An experimental test of two policies to increase donations to public projects

Raúl López-Pérez a,*, Aldo Ramírez-Zamudio b

a Institute of Public Goods and Policies, (IPP), Spanish National Research Council, (CSIC), C/Abasanz 26-28, 28037, Madrid, Spain
b Center for Economics, Banking and Finance Studies, Department of Economics, Universidad De Lima, Av. Javier Prado Este S/N - Santiago De Surco, Lima 33, Peru

A R T I C L E   I N F O

Article history:
Received 10 May 2019
Received in revised form 30 January 2020
Accepted 1 February 2020
Available online 3 February 2020

JEL classification:
C09
D91
H21
H26
H30

Keywords:
Donations
Peer effects
Public projects
Social norms
Tax morale

A B S T R A C T

This paper uses lab-in-the-field experiments and theory to explore why people give money to governments. We assume that giving is motivated by outcome-oriented or consequentialist norms, and conditional on (a) others’ behavior and (b) beliefs about how competent the government is. The evidence from a lab experiment in Peru is in line with this. On the other hand, we analyze the potential effects of two policies to increase giving, observing that less people give zero if they are informed about (i) two specific government projects (a subway line and a children’s hospital) or (ii) that some well-known Olympic medalist pays punctually her taxes, according to public information released by the Peruvian tax Agency. Our findings contribute to a burgeoning literature on tax morale and are arguably relevant to understand taxpayers’ non-selfish reasons to pay (or evade) their taxes.

© 2020 Elsevier Inc. All rights reserved.

1. Introduction

Suppose that you participate in the following experiment, together with several other subjects: The experimenter endows each of you with $10, writes down a Public Treasury bank account number in a blackboard, and explains that you can anonymously donate whatever fraction of your endowment to the Government. After all subjects have decided, the experimenter will go to the nearest bank office and make an anonymous deposit for an amount equal to the aggregate donation. If you give $2, for example, the Government will hence increase its budget in that amount. What would you do?

We think that this decision problem allows us to understand better the potential non-selfish motives why people give money to their governments. Since paying taxes is a particular instance of that behavior, further, the experiment arguably offers insights on tax compliance, a question of capital importance – e.g., (Alm and Jackson, 1993; Andreoni et al., 1998; Torgler, 2005; Luttmer and Singhal, 2014; Mascagni, 2018, Alm, 2019). Of course, taxes are typically associated to a system of fines and sanctions for those who evade them,1 and to formal laws that somehow express the will of the society or some societal groups. The experiment cited abstracts from such features, focusing on those motives to give that do not require deterrence or the existence of a legal code; we extend on this argument later, in Section 2.

1 We are grateful to Gibrán Cruz-Martínez and participants at the First International Congress in Behavioral Economics and Finance at the University of Lima in November 2018 for helpful comments and suggestions. We also gratefully acknowledge financial support from Instituto de Investigación Científica IDIC at the University of Lima and helpful research assistance by Deyvi Abanto and Christie Awa also at the University of Lima.
* Corresponding author.
E-mail addresses: raul.lopez@csic.es (R. López-Pérez), alramirez@ullima.edu.pe (A. Ramírez-Zamudio).

https://doi.org/10.1016/j.irle.2020.105892
0144-8188/© 2020 Elsevier Inc. All rights reserved.
We study four main research questions, numbered I to IV below. To start, I: do people give more if they expect the government to spend the money efficiently and competently? II: do people give more if they expect others to give as well? Our tentative answer to both I and II is “yes”, based on the utility theory developed in López-Pérez and Ramirez-Zamudio (2018). In this theory, people can internalize norms, and hence suffer a (heterogeneous) psychological cost C if they act contrary to what the norm prescribes, i.e., a negative, painful emotion like guilt or shame. People trade off this cost C with their material interest. Further, the norms are outcome-oriented or consequentialist, in that acts are prescribed conditional on their expected consequences; more precisely, if they lead to an allocation maximizing social welfare. If people have heterogeneous expectations about the consequences of some act, therefore, they will face different normative prescriptions. In our experiment, for instance, some may think that their government is highly incompetent or corrupt. Hence, they should conclude that giving money is a waste, even in terms of social welfare. Others may be more optimistic and thus infer that donating some money is the right thing to do. Our affirmative answer to question I is thus explained. Further, the cost C of deviating is not fixed, but depends directly on (a) the magnitude of the deviation, i.e., the expected ‘harm’ or decrease in social welfare caused by the deviation. Point (a) reinforces the positive effect hypothesized for question I: If an agent believes that the State is relatively efficient, she will feel rather badly for not doing the right thing (i.e., giving money, given her expectations), and hence will tend not to deviate much from the norm. In addition, C depends indirectly on (b) the average deviation among the reference group members (peer effects). Regarding question II above, therefore, we predict individual donations in our experiment to be conditional on others’ donations: If others are expected to give a lot, intuitively, an agent feels ‘inferior’ if she gives little, thus suffering a relatively larger ‘remorse’.2

Our third question is III: Do inequality concerns co-vary with giving? To clarify, think of a person who views social welfare as depending on equality (and not only, say, the social surplus). If this person believes as well that her government helps the poor or does a good job fighting inequality, she should donate ceteris paribus more money in our experiment. In effect, since she believes that giving is more socially ‘helpful’ than a person who is just concerned about the social surplus, her cost C for deviating is larger and hence should give more – recall assumption (a) above. We also explore two research questions related to III. To start, left-wing supporters are perhaps more concerned about inequality; do they give more? Additionally, a person might dislike inequality because she believes that income mostly depends on chance (and not effort), thus finding large inequalities unfair. In this respect, do beliefs about the determinants of personal income influence donations?

We explore research questions I to III (and also IV below) with the help of the experiment described above, which was run in Lima, Peru, having as recipient of the subjects’ donations the Peruvian Tax Office. In this manner, we contribute to a nascent literature on motivating tax compliance or contributions to public projects in developing countries, e.g., Mascagni (2018). It is also worth noting that participants in our study were a representative sample of the population of taxpayers in Lima, and not just university students. The data shows that (i) a substantial share of the subjects give something (around 64 % of them; N = 117), and a subject’s choice co-moves with (ii) her incentivized beliefs about the average donation by other subjects, and (iii) her level of support of the current government, arguably a proxy of her trust in that government. This is consistent with our presuppositions regarding questions I and II. On the other hand, we do not find a correlation between donations and other variables like gender, political ideology, religiosity, age, and education. Left-wing supporters, in particular, are not more likely to give. In our experiment, in fact, fairness considerations do not seem to account for behavior among the donors. In contrast, non-donors are significantly more egalitarian and believe that chance is a relatively stronger determinant of income. This is an intriguing result for which we have yet no good explanation.3 In any case, our results indicate that the relationship between giving, fairness concerns, and ideology is not a simple one. Less unexpectedly, non-donors have as well relatively high (incentivized) expectations about the tax-to-GDP ratio, i.e., they arguably perceive a stronger tax burden.

