one-sided Wilcoxon rank sum test) and full compliers (48 seconds,  $p < .001$ , one-sided Wilcoxon rank sum test). Figure 3 provides a more detailed overview of this pattern. The figure presents the EDF of response times for partial compliers in all treatments, and provides a comparison with the EDF of response times among other subjects — who either fully comply, or fully evade. The EDF of response times among partial compliers first order dominates the EDF of other subjects in all treatments (first order stochastic dominance is statistically significant in BASELINE,  $p < .001$ ; VOTE,  $p = .066$ ; VOTE+,  $p < .001$ ; and CHOICE,  $p = .023$ ; KS bootstrap tests).

## 4.2 The commitment effect of direct democracy

We now turn to the commitment effect of direct democracy, by contrasting compliance across organizations on the basis of subjects' stated preference. To that end, we compute the gap between the two compliance decisions at the individual level between the preferred/chosen organization and the remaining one. Figure 4 displays the empirical distribution functions (EDF) of this gap by gender and treatment. The gaps observed in BASELINE show that a slight preference effect remains present in the data. While we observe that a large number of men exhibit no difference of compliance between the two options (80.0%, see Figure 4.a), the average value of the gap among the remaining men is .085.<sup>8</sup> Women who show no difference in compliance are less numerous than men (68.6%, see Figure 4.b). The difference in proportion between women and men is however not significant ( $p = .630$ ). When compliances are different, the average positive gap for women is low (.145). A bootstrap KS test indicates that the EDF of men and women are not significantly

<sup>8</sup>We do not consider subjects who comply more for their least preferred option as they are too few for a proper statistical analysis.

different ( $p = .281$ ). Importantly, when we pool men and women, a Wilcoxon signed rank test cannot reject the null of a zero gap location in BASELINE. Subjects are thus slightly more likely to comply for their most preferred option in BASELINE. Although this difference is not statistically significant, we investigate commitment based on the change in the gap between BASELINE and the other treatments. This allows us to compare the EDF of the gap across treatments, which is more conservative than only testing a zero location gap.

This means that a significant but small selection effect persists in our setting. Subjects are slightly more likely to comply for their most preferred option. We thus investigate commitment based on the change in the gap between the BASELINE and other treatments.

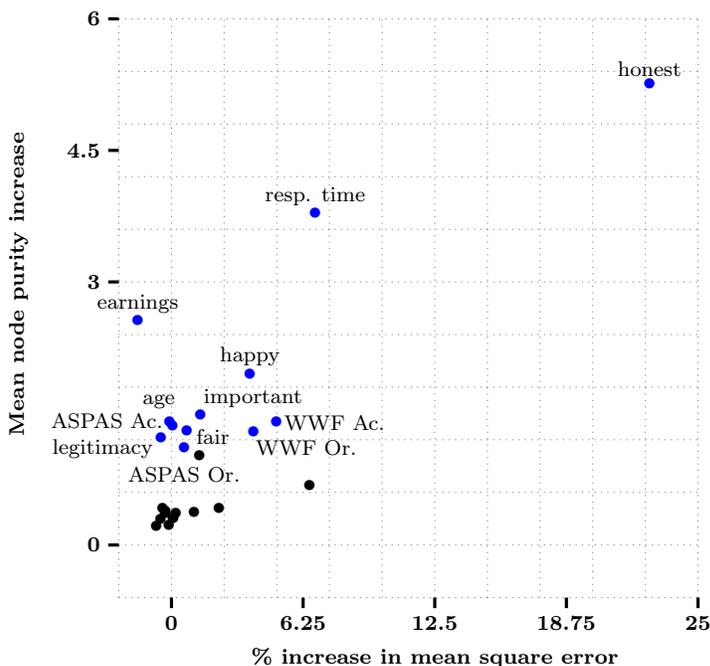
The EDF of men and women in CHOICE suggest that making an isolated choice before declaring one’s income does not induce more compliance: the distribution of the gap is statistically no different in CHOICE than in BASELINE ( $p = .350$ , KS bootstrap test on pooled data). In VOTE and VOTE+, by contrast, the EDF of the compliance gap first order dominates the EDF observed in BASELINE ( $p = .019$ ,  $p = 0.084$ ; KS bootstrap tests). Both men and women are less likely to exhibit a zero gap in these treatments: they are 47.1% (men) and 48.8% (women) in VOTE, and 66.7% (men) and 52.4% (women) in VOTE+. Among subjects whose gap is positive, the average gap is greater for men in VOTE (.201) and VOTE+ (.300) as compared to BASELINE, and greater for women in VOTE+ (.284). Our data thus support a slight commitment effect of direct democracy, that only occurs when the decision is made by vote, *i.e.*, when the preliminary behavior has a social dimension.

### 4.3 Sources of heterogeneity in compliance

We conclude this section by an investigation of the predictive power of all individual covariates presented in Table 1 as well as treatment variables. We do so by performing random forest algorithms on compliance behavior for the selected organization. Random forest is a machine learning method based on constructing decision trees that allows to rank the importance of individual variables in their ability to predict the outcome variable (Breiman, 2001). As compliance is a continuous variable, we perform a mean-prediction based on linear regressions. We compute  $p$ -values from one-sided binomial statistics that test whether the success of each predictor in the random forest task exceeds the theoretical success of a purely random predictor. We then rank the predictors according to the per cent increase in mean squared error and the mean node purity increase.

