



HAL
open science

Do Women Contribute More Effort than Men to a Real Public Good?

Ingela Alger, Laura Juarez, Miriam Juarez-Torres, Josepa Miquel-Florensa

► **To cite this version:**

Ingela Alger, Laura Juarez, Miriam Juarez-Torres, Josepa Miquel-Florensa. Do Women Contribute More Effort than Men to a Real Public Good?. *World Bank Economic Review*, 2023, 37 (2), pp.205-220. 10.1093/wber/lhac026 . hal-04050045

HAL Id: hal-04050045

<https://hal.science/hal-04050045>

Submitted on 27 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Do women contribute more effort than men to a real public good?

Ingela Alger, Laura Juarez, Miriam Juarez-Torres, and Josepa Miquel-Florensa

Abstract

This study presents evidence from a lab-in-the-field experiment, conducted in eight small, rural villages in Mexico, in which subjects choose to exert real effort to fund real health centers in their own and other localities. The results show that women are more willing than men to exert effort to fund the health center in another locality, relative to the one in their locality. This gender gap is mostly due to women who have some trust in the way the government spends taxes, and to those who benefit from a government program that targets women and fosters health care use. These results also suggest that women might be aware of their higher willingness to exert effort for a public good that does not benefit them directly, compared to men, because they seem to reduce their individual effort the more female their environment is.

JEL classification: H41, C91, O12

Keywords: public goods, gender, lab-in-the-field experiment, real effort, in- vs out-group transfers

1. Introduction

A society's ability to provide public goods, which are essential for economic development, depends on raising potentially distortionary taxes. For instance, labor taxes might disincentivize work, particularly if they do not benefit the worker directly.¹ The non-experimental literature has found that labor force participation and work hours respond to taxes, and such response varies by gender and education (Meghir and Phillips 2008). Taxes could also affect work effort, a relevant margin not measured in standard labor surveys.

This study tests whether women are more inclined than men to contribute effort towards a public good that may not benefit them directly. If confirmed, this gender gap would imply that policies that promote female labor

Ingela Alger is a CNRS Senior scientist at Toulouse School of Economics, University of Toulouse Capitole, and the Institute for Advanced Study in Toulouse, France; her email address is ingela.alger@tse-fr.eu. Laura Juarez is a professor at Centro de Estudios Economicos, El Colegio de México, Mexico City; her email address is laura.juarez@colmex.mx. Miriam Juarez-Torres is a researcher at the General Directorate of Economic Research, Banco de Mexico, Mexico City; her email address is mjuarez@banxico.org.mx. Josepa Miquel-Florensa is an associate professor at Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France; her email address is: pepita.miquel@tse-fr.eu. The research for this article was financed by the Agence Nationale de la Recherche (Chaire d'Excellence ANR-12-CHEX-0012-01; the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 789111 - ERC EvolvingEconomics), the French National Research Agency (ANR) under IAST grant ANR-17-EURE-0010 IAST (Investissements d'Avenir program). The authors thank Mariana Garcia and Ildrim Valley for outstanding research assistance, both in the field and in the office; Ingela Alger and Josepa Miquel-Florensa thank Banco de Mexico for its generous hosting. The views expressed in this article are solely those of the authors and do not necessarily reflect those of Banco de Mexico. A supplementary online appendix is available with this article at the *World Bank Economic Review* website.

¹ For instance, Summers (1989) shows, in a simple model, that labor taxes tied to mandated benefits that workers value have a lower disincentive effect on work compared to pure taxes, which do not benefit the worker directly.

market participation and attachment might enhance tax revenues with less distortions.² The data come from a lab-in-the-field experiment conducted in rural villages, inhabited mostly by low-income farmers, in the Mexican state of Yucatan.³ Gender differences in effort responses to taxes are unlikely to stem from those in exposure to real taxes in this population because of its isolation from the formal tax system.

In our experimental design, a subject's real effort in a simple but boring task determined his/her probability of generating earnings above and beyond the initial endowment; any such extra earnings were allocated to a public good. In the *Local Public Good* (LPG) game, they went to the public health center of the subject's village; in the *Non-Local Public Good* (NLPG) game, they went to the health center of another village, which may not benefit the subject directly. Thus, in this experiment, the tax rate was 100 percent, imposed by the research team, and not varied across treatments. An advantage of this design, compared to the standard public goods experiment (in which a subject's contribution also reduces his/her net payoff), is that it eliminates the income effect. It also isolates the effort choice from the transfer choice, and the former can be interpreted as the subject's willingness to contribute effort to the corresponding public good.

In each village, participants were divided randomly in two groups: *Donors* and *Recipients*. In the *Donor* treatment, any extra earnings generated by the subject increased the resources given to the public good; in the *Recipient* treatment, failure to generate such earnings reduced those resources, because the subject was compensated from them. Hence, this investigation also examines whether the gender differences in the willingness to exert effort for a public good vary by whether the subject gives/takes resources to/from the public good.

The empirical analysis consists in estimating the effort difference between the NLPG and the LPG games, and testing whether it varies significantly with gender and role (Donor vs Recipient). Arguably, this gender gap is a clean measure of the discrepancy between men's and women's willingness to contribute real effort to a public good that does not benefit them directly, because it controls for the effort they contribute to one that does, keeping the tax rate constant. To control for any individual characteristics that might affect effort, and that remained constant throughout the experimental session, all estimations include player fixed effects.

The results show that, on average, female *Donors* exert about 7% more effort than male *Donors* in the NLPG, relative to the LPG game, and this estimate is significant. Thus, female *Donors* are more willing than their male counterparts to contribute effort for the health center of another locality, relative to that in their own. Although there is no income reporting in this study's experiment, that finding is consistent with the results of Bruner et al. (2017), who show that the gender gap in tax compliance is largest when the tax revenues do not benefit participants directly. Thus, their work also implies that women are more willing to pay taxes, by reporting their income truthfully, than men in this case. For *Recipients*, no significant gender differences in effort between the NLPG and LPG games are found. All these findings are robust to controlling for the order of games in the experimental session.