Our final research question IV concerns the potential effects of two policies that aim to increase giving to the government. In the Informed Policies (IP) treatment, the instructions mention two recent public projects in Lima: The subway line 2 and the new Children’s Hospital. In line with our theoretical framework, we conjecture that people are more willing to give if they are happy with the way in which the government uses the money. Yet people can be sometimes unaware or relatively uninformed about government spending (we provide indeed evidence in this line). By giving information about two specific public projects, therefore, we expect a significant increase in median giving relative to the Control treatment, as it happens to be the case (p < 0.032).4 In the Public Figure (PF) treatment, in turn, subjects receive instead information about a famous Peruvian Olympic medalist, noting that her firm pays its taxes punctually. In this respect, our framework predicts that giving is conditional, that is, people give more if they expect others to give as well. More to the point, further, we conjecture that this conditionalities depends on the identity of the agent’s referents: Well-known and reputed public figures might be particularly influential. We hence predict that providing information about compliance by such type of figures should foster donations. In this line, the average donation equals 2.6 soles in Control (out of a 30 Soles endowment), but almost 5 soles in PF (p < 0.016). Fig. 1 below provides a more detailed picture of the treatment effects. The length of each box represents the interquartile range (IQR) in the corresponding treatment, whereas the vertical lines (“whiskers”) extend above (below) so as to include all data points within 1.5 IQR of the upper (lower) quartile, stopping at the largest such value – in turn, the dots represent data points beyond the whiskers. Clearly, subjects tend to be less generous in the control. Indeed, the median giver, represented by a line within each box, donates 0 Soles in Control, but 3 Soles in both IP and PF.

The rest of the paper proceeds as follows. Section 2 discusses how our research contributes to some related literature, while experimental design and procedures are described in Section 3. Then Section 4 reports aggregate results, presents our research hypotheses in what regards Questions I to IV, and tests them with

---

2 This hypothesis surely abstracts from much of the actual diversity in agents’ feelings. As noticed by one referee, for instance, some people might think that acting selfishly is ‘smart’, and thus feel superior to any subject who gives something in our experiment.

3 Social assistance programs, it must be noticed, account for a rather small part of the already small Peruvian public budget. In the year 2017, indeed, they amounted to 3.3% of the total budget and 0.69% of GDP. If many subjects in our study anticipate this, fairness concerns might possibly play a relatively little role in their decision whether to give.

4 As noted by two referees, subjects might give more generously in IP for other, less standard reasons than learning. On one hand, hearing about good projects could improve subjects’ positive feelings towards the government. Alternatively, describing a project like the Children’s hospital could create an “identifiable victim” effect, e.g., Laxewinstein and Lerner (2003). Further, the evidence provided could be well-known by the subjects, but became salient in their minds when it is mentioned in the instructions. In any case, the key point of question IV is not discriminating between these potential causes but exploring the effect of the policies on donations.
a regression analysis. The last section concludes with a discussion of our main contributions.

2. Related literature

This study contributes to the literature on tax morale (Luttmer and Singhal, 2014), and more particularly to the understanding of tax compliance: What are the taxpayers’ motives to pay their taxes? Our account basically says that taxpayers are more likely to comply if they trust the government and believe that others pay their taxes as well, while they feel no remorse for evading their taxes if they think that public money is often stolen or squandered, or that many people also evade.5 We offer suggestive evidence in this line from an experiment in which real taxpayers decide on a real donation to their government, carefully controlling for their beliefs as well as investigating the effects of two policies.

Needless to say, our experiment does not perfectly resemble a tax compliance decision. In particular, there were no sanctions at all for not donating. This was necessary because we wanted to isolate and investigate the taxpayers’ non-selfish motives for giving money. Yet the question naturally arises: could these motives be crowded out when taxpayers face sanctions? In other terms, are our results relevant for understanding tax compliance? We believe so, for several reasons. On one hand, evidence from the neuroscience field suggests that both taxation and voluntary giving activate similar neural substrates: Harbaugh et al. (2007) show that both voluntary donations and mandatory transfers to a charity elicit activity in the same brain region associated with processing rewards.6 Note also that the existing literature on tax morale operates under the implicit assumption that non-selfish motivations are not crowded out in the field. In effect, while most of the early developments on tax compliance relied on the model by Allingham and Sandmo (1972), who assumed rational and selfish taxpayers, a more recent strand of the literature has introduced non-standard motives like altruism, emotions, reciprocity, social image, or social norms – e.g., Andreoni et al., 1998; Frey, 2003; Bazart and Bonein, 2014; Luttmer and Singhal, 2014; Varvarigos, 2016; Olsen et al., 2018. This was motivated by abundant evidence at odds with the selfishness hypothesis, particularly the observed (high) levels of tax compliance in developed countries in spite of the relatively weak expected sanctions; Andreoni et al., 1998. For instance, the calibration in Alm et al. (1992a) indicates that a mid-range estimate of the coefficient of relative risk aversion (R = 3) implies a rate of compliance of only 13 percent, well below any audit-based estimates of compliance for most forms of business income. Only quite high (R = 5) or extraordinarily high (R = 10) levels of risk aversion could explain the estimated levels of compliance in western countries.

One of the non-selfish motives that we emphasize here is what Luttmer and Singhal (2014) call peer effects and social influences, i.e., an additional term in the utility function that depends on whether others pay their taxes or view this behavior as appropriate. In this vein, Torgler (2005) uses data from Latinobaròmetro and the WVS, reporting that an individual’s tax morale co-moves with the belief that other individuals obey the law. Del Carpio (2014) studies property tax compliance in two municipalities in Lima. In her field experiment, residents receive a letter from the municipality that informs them about the actual average rate of compliance (peer effects), the average level of municipal enforcement (dissuasion), or both. While she finds a significant joint effect of (i) disclosing the compliance rate and (ii) a payment reminder, she cannot attribute this significance to peer effects only. In turn, Andreighetto et al. (2016) report data from a tax compliance experiment in Italy and Sweden, interpreting the differences observed as the effect of dissimilar norms and cultures across those countries. Our work advances these contributions by offering controlled and detailed evidence on the relationship between subjects’ beliefs and their actual donation, i.e., question I above.

Our analysis of question I also contributes to a literature showing that tax compliance is conditional on the citizens’ trust in their government and their expectations about how public institutions use their revenues. For instance, Alm et al. (1992b) provide evidence that tax compliance increases when taxpayers are aware of a direct link between their tax payments and the provision of a desirable public good. Torgler (2003) contends that trust in the president has a positive effect on tax morale, particularly in developing countries. Further, Torgler (2005) reports that tax morale in Latin America is weak among citizens who perceive that lack of honesty and corruption are widespread. In contrast, trust in the president and government officials has a significant positive effect on tax morale, whereas the effective tax burden has a negative effect. In turn, the “slippery slope framework” of tax compliance (Kirchler et al., 2008), which postulates that compliance can be achieved through power (coercion) and from a trustworthy relationship between tax authorities and taxpayers, has been supported by a substantial body of evidence, e.g., Wahl et al., 2010; Kagler et al., 2013; and Kagler et al., 2013. See also Feld and Frey (2002) and Richardson (2008).