The resulting ranking of predictors is displayed in Figure 5, where only the individual characteristics endowed with a significant predictive power are explicitly labeled (all others appear in black on the figure). Twelve variables turn-out significant, that can be split into two groups. The first is made of variables with a very strong predictive power on compliance: self-reported honesty (*honest*), response time (*resp. time*) and, to a lesser extent, subject’s experimental earnings (*earn-*

Figure 5: Ranking of predictor variables to explain compliance behavior



ings). The second group is made of predictors that are significant, have a similar predictive power but matter less than variables in the first group: age, happiness (*happy*), Perceptions on ASPAS & WWF organizations and actions (*ASPAS Or.*, *WWF Or.*, *ASPAS Ac.* and *WWF Ac.*) and opinions about the experiment (*importance*, *legitimacy*, *fair*). None of the treatment variables are significant once those individual characteristics are taken into account. This confirms the absence of a DDE in our setting.

## 5 Conclusion

Previous experiments have found that offering participants to vote or choose the destination of their tax fund tends to increase tax compliance. The question we ask in this paper is whether this might be due to a “pure” direct democracy effect in tax evasion games. We investigate whether the increase in compliance is due to the *instrumental* value of democracy — the fact that more democracy leads to a better match between one’s own preferences and the way taxes are spent — or to some *intrinsic* value — democracy in itself leading to a change in compliance behavior (Pattanaik and Xu, 1998). Several arguments point into that direction: the process of tax collection can appear as more legitimate to tax payers, more agency can lead to less reactance against taxes, a collective decision setting can induce some social coordination, and the preliminary decision can also commit subjects to more honest behavior.

In order to isolate the democracy effect from the preference effect, we rely on a novel experimental design that mitigates self-selection by giving subjects the opportunity to select the organization that they prefer, but between very similar options. Subjects are therefore given a choice, either collectively or individually, but the preference between one or the other organization does not really matter. The two organizations are providing exactly the same social/ecological benefits and operate in similar ways by collecting donations through an adoption process and issue adoption certificates. If there is an effect of a preliminary choice on compliance behavior, it would therefore have no instrumental value. We moreover elicit compliance towards both organizations before the organization actually receiving the funds is selected, allowing to measure the gap in compliance induced by the commitment effect of the choice offered to subjects.

Our results are clear: we observe no effect of prior choice on compliance behavior when democracy has no instrumental value. Our results even suggest a potential opposite result — a decrease in full compliance in favour of partial compliance, in particular for women. Based on the compliance gap, we confirm a slight commitment effect of democracy as the gap in compliance in favour of the selected organization is significantly higher when selection results from an active choice on the part of subjects. This, however, is not enough to overcome the drop in the level of compliance and results in a negative direct democracy effect.

While this result echoes some recent evidence in the literature (see in particular Dal Bó, Foster, and Kamei, 2019), the obvious open avenue for further research is to assess the robustness of these results to those features of our design that were implemented so as to rule out the instrumental value of democracy. First, we used the strategy method and elicited compliance for both organizations. Although treatment effects observed using the strategy method have been shown to replicate well treatment effects obtained based on direct-responses (see Brandts and Charness, 2011, for a survey), it might be that asking for two compliance decisions led subjects to balance their decisions more than they would have done with direct responses. Second, following the literature on tax evasion games, we focus on the effect of democracy on a very specific feature of the tax system — the use of the collected funds. Based on the existing literature using other kinds of outcome behavior, it might well be that a direct democracy effect applies to the choice of other features of the design of tax policies — like, *e.g.*, the tax rate schedule or the audit mechanism.

Last, and more importantly, our design elicits the democracy effect in a context in which choices do not really matter, leading to the conclusion that choosing among instrumentally equivalent alternatives actually weakens honesty. This could be interpreted as a natural consequence of the fact that, in our setting, there is no added self-reliance when a choice is offered as compared to the random selection of organizations. Since self-reliance has been shown to be the main component of the intrinsic value of decision rights (Ferreira, Hanaki, and Tarrow, 2020), this feature would explain our results if subjects perceive this “fake agency” and react negatively to

it in the income declaration task. This result has obvious implications in the field. Democracy is not a tax compliance improving policy on its own: people must be given a real choice if one wants to improve compliance by adding more democracy to tax collection mechanisms. Without instrumental value, our results suggest that democracy alone is not enough to mitigate tax evasion. Whether offering subjects a choice associated with more instrumental value would also foster some intrinsic value of democracy remains an open question, that is next on the research agenda.