This study further examines whether the estimated gender gap varies with measures of informal giving and receiving, community involvement, trust, strength of the participant's network in the matched village and their own, and other individual characteristics. The findings show that most of the positive gender gap in relative effort for *Donors* is due to the behavior of women who have at least some trust in the way the government spends taxes, and to those who participate in the Oportunidades program. This program pays a cash transfer to poor, rural women, conditional on complying with family health checks and other requirements.⁴ Therefore, Oportunidades beneficiaries, women who use health services relatively often, are those who are more willing to exert effort than men for the clinic at the matched locality, relative to that in their own.

² This study is in line with the growing interest in the potential effects that women's empowerment has on economic performance, both in developed (Adams and Funk 2012; Bertrand 2011), and in developing countries (Beaman et al. 2012; Dupas and Robinson 2013; Field et al. 2016).

³ This study also adds to the literature on behaviors and network structures in small villages in Mexico (Angelucci et al. 2009, 2010).

⁴ Oportunidades, which started as Progresa in 1997, is one of the best known conditional cash transfer programs in the world. The transfer from the program is also conditioned on children's enrollment and regular attendance to school. The literature evaluating different aspects of this program is extensive. For recent surveys on the long-term effects of this program, see Millán et al. (2019), and on impact on education see Gertler et al. (2006).

Most of the literature on anti-poverty and social insurance programs has focused on the negative effects of government programs on private interventions. In contrast, the Oportunidades result suggests that some government programs could actually foster social capital and a willingness to contribute to public goods, instead of crowding them out, as other recent studies find.⁵ Adato (2000) report mixed effects of Progresa on social capital. On the one hand, the formation of a “club” of program beneficiaries reduces their interaction with the rest of their community. On the other hand, the interactive program events, such as health meetings for mothers or group community tasks, strengthen the sense of community. Attanasio et al. (2015) also document the latter in Colombia: they show that the *Familias en Acción* program increased public good contributions in a lab-in-the-field public good game.⁶

Finally, the evidence shows that the positive gender gap in relative effort between the NLPG and LPG games for *Donors* decreases with the share of female participants in the experimental session. This result suggests that women might be aware of their higher willingness to exert effort for a public good that does not benefit them directly, compared to men. Consequently, women reduce their individual effort the more female their environment is.

Two caveats regarding the external validity of these results are (Levitt and List 2007): (i) in our experiment subjects undertake effort for a short time, while in reality productive efforts take place over extended periods; (ii) since the population we sampled resides within the village-based communal land system known as *ejido*, which involves mostly men, men’s pro-sociality towards their own village may be greater than in other population groups.⁷

This work contributes to the experimental literature on gender differences in economic experiments⁸ and the disincentive effects of taxes on effort. Several authors find a negative relationship between tax rates and effort exerted in real effort experiments (Keser et al. 2019; Lévy-Garboua et al. 2009; Sillamaa 1999; Sutter and Weck-Hannemann 2003; Swenson 1988). Among them, only Keser et al. (2019) looks at the impact that the use of tax revenue has on effort choices. Most of these studies do not look at gender differences in the effort responses to taxes, and none of them studies the interaction between gender and the destination of the tax revenue collected.⁹ Thus, to our knowledge, this experimental study is the first to examine gender difference in real-effort decisions that impact a real public good, with variation on whether subjects can benefit directly from the public good or not.¹⁰ In addition, all of previous experimental studies have been carried out in developed countries. Evidence for developing countries is scarce, but relevant given that raising taxes to finance public goods appears to be a challenge particularly in these countries (Besley and Persson 2014).

This work also adds to the growing literature on how social distance impacts giving in developing countries (Adena et al. 2020; Alt et al. 2018; Candelo et al. 2018). The research closest to our study is the lab-in-the-field experiment run in rural Mexico by Candelo et al. (2018). Their design uses the dictator game and varies the type of recipient across treatments. While they detect a more pronounced generosity towards family than non-family members, subjects are on average equally generous towards a random individual in another village or their own. Alt et al. (2018) also used the standard dictator game in an experiment with university students in Indonesia. Social distance was instilled artificially through the minimal group paradigm, and their design included both a

⁵ This finding is also consistent with the theory of reciprocity between citizens and the elite proposed by Besley (2020).

⁶ Other studies on the relationship between cash transfer programs, social capital and trust in the government are Nguyen and Rieger (2017) and Evans et al. (2019) for Morocco and Tanzania, respectively.

⁷ Relatedly, Carpenter and Seki (2011) find that exposure to teamwork had a positive impact on contributions in public goods games among fishermen.

⁸ Please refer to the surveys by Croson and Gneezy (2009), and Niederle (2017).

⁹ In contrast, the literature on tax compliance shows that there is heterogeneity in tax compliance behaviors (Mascagni 2018; Mohdali et al. 2014; Slemrod 2007). In particular, this literature finds that women report a much higher fraction of their earned income than men (Bruner et al. 2017; D’Attoma et al. 2018), and that this gap is particularly large when individuals do not benefit directly from the tax revenues collected (Bruner et al. 2017).

¹⁰ The large body of experimental work on monetary contributions made by subjects in artificial public goods games in developed countries has resulted in no clear gender differences (DellaVigna et al. 2013; Niederle 2017). For example, in experiments conducted on students with standard linear public goods games, Brown-Kruse and Hummels (1993) find that women make lower contributions than men on average, while Nowell and Tinkler (1994) find the opposite.

donor and a recipient treatment (or a “give” and a “take” treatment). In line with our findings, they find that donations were higher in the give than in the take treatment. However, none of these studies use real-effort decisions or examine gender differences.

2. The effort-and-transfer experiment

Participants played three games, all with the same initial endowment and real-effort task, and completed a post-experiment questionnaire.¹¹ The effort task consisted of using a pencil to mark the outline of pre-printed squares on a sheet of paper.¹² A subject’s effort determined the probability of generating additional income. The sheet had ten lines of empty squares, each corresponding to one “level” ($\ell = 1, \dots, 10$), which required ℓ squares to be marked, to make the marginal cost of effort increasing. For each completed level, the subject increased the probability of generating additional income by 1/11; the probability being zero if no level was completed. In the *Autarky* game, any additional income was simply added to the subject’s endowment. In the *Local Public Good* (LPG) game, it went to the health center in the subject’s village; in the *Non-Local Public Good* (NLPG) game, to the health center of another village. The empirical analysis considers only these two public good games. Details on the payoffs, the experimental protocol and procedures are presented in Online Appendix S1 and S4.