In addition, our study of question IV is related to a strand of the literature interested in the design of public policies to improve tax compliance – e.g., Blumenthal et al., 2001; Stemrod et al., 2001, Kleven et al., 2011; Castro and Scarlascini, 2015; Del Carpio, 2014; Bott et al., 2014; Battiston and Gamba, 2016; Hallsworth et al., 2017; see also the excellent review in Mascagni, 2018. In a typical field experiment, taxpayers receive letters from the revenue administration, including some moral appeal or information about an upcoming audit, enforcement and potential fines, other taxpayers’ behavior, how government resources are used, etc. Mascagni (2018) notes that the evidence on the effectiveness of moral suasion or the information on the public goods is mixed; many studies often find no effects. She also stresses the need for more evidence from low- and middle-income countries. In a controlled setting and with a representative sample of Lima’s taxpayers, in this vein, we

---

5 This account follows from the utility theory in López-Pérez and Ramírez-Zamudio (2018), based in turn in López-Pérez (2008). Consult the latter paper for an extensive review of the theoretical literature on non-selfish preferences. Some models, including altruism (e.g., Andreoni and Miller, 2002), inequity aversion (e.g., Fehr and Schmidt, 1999), or the stylized model of quasi-maximin preferences in Charness and Rabin (2002), assume that utility only depends on the payoff allocation, with different functional forms depending on the model. In our theory, the payoff allocation is not the only relevant factor: people also care about how much others deviate from norms. This feature generates highly conditional behavior, which relates our model to existing theories of reciprocity, e.g., Rabin (1993).

6 This suggests that mandatory taxation under the threat of punishment does not entirely crowd-out non-selfish motives. Yet the psychological cost C may be smaller when fines are present. In this line, Harbaugh et al. (2007) also note that neural activity further increases when people make transfers voluntarily, perhaps a “warm-glow” effect related to free choice. Further research is warranted.
find a significant effect of both giving information about two public projects and about recorded compliance by a firm owned by a well-known public figure.

Finally, we would like to highlight some similarities between our results and those from the literature on VCM experiments, where subjects contribute to a public good – see Chaudhuri (2011) for a review. A first one is that subjects give more than predicted by the selfishness hypothesis: in one-shot versions of the VCM game, participants on average typically contribute between 40 % and 60 % of the optimal level, e.g., Dawes (1980). In VCM games, second, there are also wide variations in individual contributions. Akin to the peer effects that we find here, third, many participants in VCM experiments behave as conditional cooperators, i.e., their contributions are positively correlated with their beliefs about the contributions by the other group members – see for instance Keser and van Winden (2000), Fischbacher et al. (2001), or Fischbacher and Gächter (2010). Among the differences worth citing, in turn, our novel design allows us to study contributions to a real government by a representative sample of taxpayers, controlling for their support to that government and exploring the influence that (i) public figures and (ii) information about actual public projects may have on giving.  

3. Experimental design and procedures

Our experiment is a one-shot decision problem where each subject is endowed with 30 Soles (around $US 10) and can voluntarily donate some of this endowment to the Peruvian government. This donation is implemented by means of an actual bank deposit to an account of the Peruvian Public Treasury (Banco de la Nación account number 00000–299294), made anonymously by two of the experimenters after all participants have finished their choices (with two subjects acting as witnesses). Any subject’s payoff equals the initial endowment minus the donation, plus a 20 soles (around $US 7) show-up fee.

Before any session started, copies of the instructions and decision sheet were distributed in conveniently separated seats across the room. Then every subject entered the room and chose one of those seats. They first read the instructions at their own pace; subsequently, the experimenter read them aloud to ensure common information. Questions were privately clarified. Any subject was identified by an individual numerical code, included in her/his decision sheet. Although an experimenter was present in the room during each session, the instructions attempted to diminish potential demand effects or other confounds. For instance, we used neutral language and stated that there were no tricky questions and that subjects could choose as they preferred. The instructions also recalled that the Peruvian government offers different public services. In this respect, the experimenter noted verbally, while reading aloud the instructions, that the subject’s donation would be used by the Public Treasury to finance similar expenditures and public projects as taxes do; subjects were also informed in this manner about the Banco de la Nación account number mentioned above, which was written in a blackboard as well.

All decisions were taken with pencil and paper. When subjects had decided on their donation, decision sheets were collected and an elicitation sheet given. Four of the beliefs elicited are particularly relevant. First, we asked each subject to estimate the (rounded) average donation among all participants in the session. Second, we also elicited beliefs about the position of Peru in the corruption index of 2017 by Transparency International and, third, about the level of the Tax-to-GDP ratio in Peru for 2017. Finally, Question 4 had four items, all of them regarding the Peruvian government’s budget –i.e., the fraction allocated to social programs, infrastructure, public functionaries’ salaries, and other expenditures. Questions 1–4 were incentivized: At the end of each session, we randomly selected (i) one of these questions and (ii) one of the subjects who correctly answered that question.  

This subject earned 30 soles for this correct answer.  

After all subjects had their beliefs elicited, we collected the corresponding sheet. Then they answered a brief questionnaire which requested some socio-demographic information and made some questions about support to the current presidential team, concern for inequality, etc. The experiment ended with the selection of the subject who, eventually, got 30 soles for answering correctly one of Questions 1–4. Subjects were then paid in private by an assistant who, as the subjects’ instructions clarified, was not informed about the details of the experiment and was waiting in a different room during the experiment. After all subjects had been paid, two of the experimenters and the subjects acting as witnesses went to a nearby bank office, where an anonymous deposit was made for the total amount donated.

In addition to our Control treatment, we considered two slight treatment variations. In the Informed Policies (IP) treatment, the instructions included information about two specific projects recently developed by the Peruvian government in Lima (the construction of line 2 of the subway and the new Children’s Hospital). Finally, the Public Figure (PF) treatment was also identical to Control, except that the instructions mentioned a firm, owned by Peruvian Olympic medalist Natalia Málaga, and noted that such firm pays its taxes punctually and has no debts with the Tax Agency. The average payoff for the participants was 47.40 soles in Control, 46.40 in IP, and 45.04 in the PF treatment, including always the mentioned show-up fee of 20 soles. We run one session per treatment in August 2018 at Universidad de Lima, with approximately 40 participants in each one. Subjects were between 23 and 60 years old and economically active. In fact, they were selected by the University of Lima’s market research department following precise instructions, so that the random sample was representative of the taxpayer population of Metropolitan Lima regarding age, gender, and socio-economic conditions. In any case, recruiters did not disclose any detail about the experiment to the subjects at the time of recruiting, except that this was a “focus group” meeting to collect opinions about government, institutions and other social issues. Each session lasted approximately 90 min, including paying the subjects individually. No subject attended more than one treatment.

---

7 These features of our design also make a substantial difference with many tax compliance experiments, which use a similar design as Afn (1991), with the following phases: (1) individuals perform a clerical task and earn real income, (2) they pay ‘taxes’ on income voluntarily reported, (3) subjects who unreported taxes face some chance of being fined, and (4) tax revenues are used to fund a public good. There are also many differences with the typical field experiments in which letters with varied messages are sent to groups of taxpayers – e.g., Blumenthal et al., 2001; Stenrod et al., 2001; Klevén et al., 2011; Castro and Scartascini, 2015; Del Carpio, 2014. For obvious reasons, these experiments do not get feedback from subjects about their support to the government, preferences for equality, beliefs about other taxpayers’ behavior, etc.

8 The translated instructions, decision forms and questionnaires can be found in Appendix II (in Supplementary material).

9 We wrote each individual ID code in a little piece of paper. These pieces were introduced into an opaque bag, from which one subject took out the winner.