## References

- ABADIE, A. (2002): “Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Model,” *Journal of the American Statistical Association*, 97(457), 284–292.
- ALM, J., T. CHERRY, M. JONES, AND M. MCKEE (2008): “Encouraging Filing via Tax Credits and Social Safety Nets,” *The IRS Research Bulletin: Proceedings of the 2008 IRS Research Conference*, pp. 43–57.
- ALM, J., B. R. JACKSON, AND M. MCKEE (1993): “Fiscal Exchange, Collective Decision Institutions, and Tax Compliance,” *Journal of Economic Behavior & Organization*, 22(3), 285–303.
- ALM, J., G. H. MCCLELLAND, AND W. D. SCHULZE (1999): “Changing the Social Norm of Tax Compliance by Voting,” *Kyklos*, 52(2), 141–71.
- BARDHAN, P. (2000): “Irrigation and Cooperation: An Empirical Analysis of 48 Irrigation Communities in South India,” *Economic Development and Cultural Change*, 48(4), 847–865.
- BEM, D. (1972): “Self-Perception Theory,” in *Advances in Experimental Social Psychology*, ed. by L. Berkowitz, vol. 6. Academic Press, New York.
- BISCHOFF, I. (2007): “Institutional Choice versus Communication in Social Dilemmas? An Experimental Approach,” *Journal of Economic Behavior & Organization*, 62(1), 20–36.
- BODNER, R., D. PRELEC, I. BROCCAS, AND J. D. CARILLO (2003): “Self-Signaling and Diagnostic Utility in Everyday Decision Making,” in *The Psychology of Economic Decisions*, vol. 1, pp. 105–123. Oxford University Press, Oxford (UK).
- BOGLIACINO, F., L. JIMÉNEZ, G. GRIMALDA, ET AL. (2015): “Consultative Democracy and Trust,” Discussion paper, UN-RCE-CID.
- BRANDTS, J., AND G. CHARNES (2011): “The Strategy versus the Direct-Response Method: A First Survey of Experimental Comparisons,” *Experimental Economics*, 14(3), 375–398.
- BREHM, J. (1966): *A Theory of Psychological Reactance*. Academic Press, New York.
- BREIMAN, L. (2001): “Random Forests,” *Machine learning*, 45(1), 5–32.
- BROCKMANN, H., P. GENSCHEL, AND L. SEELKOPF (2016): “Happy Taxation: Increasing Tax Compliance through Positive Rewards?,” *Journal of Public Policy*, 36(3), 381–406.

- CIALDINI, R. B. (2007): *Influence: The Psychology of Persuasion*, vol. 55. Collins New York.
- CINYABUGUMA, M., T. PAGE, AND L. PUTTERMAN (2005): “Cooperation under the Threat of Expulsion in a Public Goods Experiment,” *Journal of Public Economics*, 89(8), 1421–1435.
- CZAP, H. J., N. V. CZAP, AND E. BONAKDARIAN (2010): “Walk the Talk? The Effect of Voting and Excludability in Public Goods Experiments,” *Economics Research International*, 2010.
- DAL BÓ, P. (2014): “Experimental Evidence on the Workings of Democratic Institutions,” in *Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North*, pp. 266–288. Cambridge: Cambridge University Press.
- DAL BÓ, P., A. FOSTER, AND K. KAMEI (2019): “The Democracy Effect: A Weights-Based Identification Strategy,” *NBER WP*, 25724.
- DAL BÓ, P., A. FOSTER, AND L. PUTTERMAN (2010): “Institutions and Behavior: Experimental Evidence on the Effects of Democracy,” *American Economic Review*, 100(5), 2205–2229.
- DANNENBERG, A., C. HAITA-FALAH, AND S. ZITZELSBERGER (2020): “Voting on the Threat of Exclusion in a Public Goods Experiment,” *Experimental Economics*, 23(1), 84–109.
- DEUTSCH, M., AND H. B. GERARD (1955): “A Study of Normative and Informational Social Influences upon Individual Judgment,” *The journal of abnormal and social psychology*, 51(3), 629.
- EHRHART, K.-M., AND C. FEIGE (2014): “Voting and Transfer Payments in a Threshold Public Goods Game,” *KIT Working Paper Series in Economics*.
- EHRHART, K.-M., C. FEIGE, AND J. KRÄMER (2015): “Voting on Contributions to a Threshold Public Goods Game: An Experimental Investigation,” *KIT Working Paper Series in Economics*, (60).
- FELD, L. P., AND J. R. TYRAN (2002): “Tax Evasion and Voting: An Experimental Analysis,” *Kyklos*, 55(2), 197–221.
- FERREIRA, J. V., N. HANAKI, AND B. TARROUX (2020): “On the Roots of the Intrinsic Value of Decision Rights: Experimental Evidence,” *Games and Economic Behavior*, 119, 110–122.
- FESTINGER, L. (1957): *A Theory of Cognitive Dissonance*. Stanford University Press, CA.
- GALLIER, C., M. KESTERNICH, AND B. STURM (2014): “Voting for Burden Sharing Rules in Public Goods Games,” .
- GREINER, B. (2015): “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE,” *Journal of the Economic Science Association*, 1(1), 114–125.
- GROSSMAN, G., AND D. BALDASSARRI (2012): “The Impact of Elections on Cooperation: Evidence from a Lab-in-the-Field Experiment in Uganda,” *American Journal of Political Science*, 56(4), 964–985.
- ICHNIOWSKI, C., K. SHAW, AND G. PRENNUSHI (1997): “The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines,” *American Economic Review*, 87(3), 291–313.