Our experimental design, which builds upon that in Alger et al. (2019), has several advantages. First, the subject’s effort choice in both public good games affects the resources allocated to the corresponding public good but not his/her own monetary payoff. Thus, this design avoids the income effect typically present in experiments using the classical public good or dictator games, where a subject’s contribution increases the monetary payoff for those who benefit from it and reduces his/her own. Second, because a 100% of the extra earnings were “taxed away”, a subject’s choice could be interpreted as her willingness to contribute real effort for the public good. Third, imposing the tax rate isolates the effort choice from the transfer choice, whereas in other experiments both are simultaneous. Finally, any potential social pressure to exert effort was eliminated by ensuring that individual decisions were unobservable by other participants (Perez-Truglia and Troiano 2016).

In each experimental session, the subject pool was divided randomly into two groups of equal size, usually around 20 participants. One group was allocated to the *Donor* treatment, in which the subject’s additional income, if materialized, was added to the total amount to be given to the health center. The other group was allocated to the *Recipient* treatment: if the subject failed to generate additional income, he/she received the equivalent amount from the resources to be given to the health center. The average payment was over 100 pesos, close to the daily minimum wage of an agricultural machinery operator.¹³

Online Appendix S3 presents a simple theoretical model of effort choice, in which individuals may be motivated to contribute to the public good because they may directly benefit from it, because of altruism towards other beneficiaries, and/or because of warm glow. This model shows that (1) effort in the NLPG versus the LPG game depends on the strength of the subject’s social ties in his/her own village versus the matched one (if such social ties are positive in the sense that they entail altruism towards other beneficiaries), and (2) effort is expected to be lower in the *Recipient* treatment than in the *Donor* one. The latter is due to the payoff structure, which implies a lower marginal benefit of effort in the *Recipient* treatment.

To study a population of subsistence farmers with low exposure to taxes, the experiment was conducted in 8 small rural localities (less than 2,000 inhabitants) in Southern Yucatán, Mexico, where such localities are common. Agriculture is highly heterogeneous across the state, but most of the producers, particularly in this region,

¹¹ In each session one of the games was randomly drawn to determine the payoffs of all the subjects in the session. A fourth game (the *One-to-One* game), was included in the experimental sessions but is not used in this paper.

¹² A variety of real effort tasks have been used in the literature, such as solving anagrams (Charness and Villeval (2009)), stuffing envelopes (Carpenter et al. (2010)), counting zeros (Abeler et al. (2011)), moving sliders (Gill and Prowse (2012)), and threading nuts onto bolts (Alger et al. (2019)). This is a manual task that does not require the use of computers. Given that the effort task is performed repeatedly, learning or fatigue may also arise in this experiment. Online Appendix S2 shows the robustness of the main results to the ordering of games.

¹³ In Mexico, there is no legal minimum wage for agricultural workers, but the minimum wage for an operator of agricultural machinery was 93.6 pesos per day at the time of the experiment (<http://www.conasami.gob.mx>).

operate at a small scale. Each experimental locality was matched with another similar locality for the NLPG. The characteristics of experimental subjects are comparable to those reported in the 2010 Mexican Population Census for their respective localities, with a few exceptions. Our pool seems to be relatively more female, less single, and more bilingual (Maya and Spanish), compared to the Census. This information, the match between the experimental localities and those that received the resources generated in the NLPG game and other descriptive statistics are in Online Appendix S1.

3. Descriptive evidence

We present graphical evidence on the distributions of effort differences between LPG and the NLPG games. Effort is measured as completed levels. For the whole sample, the average effort difference between the two games is negligible (7.27 versus 7.41 levels completed; Wilcoxon test $p=0.3496$).¹⁴

Figure 1 explores whether the distribution of the effort difference between those games varies systematically by role in the experimental session (*Donor vs Recipient*). As shown in the left panel, such distribution is practically identical for both (Wilcoxon test $p=0.9943$). The absence of a significant difference is compatible with a constant marginal benefit of donations, and/or systematic differences between the preferences of subjects in the *Recipient* and *Donor* treatments. This is because in the *Recipient* treatment the *smallest* possible donation to the health center is $100n$ (where n is the total number of subjects), whereas in the *Donor* treatment it is the *largest* possible donation that equals $100n$. Hence, if for subjects the marginal benefit of the donation is decreasing, *Recipients* would have been expected to exert a lower effort on average than *Donors*.

Figure 1. Distribution of the effort difference by Role (left) and by Gender (right).

The right panel of Figure 1 shows the distributions of the effort difference between games for men and women. The mean effort difference is 0.23 levels for women, whereas it is close to zero for men (0.02), so the gender gap in effort difference is positive and statistically significant ($0.23-0.02=0.21$ levels, Wilcoxon test $p=0.0365$). This result suggests that women are more willing than men to exert effort in the NLPG game, relative to the LPG one.

Figure 2 shows the distributions of the effort difference, separately by the interaction of role and gender. Female subjects seem to be more willing to exert effort than men in the NLPG versus the LPG game in both roles, as discussed before, but this gender gap is positive and significant, for *Donors* (0.5 levels, Wilcoxon test $p=0.0019$), whereas it is close to zero and not significant for *Recipients* (-0.08 levels, Wilcoxon test $p=0.9034$).

Figure 2. Distributions of effort difference by role and gender

In sum, women seem to be more willing than men to exert effort in the NLPG game, which benefits the health center of the matched locality, compared to the LPG game, which benefits the health center in their own locality. Furthermore, this gender gap in effort between those two games seems to be larger in the *Donor* treatment, but does not arise in comparisons by role alone.¹⁵

¹⁴ The higher average effort in the NLPG game than in the LPG game can be driven by learning (the NLPG game was always played last) and/or preferences (please see the theoretical section in Online Appendix S3). The interest of this study is, however, in the gender NLPG-LPG effort gap rather than in the absolute effort levels. As discussed later, the main results on this gap are robust to the inclusion of a learning proxy (the game ordering in the session).