10 The actual answers are shown in Appendix II (in Supplementary material).

11 Miss Málaga is a well-known public figure, at the present the coach of the national juvenile female volleyball team (https://es.wikipedia.org/wiki/Natalia_ M%20M%C3%A1laga). Here is a small firm, as catalogued by the tax authority of Peru. The information about a firm’s debts with the Tax Agency is publicly available and banks often use it for credit purposes. See https://e-consultaruc.sunat.gob.pe/cl-ti-itmrconsruc/frameCriterioBusqueda.jsp.
Table 1
Descriptive statistics of each treatment.

<table>
<thead>
<tr>
<th>Treatments</th>
<th>Number of Subjects</th>
<th>Average Donation</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>39</td>
<td>2.60</td>
<td>5.59</td>
</tr>
<tr>
<td>Informed Policies (IP)</td>
<td>40</td>
<td>3.60</td>
<td>4.19</td>
</tr>
<tr>
<td>Public Figure (PF)</td>
<td>38</td>
<td>4.96</td>
<td>5.26</td>
</tr>
<tr>
<td>All treatments</td>
<td>117</td>
<td>3.71</td>
<td>5.09</td>
</tr>
</tbody>
</table>

Table 1 presents some descriptive data regarding the distribution of donations in each treatment. We observe the highest average donation in PF (4.96 Soles), whereas the median donation was 3 Soles in both IP and PF, but 0 in Control; the Median test indicates a significant difference in both IP (p < 0.032) and PF (p < 0.016) with respect to Control.

For a more visual representation, Fig. 2 below depicts the percentage of subjects who give something in each treatment. This is lowest in Control and highest in PF. As we will discuss more carefully later, these figures represent one of the main differences across treatments. In contrast, the average donation among those subjects who give something is rather similar across treatments, and amounts to 5.6, 5.7, and 6.1 Soles in treatments Control, IP and PF, respectively.

In turn, Fig. 3 shows the cumulative distribution function (CDF) of donations in each treatment. Clearly, the CDF for the control lies above the other CDFs in most ranges, a signal that subjects tend to give less in that treatment. We note that a Kolmogorov-Smirnov test rejects the null that the two samples in Control and PF come from the same distribution (p = 0.016). Yet this is not true in the comparisons Control vs. IP (p = 0.122) and IP vs. PF (p = 0.843).12

4. Data analysis

This section starts with a brief summary of the subjects’ decisions in each treatment. Afterwards, we present several hypotheses in what regards Question I to IV in the introduction. We subsequently test these hypotheses and study the causes behind any potential difference in average donations across treatments.

4.1. Summary of results

Table 1 presents some descriptive data regarding the distribution of donations in each treatment. We observe the highest average donation in PF (4.96 Soles), whereas the median donation was 3 Soles in both IP and PF, but 0 in Control; the Median test indicates a significant difference in both IP (p < 0.032) and PF (p < 0.016) with respect to Control.

For a more visual representation, Fig. 2 below depicts the percentage of subjects who give something in each treatment. This is lowest in Control and highest in PF. As we will discuss more carefully later, these figures represent one of the main differences across treatments. In contrast, the average donation among those subjects who give something is rather similar across treatments, and amounts to 5.6, 5.7, and 6.1 Soles in treatments Control, IP and PF, respectively.

In turn, Fig. 3 shows the cumulative distribution function (CDF) of donations in each treatment. Clearly, the CDF for the control lies above the other CDFs in most ranges, a signal that subjects tend to give less in that treatment. We note that a Kolmogorov-Smirnov test rejects the null that the two samples in Control and PF come from the same distribution (p = 0.016). Yet this is not true in the comparisons Control vs. IP (p = 0.122) and IP vs. PF (p = 0.843).12

A Mann-Whitney, in contrast, indicates that the p-values for the comparisons (i) Control vs. PF, (ii) Control vs. IP, and (iii) IP vs. PF are: (i) 0.0022, (ii) 0.0304, and (iii) 0.3247, respectively.

We note finally that the average subject for all treatments was around 38 years-old, slightly oriented to the right side in politics (6 out of 10), and more religious than the half line (6 out of 10). Further, he/she believes that Peru is a rather corrupt country (position 131 out of 180 countries in the Transparency International ranking). We find no significant distributional differences across experimental groups with regard to gender, political ideology, socio-economic status, and religiosity (two-sample Mann-Whitney test; p > 0.3677 always except for the comparison between control and PF regarding status, where p > 0.1388; results are similar with a Kolmogorov-Smirnov test).

4.2. Determinants of the donation: testing our research hypotheses

The utility theory informally presented in the introduction can be made a bit more precise, which helps to organize the analysis. In our experiment, a subject chooses a donation level D from 0 to 30. A norm selects one or (maybe) several donations; these are the normative choices whereas remaining choices are deviations from that norm. For a prominent example, consider a norm selecting those acts that lead towards a socially desirable or ‘fair’ outcome, i.e., one maximizing some social welfare function (SWF). More specifically, assume that donations have social consequences or, more precisely, that each D leads to some allocation x(D) = [x_1(D),..., x_i(D),..., x_n(D)] of material payoffs among the I members of the society. Further, let s(x) = \sum x_i denote the social efficiency of allocation x and e: \mathbb{R}^I \rightarrow \mathbb{R} denote some measure of how equitable allocation x is. An ‘efficiency and equity’ SWF, or W^E, is defined as

---

12 The lack of differences between Control and IP seems to be driven in large part by the presence of one subject in Control who donates her whole endowment (see Fig. 1). When we remove this outlier from the sample, a K-S test finds a marginally significant difference (p = 0.088).
some increasing function of \( s(x) \) and \( s(x) \).

An associated E-norm then naturally states: ‘choose D if you think that it leads towards an allocation maximizing \( W^D \).’ We posit for simplicity that any subject finds binding an E-norm. Heterogeneity can be introduced by assuming that subjects care about different E-norms, i.e., different specifications of \( W^D \). For instance, some people may be more sensitive to equality than others.

Additionally, subjects may have different expectations regarding the outcome \( x(D) \) of a donation. This has crucial implications on what subjects find normative in our experiment. To clarify this point, let \( D_{ni} \in [0, 30] \) denote the smallest donation that subject \( i \) finds normative. On one extreme, those subjects who find the public sector to be characterized by corruption and inefficiency will think that the money is better in their purses, and that \( D_{ni} = 0 \) is in fact the normative choice. Giving money, that is, does not increase \( W^D \). In contrast, another group of subjects expect the government to be sufficiently competent, so that \( D_{ni} > 0 \) is the ‘right’ level of donation. In what follows, we refer to the first group of subjects as (government) ‘skeptics’, and the second as ‘believers’.