- JACQUEMET, N., S. LUCHINI, A. MALÉZIEUX, AND J. SHOGREN (2020): “Who’ll Stop Lying under Oath? Experimental Evidence from Tax Evasion Games,” *European Economic Review*, 20, 103369.
- JOULE, R., AND J. BEAUVOIS (1998): *La Soumission Librement Consentie*. Presses Universitaires de France, Paris.
- KAMEI, K. (2016): “Democracy and Resilient Pro-Social Behavioral Change: An Experimental Study,” *Social Choice and Welfare*, 47(2), 359–378.
- KIESLER, C. (1971): *The Psychology of Commitment. Experiments Linking Behavior to Belief*. Academic Press, New York.
- KIESLER, C. A., AND J. SAKUMURA (1966): “A Test of a Model for Commitment,” *Journal of Personality and Social Psychology*, 3(3), 349–353.
- KROLL, S., T. L. CHERRY, AND J. F. SHOGREN (2007): “Voting, Punishment, and Public Goods,” *Economic Inquiry*, 45(3), 557–570.
- LAMBERTON, C. P., J.-E. DE NEVE, AND M. I. NORTON (2018): “The Power of Voice in Stimulating Morality: Eliciting Taxpayer Preferences Increases Tax Compliance,” *Journal of Consumer Psychology*, 28(2), 310–328.
- LE SAGE, S., AND E. VAN DER HEIJDEN (2015): “The Effect of Voting on Contributions in a Public Goods Game,” *CentER Discussion Paper*, (2015-039).
- LOCKE, E. A., AND G. P. LATHAM (2002): “Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey,” *American Psychologist*, 57(9), 705–717.
- MARGREITER, M., M. SUTTER, AND D. DITTRICH (2005): “Individual and Collective Choice and Voting in Common Pool Resource Problem with Heterogeneous Actors,” *Environmental and Resource Economics*, 32(2), 241–271.
- MARKUSSEN, T., L. PUTTERMAN, AND J.-R. TYRAN (2013): “Self-Organization for Collective Action: An Experimental Study of Voting on Sanction Regimes,” *Review of Economic Studies*, 81, 301–324.
- MATSUSAKA, J. G. (2005): “Direct Democracy Works,” *Journal of Economic Perspectives*, 19(2), 185–206.
- MESSER, K. D., J. F. SUTER, AND J. YAN (2013): “Context Effects in a Negatively Framed Social Dilemma Experiment,” *Environmental & Resource Economics*, 55(3), 387–405.
- MESSER, K. D., H. ZARGHAMEE, H. M. KAISER, AND W. D. SCHULZE (2007): “New Hope for the Voluntary Contributions Mechanism: The Effects of Context,” *Journal of Public Economics*, 91(9), 1783–1799.
- PATTANAİK, P. K., AND Y. XU (1998): “On Preference and Freedom,” *Theory and Decision*, 44(2), 173–198.
- PUTTERMAN, L., J.-R. TYRAN, AND K. KAMEI (2011): “Public Goods and Voting on Formal Sanction Schemes,” *Journal of Public Economics*, 95(9), 1213–1222.

- RAUCHDOBLER, J., R. SAUSGRUBER, AND J.-R. TYRAN (2010): “Voting on Thresholds for Public Goods: Experimental Evidence,” *FinanzArchiv: Public Finance Analysis*, 66(1), 34–64.
- SEKHON, J. (2011): “Multivariate and Propensity Score Matching Software with Automated Balance Optimization,” *Journal of Statistical Software*, 42(7), 1–52.
- SUTTER, M., S. HAIGNER, AND M. G. KOCHER (2010): “Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations,” *Review of Economic Studies*, 77(4), 1540–1566.
- TYRAN, J.-R., AND L. P. FELD (2006): “Achieving Compliance When Legal Sanctions Are Non-Deterrent,” *Scandinavian Journal of Economics*, 108(1), 135–156.
- VANBERG, C. (2010): “Voting on a Sharing Norm in a Dictator Game,” *Journal of Economic Psychology*, 31(3), 285–292.
- WAHL, I., S. MUEHLBACHER, AND E. KIRCHLER (2010): “The Impact of Voting on Tax Payments,” *Kyklos*, 63(1), 144–158.
- WALKER, J. M., R. GARDNER, A. HERR, AND E. OSTROM (2000): “Collective Choice in the Commons: Experimental Results on Proposed Allocation Rules and Votes,” *Economic Journal*, 110(460), 212–234.

## Appendix

### A Voting experiments: an overview

Table A provides a summary of empirical studies measuring the DDE, *i.e.* featuring a comparison between a vote and a baseline condition.

#### A.1 Existing empirical evidence on DDE in tax evasion games

Alm, Jackson, and McKee (1993) are, to the best of our knowledge, the first to introduce such a choice before the income reporting stage. In this experiment, participants vote on one of two organizations that can receive their group tax fund. In comparison with the treatment in which the same organization is rather imposed by the experimenter, the compliance raises from 33.7% to 45.2%. A large part of this effect is due to social coordination: if the same organization is chosen by vote, but the outcome of the vote is more balanced (because the alternative organization is more favored in the sample), compliance falls to 41.1%, and to 33.2% if only the winning organization, rather than the count of the vote, is announced before the declaration stage.