¹⁵ Table S2.1 in Online Appendix S2 shows the mean and standard deviation of effort choices at each one of the games and these same patterns.

4. Main Results

Empirical Specification

To examine how the effort difference between the LPG and NLPG games varies with gender, the following specification was estimated:

$$\text{Effort}_{i,g} = \alpha + \beta \cdot \text{Non-Local}_g + \delta \cdot (\text{Non-Local}_g * \text{Female}_i) + \lambda_i + \epsilon_{ig} \quad (1)$$

where $\text{Effort}_{i,g}$ denotes the effort of subject i in game g ; Non-Local_g is a dummy equal to one for the transfer game that benefits the health center of the matched locality, and zero for the LPG game; the next term is the interaction of the *Non-Local* dummy with a *Female* dummy; λ_i is a subject fixed effect, which controls for individual characteristics affecting effort that remained constant during the session, and ϵ_{ig} is an error term. The coefficient β captures the mean difference in effort between the NLPG and LPG games for men, whereas $\beta + \delta$ captures the corresponding one for women. Thus, δ represents the gender gap in effort difference.

Equation 1 is estimated separately for *Donors* and *Recipients*. In all estimations, the standard errors are clustered at the session level to account for any within-session correlation. Given the small number of clusters (8), p-values obtained using a wild cluster bootstrap (WCB) procedure, with Webb's (2013) recommended weights, are also reported.

Regression results

Table 1 presents the main results. The *Non-Local* dummy alone is negative for *Donors* in Column 1, and positive for *Recipients* in Column 2, but none of them are statistically significant. Thus, male *Donors* and *Recipients* do not seem to put forth a significantly different effort in the two public good games. In contrast, the interaction of the *Non-Local* dummy with the *Female* dummy is positive and significant at 1 percent for *Donors* in Column 1, whereas it is small and not statistically significant for *Recipients* in Column 2. On average, female *Donors* complete about half a level (0.508) more than male *Donors* in the NLPG, relative to the LPG game. This gender gap amounts to 7 percent of the mean effort in the NLPG game (7.41 levels).

Thus, female *Donors* appear to be more willing than their male counterparts to exert effort when the transfer benefits the other village instead of their own. No comparable gender gap in relative effort is found for *Recipients*. As shown at the bottom of the table, the statistical significance of all the estimates in Table 1 remains the same when using the WCB procedure to account for the small number of clusters, except for the *Non-Local* dummy in column 1, which becomes significant at 5 percent.

Table S2.2 in Online Appendix S2 shows that the main results are similar when replacing individual fixed effects with a combination of session fixed effects and individual controls. In all sessions, the last game was always the NLPG game, whereas the LPG game was either the second or third game, depending on the session. To check whether the results are driven by learning, interactions with a dummy for the game order in the session were added to Equation 1. Table S2.2 confirms that the main results hold when controlling for the order of the games. If anything, the positive gender gap in effort for *Donors* is a bit smaller in magnitude, suggesting ordering might explain some of it but not all.

5. What explains the gender difference in effort between the NLPG and LPG games?

This section analyzes whether the positive gender gap in the effort difference between games for *Donors* varies, with measures of informal giving and receiving, community involvement and trust, participation in government programs, strength of the participant's network in her own and in the matched village, and other sociodemographic characteristics at the individual level.

In particular, the interest is to see whether women's relative willingness to exert effort in the NLPG versus the LPG game is related to their informal giving and receiving behavior. For this, dummy variables for whether the participant's household gave or received any help to or from other households in the previous year are considered.

Women could also be more willing to exert effort for others because they are more socially involved in the community and have higher trust in institutions, compared to men. For community involvement and trust, sep-

arate PCA indices were constructed. The community involvement index includes dummies for whether the respondent believes the people in her locality cooperate to solve their problems most of the time, and the number of organizations and community celebrations in which she participated in the last 12 months. The trust index includes dummies for whether the respondent believes (i) she can trust most people, (ii) most people would treat her fairly, (iii) effort is rewarded with higher income most of the time, (iv) the government spends tax revenues adequately. Given the forced nature of the transfer in our experiment, we also consider variable (iv) separately.

The main results could also be driven by prosocial preferences, such as altruism (Becker 1974). If so, a subject's effort towards the health center in her own locality, or the matched one, could be impacted by the strength of her network in each of them. We consider the following variables as proxies for the participant's network in her own locality: dummies for whether the participant knows almost all or all of the other participants in the session, whether half or more of the session participants belong to her family, whether she shares daily activities with half or more of them. The proxies for the strength of the participant's network in the matched locality are dummies for whether she knows the other locality, whether she has visited the other locality and for whether she has actually been in the health clinic of the matched locality.

This part of the analysis also considers other standard sociodemographic characteristics of subjects, like age, dummies for having at least secondary education, for being married, and for being born in the municipality where the session took place; number of siblings and number of children. Participation in government programs could also affect women's willingness to exert effort for the other village, either because of a crowding out or crowding in effect. Thus, we use dummies for whether someone in the participant's household receives Procampo or Oportunidades (the largest cash transfer programs targeting rural localities), and the number of government programs in which her household participates.¹⁶ Finally, the analysis explores whether the gap varies by household wealth by using dummies for having an asset index below the 25th percentile or above the 75th percentile.

For each of these variables, the empirical specification is the following:

$$\text{Effort}_{ig} = \alpha + \beta * \text{Non-Local}_g + \gamma_1 * \text{Non-Local}_g * \text{Female}_i + \gamma_2 * \text{Non-Local}_g * \text{Variable}_i + \gamma_3 * \text{Non-Local}_g * \text{Female}_i * \text{Variable}_i + \delta_i + \epsilon_{ig} \quad (2)$$

where most terms are defined as above. The coefficient γ_1 measures the part of the gender gap in the relative effort exerted in the LPG versus the NLPG games that is common to all women, whereas γ_3 measures the part that varies with the value of Variable_i , the particular characteristic examined in each specification.