In any case, subject i’s utility equals his monetary payoff 30- \( D \) minus an idiosyncratic cost \( C \geq 0 \), suffered only if her choice \( D \) is a deviation from her norm; in particular, if \( D < D_{ni} \). Cost \( C \) depends on three things. First, \( C \) increases with the ‘social damage’ or decrease in social welfare caused by subject’s choice \( D \), i.e., a function \( d_i(D) \) of the difference \( W^D(D_{ni}) - W^D(D) \) between the social welfare (as measured by \( W^D \)) had she chosen the normative act \( D_{ni} \) and the actual level of \( W^D \) given donation \( D \). We posit this function to be strictly convex, so that ‘large’ deviations from the norm are relatively more painful than ‘small’ deviations. Second, \( C \) depends on how well subject i’s choice compares to others’. More precisely, there is a reference group \( G = \{1, 2, \ldots, n\} \) and subject i has expectations regarding how any referent \( g \) would act if he were in i’s position. As a result, i has beliefs as well about how damaging (i.e., ‘bad’) the referents’ choices would be. We then posit that cost \( C \) inversely depends on the average damage \( d_{-i} \) caused by the referents’ (expected) choices. Intuitively, if others are expected to deviate substantially from the norm, hence causing a large damage, the subject feels less remorse by deviating as well. In short, she is conditional, in that her negative feelings (guilt, shame?) partly depend on what the referents do. Finally, we introduce a (heterogeneous) parameter \( \gamma \geq 0 \) to represent how deeply the subject has internalized the E-norm. To illustrate our assumptions, suppose that the damage of subject i’s choice, \( d_i(D) \in [0, 1] \), is measured on a scale from 0 (nil damage, when \( D = D_{ni} \)) to 1 (maximum). Further, let \( x_i \) denote subject i’s material payoff. A simple example of a utility function satisfying our assumptions so far is

\[
u_i(D) = x_i(D) - \gamma \cdot (1 - d_{-i}) \cdot d_i(D)
\]

where, the cost C discussed above hence equals \( \gamma \cdot (1 - d_{-i}) \cdot d_i(D) \). If \( \gamma = 0 \), the cost \( C \) equals zero, so that we are back to the standard economic model with selfish agents. According to this model, all subjects should give zero Soles in our experiment. Although there are 42 subjects across all treatments who act in this manner (around 36 %; see Table 1), we note however that donations are statistically significant in all treatments (Wilcoxon signed-rank test; \( p \)-value < 0.001 always). In what follows, a selfish agent is one who has \( \gamma = 0 \).

If subject i is a skeptic, her optimal choice is straightforward: As the norm recommends \( D_{ni} = 0 \), which moreover maximizes her material payoff, she gives nothing. In contrast, believers face a trade-off between their material payoff and the ‘damage’ caused if they deviate from the norm, that is, give less than their \( D_{ni} \). Since believers may have different expectations about the consequences of a donation, they may have as well different perceptions on how damaging choice \( D \) is. If the believer is not too optimistic about the government’s performance, she should give zero because the damage from not donating, i.e., \( d_i(D=0) \), is expected to be low. Otherwise she will donate something, perhaps the whole endowment – see López-Pérez and Ramírez-Zamudio (2018) for a full analysis of a believer’s utility maximization problem.

The previous discussion implies that a non-selfish subject’s donation depends on her belief that the government is competent and works for the common good, or more precisely, on how efficient it is in maximizing the social welfare. As a result, the responses to several questions in our questionnaire should be correlated with a subject’s choice. Ceteris paribus, one might expect that a subject’s donation is negatively correlated with her perceptions about the level of waste, incompetence, and corruption in the public sector. In addition, a positive relation between the subject’s donation and her/his support to the current presidential team is expected. Note yet that the potential existence of selfish subjects is possibly a source of confounds, as their decision to donate nothing is theoretically unaffected by their perceptions. If the share of selfish subjects is sufficiently large and their perceptions show enough variation, a regression analysis could then find no correlation between donations and perceptions. For this reason, we propose a more conservative hypothesis, focused on the donors.

**Hypothesis 1.** Among those subjects who give something, donations are negatively correlated with the subject’s perceptions about the level of waste and corruption in the public sector, and positively with her level of support of the current president.

**Evidence:** Table A1 in Appendix A reports the results from a regression analysis, where we pool the data from all treatments. Models 1 and 2 are linear OLS models and the dependent variable is a subject’s donation to the government (labeled D henceforth), but restricted to those subjects whose donation is larger than zero. Model 1 includes several key variables, in particular those related to Hypothesis 1. Model 2 adds some other variables collected in the experiment, but hardly seems to improve Model 1, at least in terms of the adjusted R-squared. Focusing on Model 1, we observe that the coefficient of Variable 2, support to current president, has the expected positive sign and is significant (\( p \)-value = 0.0035). For further detail, Fig. 4 depicts the distribution of Variable 2, which is defined from 0 (minimum) to 10 (maximum), conditional on the subjects’ donation. More precisely, we have defined three groups of subjects of similar size: (1) Those who donate zero (\( N = 42 \) across all treatments), (2) those who give D such that \( 0 < D < 4 \) Soles (\( N = 38 \)), and (3) those who donate more than 4 (\( N = 37 \)). In line with our regression analysis, we can see in Fig. 4 that the most generous donors, i.e., group 3, tend to be more supportive of the current presidential team than the least generous donors, i.e., group 2. For instance, the percentage of subjects in group 2 who highly disapprove the government, i.e., assign it a score between 0 and 2, equals 32 %, i.e., 12 subjects. In group 3, in turn, they represent 19 % of the group (7 subjects).

---

\(^{13}\) The exposition so far assumes for simplicity that any agent is certain about the outcome \( x(D) \) of D. Yet our ideas can be readily extended to a setting with uncertainty, where each act is associated to a probability distribution over the set of all potential social allocations; see López-Pérez and Ramírez-Zamudio (2018). In this case, \( W^D \) would be a function of the expected social efficiency and equality of all potential allocations.

\(^{14}\) We do not report the whole analysis here, as model 2 also controls for the subject’s age, weekly frequency of watching news (on TV, internet, and others), and number of cars owned. Neither of these variables is significant (\( p \)-value > 0.286 always).
We do not find significant effects for the alternative variables related to Hypothesis 1. Variable 1, based on the subject’s perception about the ranking of Peru in the TI index, is not only non-significant (p = 0.313) but has moreover a positive sign, contrary to what we expected. We have checked as well the effect of the variable waste index, using it instead of the TI index in Model 1; while the coefficient of this waste variable has a negative sign, it is hardly significant (p = 0.929). In this alternative model, Variable 2, support to current president, is still significant (p = 0.054) and has the expected positive sign. Detailed results are available upon request.

Model 3 in Table A1 in Appendix A, in turn, is a logistic regression where the dependent variable is a dummy taking value 1 when the corresponding subject gives something, i.e., D > 0. Here we observe that Variables 1 and 2 are not determinants per se of a positive donation. Fig. 4 above provides some intuition in this regard: the subjects who donate relatively little, i.e., group 2, do not express a particularly high support to the president, compared to the non-donators (group 1). In any case, note that this result is not inconsistent with our theory and in particular with Hypothesis 1, which focuses on the donators: to repeat, selfish subjects give nothing but their perceptions are not predicted to be correlated with these variables. We overall interpret the evidence in Table A1 as not rejecting Hypothesis 1, except that the variable trust in the government seems to have a more significant explanatory role than those measuring perceptions about corruption or government waste. The following result summarizes our discussion so far.

Result 1: among those who give something, donations depend (positively) on the subject’s perceptions about the competency of the government, but not on the expected ranking of Peru in the TI index (or on the waste index). Compared to a non-donator, however, a donator is not more likely to trust the government. This is perhaps a signal that many non-donators are unaffected by their perceptions in this respect.