The three other studies share the same feature of offering participants the possibility to express their opinion about the use of collected funds before they report their income. In Wahl, Muehlbacher, and Kirchler (2010), participants vote on two hypothetical scenarios about the recipient of collected taxes: either themselves or a third party. In Brockmann, Genschel, and Seelkopf (2016), participants can choose where to earmark their tax fund between three alternatives: recreational facilities on Campus and/or for the financial support of needy students and/or for the invitation of VIPs to campus. Last, in Lambertson, De Neve, and Norton (2018), participants can choose between three categories that will benefit to individuals taking part in future experiments

Table A: Literature review on voting experiments

Authors	Game	$n$	Nb. of options	Options offered	Repeated voting	Voting on	Sources mentioned
Vanberg	DG	46	10	$\{0, 1, \dots, 10\}$	No	Amount sent	Social
Bogliacino, Jiménez, Grimalda, et al.	TG	62	2	Keep all / Share equally	No (but switch roles)	The preferred course of action for trustors and trustees	Social
Margreiter, Sutter, and Dittrich	CPR	6	6	Each participants' proposition	Yes (10)	Vector of extraction	None
Walker, Gardner, Herr, and Ostrom	CPR	7	7	Each participants' proposition	Yes (10)	Allocation rule to the pool	Social
Bischoff	CPR	5	10 and 6	10 level of extraction / 6 probability of detection	Yes (7)	Recommended level of extraction and probability of detection	None
Dal Bó, Foster, and Putterman	PD	4	2	Modify / Not modify	No	Whether to modify the payoffs	Social
Feld and Tyran	PGG	3	2	Yes / No	No	Mild fine punishing free-riders	Social, Commitment, Legitimacy
Cinyabuguma, Page, and Putterman	PGG	16	2	Expel / Not expel	Yes (15)	Members to expel from the group	None
Tyran and Feld	PGG	3	2	Yes / No	No	Mild fine punishing free-riders	Social, Commitment
Kroll, Cherry, and Shogren	PGG	5	5	Each participants' proposition	Yes (20)	Minimal contribution	Social
Czap, Czap, and Bonakdarian	PGG	5	2	Project 1 / 2	Yes (20)	Two equivalent projects (with different provision point and rates of return)	Social, Commitment
Rauchdobler, Sausgruber, and Tyran	PGG	3	2	High / Low	No	Threshold that participants must reach to have their share of the re-fund	Social, Commitment
Ehrhart and Feige	PGG	5	5	Each participants' proposition	No	Top-down treatment: binding total contribution to the PGG and share of contribution among players. Bottom-up treatment: contribution vector	None
Ehrhart, Feige, and Krämer	PGG	4	4	Each participants' proposition	No	Contribution vector and transfer vector	None

Table A Continued

Authors	Game	$n$	Nb. of options	Options offered	Repeated voting	Voting on	Sources mentioned
Gallier, Kesternich, and Sturm	PGG	3	3	Equal contribution / equal payoffs / contribution proportional to endowment	No	Rule-based contribution scheme	None
Grossman and Bal-dassarri	PGG	8 to 12	8 to 12	Each participants	No	Individual who will be a monitor	Social, Legitimacy
Le Sage and Van der Heijden	PGG	3	2	Yes / No	Yes (10)	Acceptation of the group total contribution	None
Markussen, Putterman, and Tyran	PGG	5	2	No sanction / informal sanction / formal sanction	Yes (6)	Sanctioning institution	None
Messer, Zarghamee, Kaiser, and Schulze	PGG	7	2	VCM / Private lottery	Yes (10)	Participation to the VCM	Social
Messer, Suter, and Yan	PGG	7	2	VCM/ Private lottery	Yes (20)	Participation to the VCM	None
Putterman, Tyran, and Kamei	PGG	5	2, 20, 20	Public / Private; $\{0, 1, \dots, 20\}$ ; $\{0, 1, \dots, 20\}$	Yes (4)	(a) whether contributing to the public or to the private account is subject to a penalty, (b) what level of contribution is exempt from penalty, and (c) maximum amount of the penalty	None
Sutter, Haigner, and Kocher	PGG	4	3	Standard VCM / With reward / With punishment	No	Institution promoting cooperation	Social
Kamei	PGG (*2)	2	2	Yes / No	No	Mild fine punishing free-riders	None
Dannenberg, Haita-Falah, and Zitzelsberger	PGG	5, 4, 3	2 / $n$	Standard PGG / PGG with exclusion and player to be excluded	Yes ( $\geq 4$ )	Type of PGG to play and player to be excluded	None
Alm, Jackson, and McKeel	TEG	15	2	Two organizations	Yes (25)	Recipient of the taxes collected	Social, Commitment
Alm, McClelland, and Schulze	TEG	11	2	High / Low	Yes (30)	Tax rate, fine rate, probability of detection	Social
Wahl, Muehlbacher, and Kirchner	TEG	3	2	Two choices	Yes (3)	Experiment 1: Two equivalent probabilities of return rates. Experiment 2: The use of the taxes (two different scenarios)	Social, Legitimacy

Table A Continued

Authors	Game	$n$	Nb. of options	Options offered	Repeated voting	Voting on	Sources mentioned
Lamberton, De Neve, and Norton	TEG	1	3	Three choices	No	Use of taxes	Tax-payer agency, legitimacy
Brockmann, Genschel, and Seelkopf	TEG	1	3	Three choices	Yes (6)	Use of taxes	Tax-payer agency, legitimacy

**Note.** Literature review of lab. experiments featuring a behavioral measure of the effect of a vote treatment. The second column reports the type of game implemented — CPR: Common Pool Resource; DG: Dictator Game; PD: Prisoner's Dilemma; PGG: Public Good Game; TEG: Tax Evasion Game; TG: Trust Game. The third column reports the size of the voting group,  $n$ . The next four columns describe the vote treatment: the number of choices submitted to the vote, the nature of those choices, whether the vote is repeated and, last, what feature of the environment the vote is about. A fine is 'mild' if it is not enough to counteract the monetary incentives to behave selfishly. The last column reports whether ('None') the main mechanisms lying behind a DDE are explicitly discussed in the paper.