Table S2.3 in Online Appendix S2 shows the unconditional mean differences by gender for all the variables used in the triple interactions of equation 2. Most of these differences are not statistically significant. Informal giving and receiving patterns (Panel A), the participant's trust in government (Panel B) and strength of her social networks in both localities (Panel C) do not seem to differ by gender. Conversely, on average, women seem to be significantly more socially involved in their community, but also display less trust towards other people, compared to men (Panel B).

Most of the unconditional mean differences in standard sociodemographic characteristics by gender are not statistically significant (Panel D), except for the probability of having at least some secondary education, the probability of being married, and the participant's number of children. About 48 percent of men have at least some secondary education, compared to only 26 percent of women, and this difference is significant at 1 percent. Women are 10 percentage points more likely to be married than men, and this difference is marginally significant. On average, female participants have more children (3.8) than male ones (3.0), and this difference is significant at 5 percent. Note that if women care more about their own children's welfare than men, they could be more willing to exert effort in the LPG game. However, given that the additional earnings in the NLPG game do not benefit the subjects' own children directly, this discrepancy in parental altruism could hardly explain the gender gap in relative effort for *Donors*.

¹⁶ In the sample, about 62 and 85 percent of subjects report that their household receives support from Procampo and Oportunidades, respectively. On average, the subjects' households receive support from 2.5 government programs.

Tables 2 and 3 show the estimation results for equation 2 using some of the above variables. For brevity, the tables report results only for informal giving and receiving, community involvement, trust, participation in government programs and household wealth. Results for other social network and sociodemographic variables are in Online Appendix S2. Overall, no significant patterns explain the main results in these excluded variables.

In Table 2, most of the main results remain after including measures of informal giving and receiving, and the community involvement and trust indices (Columns 1 to 4). In those columns, the estimate for $Non - Local_g * Female$ for *Donors* is positive after including each of these variables, and mostly statistically significant at conventional levels (Panel A). For *Donors*, the estimates of the triple interactions are not significant in columns 1 to 4, suggesting that the gender gap in the effort difference between the NLPG and LPG games does not vary with these indicators. The only exception is the triple interaction with the dummy for having some trust in government spending taxes adequately (column 5), which is positive but marginally significant, whereas the estimate of γ_1 in that column turns insignificant. This evidence suggests that, for *Donors*, most of the positive gender gap in effort between the two public good games could be due to the behavior of women who have some trust in the government.

In Panel B, most of the estimates for $Non - Local_g * Female$ are not significant for *Recipients*, as in the main results. Some triple interactions are marginally significant, suggesting that women in this role provided relatively more effort in the NLPG versus the LPG games if, for instance, their household gave help, and less the higher their trust is overall and in the government in particular.

The estimates for participation in government programs and household wealth are in Table 3. For *Donors*, panel A shows that the positive gender gap in effort between the two public good games is due to women who participate in the Oportunidades program, a conditional cash transfer program for poor rural households. In column 1, γ_1 is not significant, and it is smaller than the corresponding estimates in other columns, whereas γ_3 is positive and significant at 10 percent. Oportunidades, formerly known as Progresa, has two features that are consistent with these results. First, the transfer from the program is paid to the mother or most senior woman in the household. Second, part of the transfer is conditioned on complying with scheduled health checks for children and adults in beneficiary households. Thus, Oportunidades beneficiaries, women who use health services relatively often, are those who are more willing to exert effort for the clinic at the matched locality, compared to men. Oportunidades also has other activities that female beneficiaries do together, which could further strengthen their social ties. No similar patterns are observed for the receipt of noncontributory pensions, which are unconditional cash transfers paid to the rural elderly, regardless of their gender (column 8). Column 9 shows that the positive effort gender gap is smaller for *Donors* who participate in the Procampo program. These are probably women whose household owns land, because this program pays a transfer per hectare of land held and sown in a given agricultural cycle. Column 10 shows that participation in more government programs decreases the relative willingness of women to exert effort for the other locality, compared to men. For *Recipients*, Panel B shows once again no significant estimates for γ_1 and γ_3 . Finally, the interactions with an indicator for low or high household wealth are not significant in both panels.

In sum, the positive gender gap in effort between the two public good games does not vary with informal giving and receiving patterns, measures of community involvement and trust, proxies for the strength of the local and non-local networks of players, or other standard sociodemographic characteristics at the individual level, except for the trust of the participant in the way the government spends taxes, and the participation in government programs. Specifically, the Oportunidades result suggests that some government programs could actually foster social capital and a willingness to contribute to a public good, instead of crowding them out. Most of the literature on anti-poverty and social insurance programs has tended to focus on the substitutability between government and private interventions. However, as discussed in the introduction, this research adds to other studies that, instead, find some positive effects of public cash transfers on social capital, trust in the government, and public good contributions in developing countries.

Are women aware of their higher willingness to exert effort for others?

The final piece of analysis examines whether the gender of other participants affects the willingness of women to exert effort in the LPG versus NLPG game. This involves estimating equation 2 including a triple interaction of *Non-Local* and *Female * Variable* with each of these two variables: the share of women in the experimental session and in the room. The first measure pools *Donors* and *Recipients* in a given session when calculating the share, whereas the second captures the gender composition within each of these treatments. Sessions had on average 59% of female participants, with a minimum of 43% and a maximum of 74%, and the distribution in the two treatments in each session is similar.

Table 4 continues to show a positive and significant gender gap in effort for *Donors*, which decreases significantly with the share of women in the session, but not in the room. These findings suggest that women reduce their individual effort the more female their environment is. A possible explanation is that they know of their higher willingness to exert effort for a public good that does not benefit them directly, compared to men. In this case, women might react more to the gender composition of the whole session and not of the room because they were more aware of the former. Recall that participants hung out for a while outside of the room, while the research team waited for participants to arrive and carried out the final preparations for the session, and so they had plenty of time to look at each other. After this, they were randomly divided into two groups and the first one immediately entered the room, where participants were all the time facing the front of the room and had dividers between them to ensure privacy.