Consider now Question II in the introduction, which refers to a subject’s beliefs about others’ choices. More precisely, any subject has expectations regarding how any referent would act if he were in her position. As a result, she has beliefs about how damaging (i.e., ‘bad’) the referents’ choices would be. We then posit that cost C inversely depends on the average damage caused by the referents’ (expected) choices; see utility function (1). Intuitively, if others are expected to deviate substantially from the norm, hence causing a large damage, the agent feels less remorse by deviating as well. More precisely, we distinguish two cases. On one hand, the skeptics believe that DKi = 0 is the normative option so that their choice should be unaffected by their beliefs. For a believer, on the other hand, giving less than her DKi > 0 entails a cost C > 0 conditional on the referents’ average damage. Since the referents’ (expected) average donation to the government is a natural proxy for this, the following hypothesis is direct.

Hypothesis 2 (peer effects). Among the donators, the amount given directly depends on the subject’s belief about the average donation from others. This correlation may be absent when we focus specifically on those subjects who give nothing.

Evidence: Focusing on those subjects who give something, Models 1 and 2 in Table A1 in Appendix A report significant and positive coefficients for Variable 3 ‘average belief’ (p = 0.007 and 0.014, respectively). When non-donators are included in the analysis, see Model 3, the relationship is preserved: People who give a positive amount expect a higher donation average than the non-donators. We conclude therefore that Result 2: Donations co-move with beliefs about the average donation, and the relation is highly significant.

The previous result is of a correlational nature. Hence, it is a natural question whether beliefs cause behavior or the other way around, e.g., people might tend to believe that others donate roughly the same as them, due to a false consensus effect. In this respect, we note that in López-Pérez and Ramirez-Zamudio (2018) we have further explored this issue using a different dataset and treatments. In summary, we find there a treatment effect when the beliefs about the average donation are exogenously fixed, thus suggesting a causal role of those beliefs.

We move now to a different issue, that is, Question III in the introduction. The idea is that a subject’s distributive values, i.e., the consequences that the norm takes into account, may include not only the social efficiency of the allocation obtained, but also how unequal it is. This is indeed the intuition behind the E-norm presented in Section 4.2. As we noted, furthermore, subjects’ norms can be heterogeneous in how sensitive they are to inequality. This can be relevant to predict behavior. In effect, if a subject is highly concerned about inequality and moreover believes that a large share of the government money funds redistribution and anti-poverty programs, she should contribute relatively more because she believes that her taxes are more ‘useful’ in social welfare terms, as they help to reduce poverty and inequality. For such a person, formally, deviations from the normative choice DKi are relatively painful, i.e.,

15 Results are robust to the inclusion of controls as in Model 2.
16 Note that there is a non-trivial question in this respect: Who are a subject’s referents? That is, to whom does she compare with? As a first approximation, we posit that participants in an experiment compare with other participants; in the PF treatment, however, the Olympic medalist cited in the instructions is arguably a clear referent as well.
entail a large damage $d_i$ (D) and hence an increased cost $C$. As with Hypothesis 1, we propose a conservative hypothesis, focused on the donators. The reason is that, for those subjects who are selfish or have a very low ‘internalization parameter’ $\gamma$, their decision (not) to donate is unaffected by their inequality concerns (if any). Strictly speaking, therefore, our model unconditionally predicts a correlation between donations and inequality concerns only for the donators.

Hypothesis 3. Among those subjects who donate something, the amount donated directly depends on their concern about inequality, provided that they have sufficiently large expectations about the share of the public budget focused on social programs.

Evidence: In our questionnaire (see Appendix II (in Supplementary material)), we elicited a subject’s belief about the percentage allocated to social programs in the Peruvian national budget. We also asked the following “equality” question: “Do you believe that the distribution of income in a society should be as egalitarian as possible?”, which subjects had to answer using a number from 0 (completely disagree) to 10 (completely agree). This is Variable 5 in Table A1, Appendix A. In Model 1, it is not statistically significant ($p = 0.279$) and has moreover the unpredicted sign, i.e., negative. Still this analysis does not consider the possibility that the effect predicted by Hypothesis 3 happens only for those subjects with large enough expectations about the share that social programs take in the public budget. Hence, we divide the sample of donators ($N = 75$) in two halves, depending on whether their expectations are below or above the median belief, which is 20%. Then we explore a regression model similar to Model 2 but where, instead of Variable 5, we use an interaction with a dummy that takes value 1 if the subject’s expectation is above the median. In other words, we focus on those subjects who believe that social programs are a relatively large part of the budget. Even in this case, though, the variable takes a negative sign and is non-significant ($p = 0.340$).

What about the people who gave zero Solés? To our surprise, Model 3 in Table A1 indicates that donators are significantly less egalitarian than non-donators. Note well that this effect remains significant even after controlling for the subject’s trust in the current government and political ideology. We do not have a good explanation for this intriguing result; see footnote 4, though. Fig. 5 below illustrates further the previous discussion about the role of fairness concerns on donations. We represent there the distribution of subjects’ responses to the equality question, conditional on the subjects’ donations (we define three groups 1, 2, and 3, as in Fig. 4 above). Clearly, the non-donators, i.e., Group 1, tend to be more egalitarian. For instance, 64% of them give an answer higher or equal than 7 to the equality question, whereas less than 40% answered so in each of the other two groups. In turn, the differences between groups 2 and 3, i.e., between those who donate less and more than 4 Solés, respectively, appear to be hardly significant.

Hypothesis 4. Among those subjects who do not believe that inequality Variable 5 has no statistical significance in Model 1.

An even more subtle determinant of donations is a subject’s beliefs about the determinants of personal income. Subjects who believe that income is mostly determined by effort, and not by chance, might conclude that social programs are not fair, as they reward those who do not deserve it. This is basically equivalent in our account to having a small concern about equality. On the opposite, those subjects who believe that chance is the most relevant factor should be more willing to give money to the government, at least if they think that redistribution programs are significant enough.

Evidence: In the questionnaire, subjects indicated numerically their opinion about the relative importance of two factors in determining a person’s income: (a) chance and influences of other people and (b) the extent to which the person strives to work hard in life. Opinions could range from 0, that is, factor (a) is the only important one, to 10, signifying that personal endeavor (b) is the unique determinant. Subjects tend to believe that effort is the main determinant of personal income, (the median and average responses are 7 and 6.79, respectively). Contrary to what Hypothesis 4 states, on the other hand, Model 1 in Table A1, Appendix A, shows that the sign of variable 4 is positive ($p = 0.104$). We get the same result ($p = 0.267$) if we instead assume that the effect is restricted to those subjects with beliefs about the relevance of social programs above the median, as we did before for variable 5 (equality). As further evidence against Hypothesis 4, donators assign in average a relatively larger role to effort as a determinant of personal income; see Model 3 in Table A1. In summary, we fail to find significant support for Hypotheses 3 and 4 and sum up our discussion as follows:

Result 3: Neither a preference for equality nor a perception that income depends on chance (and not effort) explain the amount donated among those subjects who donate something. It seems therefore that fairness considerations do not account for behavior within that group. In contrast, non-donators are significantly more egalitarian and believe more that income depends on chance.
We have also checked the potential effect on donations of the subjects’ beliefs about the Tax-to-GDP ratio and the composition of the public budget, as measured respectively by Questions 3 and 4 in our incentivized questionnaire. We had several conjectures in this respect: (i) a relatively large expectation about the tax rate reduces a subject’s donation, (ii) a relatively large expectation about the percentage that goes to salaries in the public budget reduces a subject’s donation, and (iii) when a subject states large differences between the percentages going to social programs or infrastructures in the actual budget (Question 4) and the respective percentages in the ideal budget (Question 5), her donation is reduced. Nevertheless, when we run a regression similar to Model 1 in Table A1, Appendix A, but including now additional variables as the subject’s responses to Questions 3 and 4, and the differences between the answers to Questions 4 and 5, we find no support for conjectures (i) to (iii). The decision of how much to donate, therefore, does not seem to be motivated by such considerations.