Table B: Estimations of a DDE based on individual data from existing studies

	(1)		(2)		(3)	
	<b>Compliance</b>		<b>Extensive margin</b>		<b>Intensive margin</b>	
	(Tobit)		(Probit)		(Tobit)	
	Coef.	Marg. Effect	Coef.	Marg. Effect	Coef.	Marg. Effect
Direct Democracy	0.13*	0.07*	0.10	0.03	0.11*	0.07*
	(.079)	(.042)	(.147)	(.046)	(.060)	(.042)
Constant	0.62***		-0.48***		0.27***	
	(.096)		(.175)		(.063)	
$\sigma$	0.64***				0.41***	
	(.035)				(.016)	
Nb of obs.	1,483		2,714		1,957	
Nb of clusters	145		347		223	

**Note.** Regressions of the compliance rate observed in Wahl, Muehlbacher, and Kirchler (2010); Brockmann, Genschel, and Seelkopf (2016); Lambertson, De Neve, and Norton (2018) on the direct democracy treatment variable. Model (1) provides the results of a Tobit regression of compliance observed in pooled data; Model (2) provides Probit estimates on the probability to fully comply or not; Model (3) provides Tobit estimates on the compliance observed among the subsample of subjects who do not fully comply. All models include study fixed effects, round fixed-effects, and cluster the standard errors at the subject level. **Legend.** Significance levels: \*: 10%, \*\*: 5%, \*\*\*: 1%

(beverages, snacks or enhanced incentives). We were able to retrieve the individual data of these three studies — unfortunately, we were not able to retrieve the original data from Alm et al. study, and only average effects are available. We provide evidence on the DDE elicited in these studies thanks to a normalization of tax compliance measured as the share of the income that is reported. The pooled data results in a sample of 2,714 observations from 347 different subjects, among which 180 subjects (1,475 observations) participate in the baseline.

Table B reports the results of several regression models on this pooled data. The main variable of interest, *Direct Democracy*, is a dummy variable equal to 1 when compliance is decided after an explicit choice about the use of collected funds (the treatment is implemented between-subjects in all three studies). In all regressions, we include study fixed effects to account for design differences leading to changes in the average level of compliance in the baseline. The tax reporting task is repeated in all these studies. We thus account for possible learning over time by including round fixed effects, and account for multiple observations arising from the same subject by clustering the standard errors at the individual level. Model (1) reports the results of a Tobit regression on the compliance variable (so as to take into account the censoring of the outcome variable). The effect of direct democracy is only significant at the 10% level, and, more importantly, the marginal effect is rather small with an expected increase of 7% in compliance rate. In Models (2) and (3) we distinguish the extensive margin (the decision to either fully comply, or not) and the intensive margin (the compliance rate conditional on not complying fully). This decomposition shows that the weak DDE observed on pooled data is driven by a small increase in compliance among subjects who do not fully comply.

## B Experimental instructions

We provide below an English translation of the original instructions in French. The instructions for each part are only distributed and read aloud to subjects once the previous part is over.

### B.1 Baseline experiment

You are taking part in an experiment in which you might win some money. The amount of money that you will win depends on your decisions. **This experiment is a fiscal simulation.**

#### INSTRUCTIONS OF THE EXPERIMENT

The experiment will take place in three stages. The instructions for each stage will be given to you at the beginning of that stage. The currency used in the experiment is ECU (for *Experimental Currency Unit*). Its value will be described below.

#### HOW WILL YOU MAKE YOUR DECISIONS?

You will make your decisions thanks to the computer in front of you. All the informations that will be useful to make your decisions will appear on the screen. To take your decisions, click on the buttons on the screen once you acknowledge these informations.

#### CALCULATION AND PAYMENT OF YOUR EARNINGS

Your earnings during the experiment is expressed in ECU. These earnings in ECU will be converted into euros at the rate of  $25 \text{ ECU} = 1 \text{ euro}$ . You will be paid the sum corresponding to this total individually, in cash, at the end of the experiment.

For scientific reasons which everyone will understand, **it is essential that you do not talk together during the experiment.** Unfortunately, any participant who fails to respect this rule will be asked to leave the room without any possibility of collecting their earnings.

**Thank you for your participation.**

\*\*\*\*\*

## PROCEDURES FOR THE FIRST PART

In the first stage, you will earn an income in ECU participating to a series of 5 tasks. **The sum that you will earned corresponds to your annual income.**

Each of the tasks take place as follow: on your screen appear two grids of 3x3 boxes, composed of 9 squares. When you click on Start, on the right hand grid appear digits between 1 and 9 in a random order. The goal is to click on the digits in an increasing order, thanks to the mouse. Thus, you have to click first on the digit 1, then on the digit 2, then on the 3 etc. When you click on 9, the task is finished. When you click on each digits, those have to appear in the left hand grid of the screen to be validated. The chrono starts when you click on Start. The time elapsing while you accomplish the task appears on the bottom of the screen: the faster you get to 9, the higher the sum that you will win.

Indeed, you are given at the beginning of this stage a sum equal to 150 ECU. This amount diminished by 13 ECU at each seconds.

Your earning in ECU at the end of each tasks is equal to

$$150 \text{ ECU} - (\text{your time} * 13).$$

Note that you cannot have a negative earning: **if your score at a task becomes negative**, it will be fixed at 0.