6. Concluding discussion

Are women more inclined than men to contribute to public goods from which they do not benefit personally? The data from a lab-in-the-field experiment, in which subjects choose to exert real effort to increase the contributions to health centers existing in other localities and their own, suggests that this may be the case. This study finds a positive and statistically significant gender gap in the difference between effort in the LPG and NLPG games for *Donors*, whose effort could have the effect of adding money to the amount given to the health centers. This implies that female *Donors* are more willing than their male counterparts to exert effort when it benefits the health center of another locality, compared to when it benefits the one in their own locality. In contrast, the gender effort gap between those two games for *Recipients*, whose effort could have the effect of avoiding to withdraw money from the contributions to health centers, is close to zero and not statistically significant. Although this work does not formally establish the external validity of these results, the gender gap in relative effort for *Donors* is mostly due to women who report having at least some trust in the way the government spends taxes (see Section 5B), and to women who benefit from a government program that targets women and fosters health care use (see Section 5D), which lends credibility to them.

As mentioned in the introduction, the evidence on gender differences regarding monetary contributions in the classic public goods game in laboratory experiments is mixed. By contrast, studies that examine gender differences in the willingness to pay taxes show that women's willingness exceeds that of men more consistently. This study complements these two strands of the literature by using a real-effort experiment in which proceeds went to actual public goods. As a caveat, the motivation of an individual for complying with the tax code might be different from the one for voluntarily exerting effort to contribute to a public good (s)he does not benefit from personally. The fear of being caught and punished might play a role in the former, whereas not in the latter, especially when the effort choice is unobservable to others, as in this study's experiment. Thus, from the results, it cannot generally be concluded that women are more willing than men to pay taxes in the studied population. Nevertheless, the evidence suggests that it would be interesting to disentangle the relative contribution of these two motivations to the gender difference in tax compliance found in the literature. Furthermore, while discrepancies in the behavior of women and men may be driven by social norms and culture (Fernández 2013), evolutionary theory suggests that they may also stem from deeper differences (Eswaran 2014). These findings are in line with the evolutionary theory which predicts that in- vs out-group attitudes are expected to differ between men and women, with men being expected to favor in-group members to a larger extent than women (Low 1993; Manson

et al. 1991; Micheletti et al. 2018).

However, the finding that the gender effort gap is mostly due to women who trust the government and/or benefit from a government program that targets women and fosters health care use, is also in line with a recent theory according to which reciprocity can drive willingness to pay taxes (Besley 2020). Further research on the competing explanations for gender differences in behavior is necessary, including those concerning their willingness to contribute towards real public goods (as *Donors*) or benefit from them (as *Recipients*), as suggested by the main results in this work.

References

- Abeler, J., Falk, A., Goette, L., and Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, 101(2):470–92.
- Adams, R. B. and Funk, P. (2012). Beyond the glass ceiling: Does gender matter? *Management Science*, 58(2):219–235.
- Adato, M. (2000). The impact of progressa on community social relationships. Technical report.
- Adena, M., Hakimov, R., and Huck, S. (2020). Charitable giving by the poor: A field experiment in kyrgyzstan. WZB Discussion Paper SP II 2019-305r, Berlin.
- Alger, I., Miquel-Florensa, J., Juarez, L., and Juarez-Torres, M. (2019). Do informal transfers induce lower efforts? evidence from lab-in-the-field experiments in rural mexico. *Economic Development and Cultural Change*.
- Alt, M., Gallier, C., Schlüter, A., Nelson, K., and Angraini, E. (2018). Giving to versus Taking from In- and Out-Group Members. *Games*, 9(3):57.
- Angelucci, M., De Giorgi, G., Rangel, M., and Rasul, I. (2009). Village economies and the structure of extended family networks. *The BE Journal of Economic Analysis & Policy*, 9(1).
- Angelucci, M., De Giorgi, G., Rangel, M. A., and Rasul, I. (2010). Family networks and school enrolment: Evidence from a randomized social experiment. *Journal of public Economics*, 94(3-4):197–221.
- Attanasio, O., Polania-Reyes, S., and Pellerano, L. (2015). Building social capital: Conditional cash transfers and cooperation. *Journal of Economic Behavior & Organization*, 118:22–39.
- Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *science*, 335(6068):582–586.
- Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy*, 82(6):1063–1093.
- Bertrand, M. (2011). New perspectives on gender. In *Handbook of labor economics*, volume 4, pages 1543–1590. Elsevier.
- Besley, T. (2020). State capacity, reciprocity, and the social contract. *Econometrica*, 88(4):1307–1335.
- Besley, T. and Persson, T. (2014). Why do developing countries tax so little? *Journal of Economic Perspectives*, 28(4):99–120.
- Brown-Kruse, J. and Hummels, D. (1993). Gender effects in laboratory public goods contribution: Do individuals put their money where their mouth is? *Journal of Economic Behavior & Organization*, 22(3):255–267.
- Bruner, D. M., D’Attoma, J., and Steinmo, S. (2017). The role of gender in the provision of public goods through tax compliance. *Journal of Behavioral and Experimental Economics*, 71:45–55.
- Candelo, N., Eckel, C., and Johnson, C. (2018). Social Distance Matters in Dictator Games: Evidence from 11 Mexican Villages. *Games*, 9(4):77.
- Carpenter, J., Matthews, P. H., and Schirm, J. (2010). Tournaments and office politics: Evidence from a real effort experiment. *American Economic Review*, 100(1):504–17.
- Carpenter, J. and Seki, E. (2011). Do social preferences increase productivity? field experimental evidence from fishermen in toyama bay. *Economic Inquiry*, 49(2):612–630.
- Charness, G. and Villeval, M.-C. (2009). Cooperation and competition in intergenerational experiments in the field and the laboratory. *American Economic Review*, 99(3):956–78.
- Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, 47(2):448–74.
- D’Attoma, J., Volintiru, C., and Malezieux, A. (2018). Gender, Social Value Orientation, and Tax Compliance. Technical report.
- DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2013). The importance of being marginal: Gender differences in generosity. *American Economic Review*, 103(3):586–90.
- Dupas, P. and Robinson, J. (2013). Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics*, 5(1):163–192.
- Eswaran, M. (2014). *Why Gender Matters in Economics*. Princeton University Press, Princeton, NJ.
- Evans, D. K., Holtemeyer, B., and Kosec, K. (2019). Cash transfers increase trust in local government. *World Development*,