Interestingly, the decision whether to donate appears to be influenced by these variables. To explore this point, we run a logit regression similar to Model 3 in Table A1, but including now the variables cited in the prior paragraph – results are available upon request. Here we observe that a subject’s expectation about the tax rate has a negative and significant coefficient (p-value = 0.026). The interpretation is that having a relatively large expectation about tax-to-GDP ratio increases the chances that one does not give to the government.\(^\text{17}\) We also observe that subjects who state relatively large differences between (a) the expected percentage going to infrastructure investments and (b) the desired, ideal one, are more likely not to give, although the effect is not significant (p-value = 0.202). One potential implication of these findings is that variables, like (beliefs about) the tax rate, have a non-linear effect, so that they affect the decision whether to donate, but not the exact amount donated (among those who decide to donate).

4.3. Exploring the differences across treatments: discussion

As we explained in Section 4.1, there is a significant treatment effect in both IP and PF. To recall, the median donation was 3 Soles in both treatments, but just 0 in the control. As we have explained before, we think that the higher donations in the Public Figure treatment are the effect of the information given about the tax compliance behavior of a well-known and reputed Olympian. In treatment IP, in turn, one possibility (among many others; see footnote 4) is that the information about the two specific public projects leads subjects to update their beliefs about the competency of the government, making them more accurate and positive. This would lead in turn to higher donations, compared to the control. We provide here some additional evidence on how these policies might potentially operate.

First, the treatment effect basically operates by increasing the fraction of subjects who give something. This was somehow anticipated in Section 4.1 above and is more forcefully indicated by Model 3 in Table A1, Appendix A, analyzing the determinants of a positive donation. The ‘treatment’ variable 10 takes value 1 if the subject participated in IP or PF, and its coefficient is highly significant (p = 0.002) and moreover of a positive sign. In contrast, this same variable is irrelevant in Models 1 and 2 that focus on those who give something: The treatment apparently affects the decision whether to give, but not the amount given.

One potential rationale for the IP treatment, second, was our conjecture that people are in general ‘too’ pessimistic about the performance of the public sector, so that informing them about some successful projects should make them less pessimistic. In line, we guessed that subjects in all treatments would over-estimate the position of Peru in the TI index and the tax-to-GDP ratio. These conjectures are validated by our data. To start, Peru was in the 96th position of the Transparency International corruption perceptions index for 2017, out of 180 countries. Of our 117 subjects, however, 76.92% of them over-estimated this position. Indeed, the average estimated position was 130.6, and the average error equal to 56. Interestingly, this over-estimation was less pronounced in the IP treatment, whereas the differences between the control and the PF treatments were not significant (Table 2 below offers evidence in this regard). With respect to the tax-to-GDP ratio of Peru in 2017, the actual figure was 12.9 %, but subjects made an average evaluation of 44.8. Indeed, a percentage of 91.5 % over-estimated the actual figure, and the mean error was 26.5.

As another rationale for the IP treatment, third, we also conjectured that subjects would under-estimate social spending and public investment in infrastructures, items that are often considered in good terms. This does not seem to be supported by our data. The fractions actually allocated in the 2016 public budget to social programs, infrastructure, salaries, and other expenditures were respectively of 8.16, 16.9, 34.66, and 40.28 %. Yet the subjects’ average beliefs were 21.2, 26.4, 36.5 and 15.9 %, respectively. Contrary to our a priori, therefore, subjects do not under-estimate social spending and public investment in infrastructures. Note however that the actual picture is possibly more complex, as at the same time the average subject grossly underestimates the item called ‘other expenditures’. We believe that this underestimation is possibly due to the subjects’ lack of awareness of the many sub-items included within this large category, and guess that they would have reported much lower figures for the other categories had they been aware of that issue. In any case, we note that the average difference between the subjects’ beliefs about each budget share and the corresponding ideal share are –9.9, –6.5, 13.8 and 2.3, respectively. Since the two first numbers are negative, people do not seem to be satisfied with the percentages devoted to social and infrastructure programs in the public budget.

In summary, subjects do not seem very optimistic or satisfied with the performance of the Peruvian public sector, and this is relevant because our prior analysis shows that some variables like trust in the government play a significant explanatory role in our experiment. A natural question is then: Do subjects give more in the PF or IP treatments due to a change in the subjects’ perceptions, as measured by our questionnaire? Overall, the evidence suggests a negative answer, as we find few significant differences across treatments in the distributions of the individual variables. In effect, Table 2 depicts the mean and median answers to several questions appearing in Parts 2 and 3 of the experiment, and the only prominent differences appear when comparing the beliefs about corruption, tax-to-GDP ratio, and the percentage of the public budget going to social programs. However, some of these variables, e.g., corruption as measured by the TI index, have no effect in the regression analysis reported in Table A1. Hence it seems unlikely that their change explains any treatment effect. The individual perceptions about tax-to-GDP ratio, in turn, were negatively correlated with the probability of giving something (see Model 3 in Table A1). Since these perceptions have a larger average and median value in treatment IP, however, they can hardly explain why subjects give more in IP than in Control.

\(^{17}\) As noted by one referee, this is perhaps a signal that the motives to give in our experiment are crowded out when taxes are high. As we suggested in Section 2, hence, some caution is warranted when extrapolating our results to the field. Still, we also note that the average belief about the Tax-to-GDP ratio among the non-donors equals 53.4%, whereas the actual figure was 12.9% in 2017 (for the donors, the average belief equals 39.9%). This suggests that the crowding out (if it exists in Peru) requires very high and inaccurate expectations about the tax rate. Alternatively, people with very high expectations about the tax rate might conclude that the government is inefficient, and thus decide to give nothing, consistent with our utility theory.
Table 2
Average and median responses to several questions, conditional on treatment.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Control</th>
<th>PF</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison</strong></td>
<td>Control-IP</td>
<td>Control-PF</td>
<td>IP-PF</td>
</tr>
<tr>
<td><strong>Statistics</strong></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td>1. Corruption (1: least – 180: most)</td>
<td>150</td>
<td>160***</td>
<td>132</td>
</tr>
<tr>
<td>2. Support to current presidential team</td>
<td>4.15</td>
<td>5</td>
<td>5.08</td>
</tr>
<tr>
<td>3. Beliefs about average donation by others</td>
<td>9.36</td>
<td>5</td>
<td>8.63</td>
</tr>
<tr>
<td>4. Role of effort (0: only luck, 10: only personal effort)</td>
<td>7.05</td>
<td>7</td>
<td>6.42</td>
</tr>
<tr>
<td>5. Equality preferences (0: None – 10: Maximum equality)</td>
<td>6.03</td>
<td>6</td>
<td>5.5</td>
</tr>
<tr>
<td>6. Waste Index</td>
<td>7.159</td>
<td>80</td>
<td>66</td>
</tr>
<tr>
<td>7. Beliefs of Tax-to-GDP ratio</td>
<td>41.96</td>
<td>40*</td>
<td>40.11</td>
</tr>
<tr>
<td>8. % going to social programs</td>
<td>20.97</td>
<td>20*</td>
<td>17.95</td>
</tr>
<tr>
<td>9. % going to infrastructures</td>
<td>26.03</td>
<td>20</td>
<td>26.68</td>
</tr>
<tr>
<td>10. % going to salaries</td>
<td>38.38</td>
<td>40</td>
<td>38.16</td>
</tr>
<tr>
<td>11. % going to other expenditures</td>
<td>14.62</td>
<td>10</td>
<td>17.21</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>39</td>
<td>38</td>
<td>38</td>
</tr>
</tbody>
</table>