At the end of this stage, we will compute **the total of your ECU earned through the 5 tasks**. The longer you spend to do the five tasks, the lower will your gain be for this stage. At the end of this stage, a message will tell you the sum in ECU that you have earned through the five tasks for this stage. Click on OK once you learned about this information.

It is very important that you understand the rules of the experiment perfectly. If you have any questions, please raise your hand, somebody will come and answer them. Thank you for following these instructions.

\*\*\*\*\*

PLEASE COMPLETE THE FOLLOWING QUESTIONNAIRE:

1. The first stage has \_\_\_\_\_ tasks.

2. In the task, the goal is to click on the squares in the increasing order, from 1 to 9.

YES  NO

3. My gain for this stage depends on the time spend for the execution of the tasks: the quicker I am, the higher my gain.

YES  NO

4. My gain for this stage correspond to my annual income.

YES  NO

5. The earnings in Euros that I will be paid at the end of the experiment depend on my decisions and on my performance during this stage.

YES  NO

**If any of these points remain unclear, do not hesitate to ask questions.**

\*\*\*\*\*

## PROCEDURES FOR THE SECOND PART

[*The instructions for this part are treatment specific, see Section B.2.*]

\*\*\*\*\*

PLEASE COMPLETE THE FOLLOWING QUESTIONNAIRE:

1. I determine the unique organization, WWF or ASPAS, that will get my tax.  
 YES  NO
2. I declare two times my income, once if WWF is selected, and once if ASPAS is selected.  
 YES  NO
3. The sum of income corresponds at the amount of my earnings in ECU from the first stage  
 YES  NO
4. The amount of tax applied to my income will be deducted from my earnings in the experiment.  
 YES  NO
5. The sum of tax collected will be entirely given to WWF or to ASPAS to support its actions for environmental protection.  
 YES  NO

**If any of these points remain unclear, do not hesitate to ask questions.**

\*\*\*\*\*

## PROCEDURES FOR THE THIRD PART

In this stage, we will ask you to answer at a certain number of questions. **All these information, along your earnings will remain strictly confidential.**

The first questions will allow us to get your reaction about this experiment. The second questions will help us to know you better (your age, professional activity, etc.). In the following questionnaires, we will ask you to check the boxes that best correspond to your opinion, among the different propositions.

Take your time to answer these questions. In order to thank you for the time dedicated to these questionnaires, a fixed amount of 10 euros will be added to your gains at the end of this stage.

### B.2 Treatment specific instructions for the second part

The instructions for the second part are treatment specific. We provide below the text common to all treatments, along with treatment-specific instructions in italics and in brackets.

## PROCEDURES FOR THE SECOND PART

A few information are needed to understand this stage properly.

The *World Wide Fund for Nature*, better known as the WWF, is an international non-governmental organization for the protection of nature and the environment, strongly involved in sustainable development. This organization is based in Gland, in Switzerland, and has more than 4.7 million members throughout the world and has an operational network in 96 countries. The aim of this private organization is to protect fauna, habitats, and nature in general, and to this end it collects funds for one-off actions. Its main activities are monitoring the application of international regulations, restoring damaged natural spaces and training.

The *Organization for the Protection of Wild Animals*, better known as ASPAS, is a french non-governmental organization for the protection of wildlife and the natural heritage. It is based in Crest, in France. The aim of this private organization is to protect fauna, preservation of natural heritage, defense of nature users' rights, and to this end it collects funds for one-off actions. Its main activities are to run informations campaigns to mobilize public opinion, to interpellate elected and to launch petitions to enforce environmental law.

To finance their environmental protection of environment activities, these two organizations propose to individuals to participate in their operations' fundings. The funds that are collected in this way allow them to continue its efforts in terms of environmental protection and the conservation of bio-diversity.

**During this stage, we ask you to declare the amount of your income.** Your income corresponds to the sum that you have earned in the first stage. The amount of income that you report will be taxed at a 35% rate. The amount of taxes collected will be collected from your income and given to the organization that you will select. Your role is thus to determine between WWF and ASPAS which organization will be the beneficiary of the amount of income collected thanks to the tax.

[CHOICE/VOTE: *This stage is implemented in two phases. In the first phase, we ask you to select this organization to which the tax collected will be given. In a second phase, we ask you to declare your income.*]

[VOTE+: *This stage is implemented in three phases. In the first phase, you will discuss as a group about the organization that you would like to see become the beneficiary of the tax collected. In the second phase, we ask you to select this organization to which the tax collected will be given. In a third phase, we ask you to declare your income.*]

Procedure of the first phase:

[CHOICE: *In this phase, you have to choose the organization that will get the amount of taxes collected. To do this, you have to choose between two options: the choice A and the choice B. These two choices have a part of chance. Once you have made your choice, a draw will determine the organization that will be effectively selected in the following way:*

- *If you choose Choice A: ASPAS is selected with one chance out of three (33,3%) and WWF is selected with two chances out of three (66,6%).*
- *If you choose Choice B: WWF is selected with one chance out of three (33,3%) and ASPAS is selected with two chances out of three (66,6%).]*