- 114:138–155.
- Fernández, R. (2013). Cultural Change as Learning : The Evolution of Female Labor Force Participation over a Century. *American Economic Review*, 103(1):472–500.
- Field, E., Holland, A., Pande, R., et al. (2016). Microcredit: Points of promise. Technical report.
- Gertler, P., Patrinos, H., and Rubio-Codina, M. (2006). *Empowering parents to improve education: Evidence from rural Mexico*. The World Bank.
- Gill, D. and Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, 102(1):469–503.
- Keser, C., Masclet, D., and Montmarquette, C. (2019). Labor supply, taxation and the use of the tax revenues: A real-effort experiment in Canada, France, and Germany. *CEPR Discussion Papers 377*, Göttingen.
- Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2):153–174.
- Low, B. S. (1993). An evolutionary perspective on war. In Jacobson, H. and Zimmerman, W., editors, *Behavior, Culture, and Conflict in World Politics*, pages 13–56. University of Michigan Press, Ann Arbor, MI.
- Lévy-Garboua, L., Masclet, D., and Montmarquette, C. (2009). A behavioral Laffer curve: Emergence of a social norm of fairness in a real effort experiment. *Journal of Economic Psychology*, 30(2):147–161.
- Manson, J. H., Wrangham, R. W., Boone, J. L., Chapais, B., Dunbar, R., Ember, C. R., Irons, W., Marchant, L., McGrew, W., Nishida, T., et al. (1991). Intergroup aggression in chimpanzees and humans [and comments and replies]. *Current Anthropology*, 32(4):369–390.
- Mascagni, G. (2018). From the lab to the field: A review of tax experiments. *Journal of Economic Surveys*, 32(2):273–301.
- Meghir, C. and Phillips, D. (2008). Labour supply and taxes. *Working Paper Series*.
- Micheletti, A. J., Ruxton, G. D., and Gardner, A. (2018). Why war is a man's game. *Proceedings of the Royal Society B: Biological Sciences*, 285(1884):20180975.
- Millán, T. M., Barham, T., Macours, K., Maluccio, J. A., and Stampini, M. (2019). Long-term impacts of conditional cash transfers: review of the evidence. *The World Bank Research Observer*, 34(1):119–159.
- Mohdali, R., Isa, K., and Yusoff, S. H. (2014). The impact of threat of punishment on tax compliance and non-compliance attitudes in Malaysia. *Procedia-Social and Behavioral Sciences*, 164:291–297.
- Nguyen, T. C. and Rieger, M. (2017). Community-driven development and social capital: Evidence from Morocco. *World Development*, 91:28–52.
- Niederle, M. (2017). A gender agenda: a progress report on competitiveness. *American Economic Review*, 107(5):115–19.
- Nowell, C. and Tinkler, S. (1994). The influence of gender on the provision of a public good. *Journal of Economic Behavior & Organization*, 25(1):25–36.
- Perez-Truglia, R. and Troiano, U. (2016). Shaming tax delinquents: Evidence from a field experiment in the United States. *SSRN Scholarly Paper. Rochester, NY: Social Science Research Network*, 1.
- Sillamaa, M. (1999). How work effort responds to wage taxation: An experimental test of a zero top marginal tax rate. *Journal of Public Economics*, 73(1):125–134.
- Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives*, 21(1):25–48.
- Summers, L. H. (1989). Some simple economics of mandated benefits. *American Economic Review*, 79(2):177–183.
- Sutter, M. and Weck-Hannemann, H. (2003). Taxation and the veil of ignorance: A real effort experiment on the Laffer curve. *Public Choice*, 115(1/2):217–240.
- Swenson, C. (1988). Taxpayer behavior in response to taxation. An experimental analysis. *Journal of Accounting and Public Policy*, 7:1–28.

Main Tables

Table 1. Effort exerted in NLPG and LPG games

Sample:	(1) Donors	(2) Recipients
Non-Local	-0.1667 (0.097)	0.1923 (0.263)
Non-Local * Female	0.5088*** (0.074)	-0.0747 (0.301)
Constant	7.2016*** (0.048)	7.3417*** (0.059)
Observations	248	240
R^2	0.109	0.014
Number of players	124	120
H_0 : Non-Local=0 (WCB p-value)	0.050	0.549
H_0 : Non-Local*Female=0 (WCB p-value)	0.001	0.856

Source: Data for the study comes from the lab-in-the-field experiment conducted by the authors in 8 rural villages in the southern part of Yucatan, Mexico, in August 2014.

Notes: Robust standard errors in parentheses, cluster session. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Unit of observation is a player's effort choice in a game. Reference LPG game. Subject fixed effects included in all specifications. P-values reported at the bottom of the table were obtained using a wild cluster bootstrap (WCB) with Webb's Weights.

Table 2. Interaction with informal giving and receiving, trust, and community involvement

Panel A: Donors					
Variable interacted:	(1) HH giving	(2) HH received	(3) Community index	(4) Trust Index	(5) Trust gov
Non-Local	0.0000 (0.155)	0.0455 (0.213)	-0.184* (0.092)	-0.1400 (0.107)	0.0833 (0.235)
Non-Local * Female	0.3571** (0.143)	0.3260 (0.185)	0.538*** (0.080)	0.486*** (0.063)	0.1944 (0.162)
Non-Local * Variable	-0.0690 (0.196)	-0.3916 (0.278)	-0.083 (0.171)	-0.117** (0.035)	-0.3333 (0.196)
Non-Local * Female * Variable	0.0354 (0.170)	0.3202 (0.318)	0.039 (0.178)	0.087 (0.090)	0.4240* (0.220)
Constant	7.1639*** (0.054)	7.2033*** (0.049)	7.202*** (0.049)	7.203*** (0.049)	7.2033*** (0.048)
Observations	244	246	248	246	246
R ²	0.133	0.125	0.115	0.122	0.122
Number of player	122	123	124	124	123
Panel B: Recipients					
Variable interacted:	(1) HH giving	(2) HH received	(3) Community index	(4) Trust Index	(5) Trus gov.
Non-Local	0.6923 (0.418)	0.4286 (0.542)	0.2200 (0.234)	0.1580 (0.317)	-0.8000 (0.582)
Non-Local * Female	-0.6611 (0.577)	-0.2394 (0.549)	-0.1020 (0.274)	-0.1260 (0.408)	1.1684* (0.597)
Non-Local * Variable	-1.0673** (0.434)	-0.2562 (0.584)	0.1730 (0.127)	0.1140 (0.232)	1.2286* (0.639)
Non-Local * Female * Variable	1.2482* (0.598)	0.0670 (0.431)	-0.2030 (0.150)	-0.392* (0.191)	-1.5766* (0.709)
Constant	7.3304*** (0.049)	7.2931*** (0.046)	7.3420*** (0.056)	7.3420*** (0.055)	7.3417*** (0.044)
Observations	230	232	240	240	240
R ²	0.088	0.035	0.025	0.049	0.084
Number of player	115	116	120	120	120