Note: Each cell shows the average (left) and median (right) value of the corresponding variable in the corresponding treatment, indicated in columns. In column Control/ PF/ IP symbols *** , ** , and * indicate a significant difference at 1 %, 5 %, and 10 % levels, respectively, between the distributions of the corresponding variable in treatments Control vs. IP/ Control vs. PF/ IP vs. PF.

5. Conclusions

This paper makes three main contributions to the literature on tax morale. First of all, we explore why people are willing to give money to their government. We conjecture that people should be more willing to contribute to the public sector if they believe that taxes are well used, i.e., foster social welfare. To explore this point, we run incentivized experiments in Peru and observe that a subject’s donation co-moves with her trust in the current government, as well as her beliefs about the average donation by other participants. In turn, non-donators tend to believe that others will give relatively less. Intriguingly, they are more egalitarian and assign a larger role to good luck in determining personal income. Non-donators also have relatively high perceptions of the tax burden.

Second, we explore two policies to increase giving to the government. In the IP treatment, subjects receive information about two public projects recently developed in Lima. In PF, the instructions mention a firm, owned by a famous Olympic medalist, remarking that it pays its taxes punctually and has no debts with the Tax Agency. In both cases, we observe a significant increase in the median donation, relative to the baseline. Since we also have evidence that people have in average rather pessimistic beliefs about the public sector, the evidence from the IP treatment readily suggests a policy that could improve tax compliance: Inform taxpayers about specific, successful public projects. The implications from the PF treatment are also straightforward, in particular in what regards the use of respected and well-known public figures in campaigns to promote tax compliance. From our statistical analysis, we cannot reject the hypothesis that both treatments are equally effective, although Fig. 2 suggests that PF might be a bit more effective in reducing the fraction of people who give nothing.

Third, we argue that social norms play a non-negligible role in our experiment, and possibly in explaining tax compliance. In our account, people have internalized a norm prescribing any act expected to lead towards a socially desirable outcome. Further, they dislike deviating from what the norm prescribes. In our setting, the norm recommends giving money, but only if it is expected to be used for the public good. If the psychological cost of deviating is large, therefore, people will give. Furthermore, we are specific about the determinants of this cost: It depends directly on the social damage generated by the deviation (relative to the optimum), but also on how others behave. As we have noted, many of our results happen to be coherent with the model. Yet we also find that fairness considerations do not co-move with donations, an issue that demands further research.

Appendix A.

Table A1
Regression analysis of determinants of donations.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>D</th>
<th>D &gt; 0</th>
<th>Dummy (takes value 1 when D &gt; 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>Coefficient (left column)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Odd ratios (right column)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(3)</td>
</tr>
<tr>
<td>1. Corruption (1: least – 180: most)</td>
<td>0.010525</td>
<td>0.01015</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>(0.1035)</td>
<td>(0.0117)</td>
<td></td>
</tr>
<tr>
<td>2. Support to current president (0: not at all, 10: entirely)</td>
<td>0.46527**</td>
<td>0.40625*</td>
<td>0.0491</td>
</tr>
<tr>
<td></td>
<td>(0.2165)</td>
<td>(0.2339)</td>
<td></td>
</tr>
<tr>
<td>3. Belief average donation others</td>
<td>0.20578***</td>
<td>0.19874**</td>
<td>0.074**</td>
</tr>
<tr>
<td></td>
<td>(0.0735)</td>
<td>(0.0787)</td>
<td></td>
</tr>
<tr>
<td>4. Role of effort (0: only luck, 10: only personal effort)</td>
<td>0.34746</td>
<td>0.28987</td>
<td>0.1617**</td>
</tr>
<tr>
<td></td>
<td>(0.2107)</td>
<td>(0.2292)</td>
<td></td>
</tr>
<tr>
<td>5. Preferences for equality (0: none, 10: maximum equality)</td>
<td>-0.22911</td>
<td>-0.26016</td>
<td>-0.2489***</td>
</tr>
<tr>
<td></td>
<td>(0.2101)</td>
<td>(0.2262)</td>
<td></td>
</tr>
<tr>
<td>6. Gender (0: Male, 1: Female)</td>
<td>-1.71856</td>
<td>-1.71856</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.2559)</td>
<td>(1.2559)</td>
<td></td>
</tr>
<tr>
<td>7. Political preferences (0: extreme left, 10: extreme right)</td>
<td>-0.49785</td>
<td>-0.49785</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.4499)</td>
<td>(0.4499)</td>
<td></td>
</tr>
<tr>
<td>8. Education</td>
<td>-0.44228</td>
<td>-0.44228</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.9400)</td>
<td>(0.9400)</td>
<td></td>
</tr>
</tbody>
</table>
### Table A1 (Continued)

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>D</th>
<th>D &gt; 0</th>
<th>Dummy (takes value 1 when D &gt; 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>Coefficient (left column)</td>
</tr>
<tr>
<td>9. Socio-economic level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Treatment (1: participated in IP or PF; 0: otherwise)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.8274</td>
<td>2.9643</td>
<td>0.9921</td>
</tr>
<tr>
<td>Obs.</td>
<td>75</td>
<td>75</td>
<td>117</td>
</tr>
<tr>
<td>Adjusted R-square</td>
<td>0.1344</td>
<td>0.1189</td>
<td>0.1759</td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. All models apparently satisfy non-multicollinearity, as signaled by a VIF of less than 1.32 for any coefficient in all regressions. ***, **, and *" indicate significance at 1%, 5%, and 10% levels, respectively. In Model (3) we report odds ratios, to be interpreted as the probability of success divided by the probability of failure. For example, the odd ratio for Variable 4 ("role of effort") says that, among those subjects who believe that effort is the only cause of personal success, the probability of a positive donation is 1.17 times higher than the probability that a subject with completely opposite beliefs donates something.

### Appendix B. Supplementary data

Supplementary material related to this article can be found in the online version, at doi:10.1016/j.ijrle.2020.105892.

### References


Battiston, P., Gamba, S., 2016. The impact of social pressure on tax compliance: a field experiment. Int. Rev. Law Econ. 46 (C), 78–85.


Del Carpio, L., 2014. Are the neighbors cheating? Evidence from a social norm experiment on property taxes in Peru. mimeo.


Mascagni, G., 2018. From the lab to the field: a review of tax experiments. J. Econ. Surv. 32 (2), 273–301.