[VOTE: *At the beginning of this phase, groups of three participants are formed randomly: yourself and two other participants. Inside each group, a vote determines the organization that will get collected taxes. To do this, each member of the group has to choose to vote for WWF or ASPAS. In each group, the organization selected is the one that get at least two votes. This organization will receive the sum of taxes given by the three group members in the second phase of this stage.*]

[VOTE+: *At the beginning of this phase, groups of three participants are formed randomly: yourself and two other participants. Inside each group, a vote determines the organization that will get collected taxes. In the first phase, you have the possibility to discuss with the other members of your group, in order to agree on the organization which will receive your tax. However, there are limits to the messages you can send: you must not send a message that allows you to locate yourself in the room, or give information that would identify you or your partners. There can also be no threats or jokes. If you do not respect these rules, you will be excluded from the experience and from any payment. This discussion opportunity will last 3 minutes. ]*

Procedure of the second phase:

[CHOICE: *The organization that have been selected will be announced uniquely at the end of this stage. In this phase, we ask you to declare your income in two different situations: first in the hypothesis where WWF is the tax beneficiary ; and then in the hypothesis where ASPAS is the tax beneficiary.*]

[VOTE: *The organization that have been selected will be announced uniquely at the end of this stage. In this phase, we ask you to declare your income in two different situations: first in the hypothesis where WWF is the tax beneficiary (it means that at least two participants of your group have voted for WWF) ; and then in the hypothesis where ASPAS is the tax beneficiary (it means that at least two participants of your group have voted for ASPAS).*]

[VOTE+: *In this phase, the vote actually takes place. To do this, each member of the group has to choose to vote for WWF or ASPAS. In each group, the organization selected is the one*

that get at least two votes. This organization will receive the sum of taxes given by the three group members in the third phase of this stage.]

[VOTE+: Procedure of the third phase:

The organization that have been selected will be announced uniquely at the end of this stage. In this phase, we ask you to declare your income in two different situations: first in the hypothesis where WWF is the tax beneficiary (it means that at least two participants of your group have voted for WWF) ; and then in the hypothesis where ASPAS is the tax beneficiary (it means that at least two participants of your group have voted for ASPAS).]

The amount of income that you report will be taxed at a 35% rate. To declare your income, move the slider in the gray zone till the amount of income that you wish to declare appear. Once you made your choice click OK to validate.

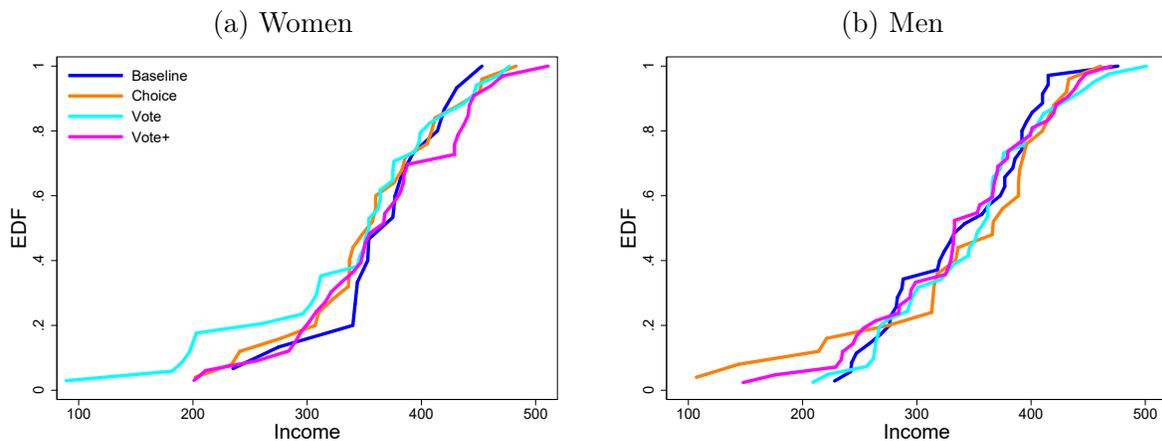
Your gain at the end of this stage corresponds to:

**Your income - (the income that you declare for the chosen organization \* 0.35)**

The total sum of money collected thanks to the tax correspond to your choice of declaration of the organization that have been selected in the first phase. This sum will be given by us to the chosen organization, to support its actions of environmental protection. It is either WWF or ASPAS, according to the chosen organization in the first phase. This donation to WWF or ASPAS will be certified by official certificates. For your information, these certificates will be send to you by email, at the latest in 3 weeks from today.

## C Additional data

### C.1 Distribution of earned income, by treatment and gender



## C.2 Average compliance by gender and treatment

	BASELINE	VOTE	CHOICE	VOTE+
Mean compliance (Men)	0.36	0.26	0.40	0.43
Mean compliance (Women)	0.56	0.34	0.30	0.42
Men				
Full evaders (%)	20.0	8.8	8.0	18.2
Partial compliers (Mean)	0.35	0.27	0.29	0.33
Full compliers (%)	13.3	2.9	20.0	24.2
Women				
Full evaders (%)	2.9	0.0	4.0	4.8
Partial compliers (Mean)	0.36	0.26	0.31	0.35
Full compliers (%)	34.3	12.2	0.0	14.3

## C.3 Median response times by treatment (in seconds)

	Full evaders	Partial compliers	Full compliers
BASELINE	43.0	94.0	45.0
VOTE	55.0	69.0	47.5
VOTE+	20.5	84.0	58.5
CHOICE	55.0	90.0	43.0