Source: Data for the study comes from the lab-in-the-field experiment conducted by the authors in 8 rural villages in the southern part of Yucatan, Mexico, in August 2014.

Notes: Robust standard errors in parentheses, cluster session. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Unit of observation is a player's effort choice in a game. Reference LPG game. Subject fixed effects included in all specifications. As measures of giving and receiving patterns, we use dummy variables for whether the participant's household gave any help to other households in the previous year (Column 1), whether help was given to someone in the respondent's locality or in another locality (Column 2). An index of community involvement measures based on cooperation, whether the respondent believes the people in her locality cooperate to solve their problems most of the time, the number of organizations in which the respondent participated in the last 12 months (social, political, religious), and the number of community celebrations in which she participates (Column 3). An index of community trust that measures if player can trust most people, whether most people would treat her fairly, and whether effort is rewarded with higher income most of the time (Column 4). Dummies for whether she has some confidence that the government will adequately spend tax revenues, and for whether she has high confidence on the same issue (Column 5).

Table 3. Interaction with participation in government programs

Panel A: Donors				
Variables interacted:	(1) Oportunidades	(2) Elderly pension	(3) Procampo	(4) Count programs
Non-Local	-0.1000 (0.173)	-0.1628 (0.087)	-0.2000 (0.149)	-0.4590** (0.141)
Non-Local * Female	0.1000 (0.173)	0.5128*** (0.100)	0.7312*** (0.085)	0.9667*** (0.168)
Non-Local * Variable	-0.0842 (0.152)	-0.0372 (0.135)	0.0571 (0.143)	0.1362** (0.041)
Non-Local * Female * Variable	0.4556* (0.224)	-0.0003 (0.268)	-0.3838** (0.115)	-0.2028** (0.068)
Constant	7.2016*** (0.047)	7.2016*** (0.047)	7.2016*** (0.048)	7.2016*** (0.047)
Observations	248	248	248	248
R^2	0.118	0.109	0.130	0.133
Number of player	124	124	124	124
Panel B: Recipients				
Variables interacted:	(1) Oportunidades	(2) Elderly pension	(3) Procampo	(4) Count programs
Non-Local	0.2500 (0.165)	0.3684 (0.385)	-0.0714 (0.109)	0.0639 (0.356)
Non-Local * Female	-0.2500 (0.163)	-0.2498 (0.429)	0.1914 (0.365)	0.0923 (0.519)
Non-Local * Variable	-0.0682 (0.195)	-0.6541 (0.463)	0.3609 (0.257)	0.0439 (0.043)
Non-Local * Female * Variable	0.2085 (0.352)	0.6466 (0.547)	-0.3646 (0.321)	-0.0596 (0.109)
Constant	7.3417*** (0.059)	7.3417*** (0.062)	7.3417*** (0.056)	7.3417*** (0.057)
Observations	240	240	240	240
R^2	0.015	0.036	0.021	0.016
Number of player	120	120	120	120

Source: Data for the study comes from the lab-in-the-field experiment conducted by the authors in 8 rural villages in the southern part of Yucatan, Mexico, in August 2014.

Notes: Robust standard errors in parentheses, cluster session. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Unit of observation is a player's effort choice in a game. Reference LPG game. Subject fixed effects included in all specifications. Dummies for whether someone in the participant's household receives Oportunidades (Column 1), Rural elderly program (Column 2) or Procampo (Column 3), and the number of government programs that the participant's household receives benefits from (Column 4) are also considered. With respect to income, dummies for having a household asset index in lower than the 25th percentile or higher than the 75th percentile are included (Columns 5 and 6, respectively).

Table 4. Interaction with share of women in the room and session

	(1)		(2)	(3)	(4)	
	Session composition			Donors	Room composition	
	Donors	Recipients		Donors		Recipients
Non-Local game	-0.9531 (0.611)	1.2589 (1.674)		-0.7279 (0.431)		0.6721 (0.551)
Non-Local game * female	1.2443*** (0.279)	-0.6412 (2.046)		0.7856 (0.442)		-0.2909 (0.921)
Non-Local * Session share	1.3928 (0.967)	-1.9378 (2.839)				
Non-Local * female * Session share	-1.3069** (0.432)	1.1110 (3.282)				
Non-Local * Room share				1.0013 (0.642)		-0.9850 (1.017)
Non-Local * female * Room share				-0.5528 (0.606)		0.5548 (1.355)
Constant	7.2016*** (0.047)	7.3417*** (0.055)		7.2016*** (0.046)		7.3417*** (0.055)
Observations	248	240		248		240
R^2	0.119	0.026		0.120		0.023
Number of player	124	120		124		120

Source: Data for the study comes from the lab-in-the-field experiment conducted by the authors in 8 rural villages in the southern part of Yucatan, Mexico, in August 2014.

Notes: Robust standard errors in parentheses, cluster session. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Unit of observation is a player's effort choice in a game. Reference LPG game. Subject fixed effects included in all specifications.

The variable *Session share* denotes the share of female participants in the session. The variable *Room share* denotes the share of females in the room, i.e. in the *Donor* or *Recipient* treatment. Given that allocation to treatments is random, both shares can differ in a given session.