Original Article

Tarek Jaber-Lopez*, Alexandra Baier, and Brent J. Davis

In-group, out-group effects in distributional preferences: the case of gender

https://doi.org/10.1515/ger-2019-0119

Abstract: We examine gender differences when eliciting distributional preferences as conducted by the Equality Equivalence Test, which has the ability to classify subjects into preferences types. Preferences are elicited when individuals interact with an individual of the same gender and with an individual of the opposite gender. We find elicited preferences are robust across both in-group (same gender) and out-group (opposite gender) interactions. When analyzing the intensity of benevolence (or malevolence) we find that overall women exhibit more malevolence than men, but there is no gender difference for benevolence. Furthermore, women exhibit a higher level of in-group favoritism than men.

Keywords: Distributional preferences, gender, experiment, in-group-out-group

1 Introduction

A plethora of work in economics and psychology has documented the importance of social preferences in various environments. As economic science continues to improve in its methodology in eliciting social preferences, both in measuring preference type and intensity, it then becomes relevant to understand if such methodology produce consistent measurements across different economic environments. Recent experimental research has investigated the robustness and stability of individual preferences, examples include time preferences (Meier and Sprenger, 2015), risk (Andersen et al., 2008), charitable preferences (Benz and Meier, 2008), and cooperative preferences (de Oliveira et al., 2012; Volk et al.,

***Corresponding author: Tarek Jaber-Lopez,** EconomiX, UPL, Univ Paris Nanterre, CNRS, F92000 Nanterre, France; and Department of Public Finance, University of Innsbruck, Universitätsstrasse 15, 6020 Innsbruck, Austria, e-mail: tarek.jl@parisnanterre.fr, ORCID: https://orcid.org/0000-0002-8361-3261
Alexandra Baier, Brent J. Davis, Department of Public Finance, University of Innsbruck, Universitätsstrasse 15, 6020 Innsbruck, Austria, e-mails: alexandra.baier@uibk.ac.at, bdavis477@gmail.com

2012) among others. In this paper, we examine gender effects of elicited distributional preferences across in- and out-groups via the Equality Equivalence Test (EET) with a gender domain in a laboratory experiment under the assumption that the EET is a stable measurement tool.

Why is examining the stability of individual preferences in a group context important? Groups, in various forms, are prevalent in nearly all aspect of society. Much research has investigated how (non)membership in groups affects decision making in various economic environments. A large portion of this research examines the willingness of group members to engage in pro-social behavior, and whether or not group members exhibit in-group favoritism and/or out-group bias. Identification with groups can elicit different behavior from individuals (Akerlof and Kranton, 2000), and even when the groups are induced in the laboratory through minimal group paradigm.[1] Due to this effect on behavior, and that individuals may form distributional preferences in the context of groups (or the measurement of distributional preferences cannot be done without a group context), we ask how, or if, elicited distributional preferences are affected by group identification, which in our paper is gender.

We examine gender differences and in-group-out-group bias in distributional preferences. Several different experimental tests to identify distributional preference types have been used by economists since the 1990s.[2] The distributional preferences archetype and test we examine, the Equality Equivalence Test (henceforth – EET), was developed by Kerschbamer (2015). We acknowledge that other distributional preferences tests are equally valid to answer our research question but we chose the EET test for the reasons described below. Our study is new in that we elicit social preferences from subjects when they play with an individual of the same gender and that of the opposite gender. This test has been used in experiments on bidding behavior (Flynn et al. 2016), ego depletion (Balafoutas et al., 2018), political attitudes (Kerschbamer and Muller, 2020), and competitive preferences (Balafoutas et al., 2012). As this test has been used to understand behavior in a several different economic environments it is therefore important to understand how these preferences change (or not change) when individuals interact with their in- and out-group. If individual preference types vary by in-group and out-group context in the distribution test we examine herein, then this would raise

1 Some prominent examples of the minimal group paradigm in experimental economics are Sutter (2009), Hargreaves-Heap and Zizzo (2009), Chen and Li (2009), and Goette et al. (2012).

2 Some examples are Charness and Rabin (2002), Cabrales et al. (2010), and Iriberri and Rey-Biel (2013). The ring-test was developed by Griesinger and Livingston (1973) and Liebrand (1984) to study social value orientations and later employed by economists to examine distributional preferences. See for example, Offerman et al. (1996), Brandts et al. (2009), and Sutter et al. (2010).

additional questions regarding the robustness of social preferences across contexts and potentially for other distributional tests used by the field. To our knowledge, this variation when performing the EET has not been explored. Although explaining the advantages of using the EET test is beyond the scope of this study, we would like to quote some of the advantages mentioned in the original paper by Kerschbamer (2015), page 94:

> (i) that it is simple and short as subjects' task is to make a small set of diagnostic choices without feedback;[...] (iv) that it is flexible as test size and test design can easily be fine-tuned to the research question of interest; (v) that it is precise as it identifies the archetypes of distributional concerns with arbitrary precision and also gives an index of preference intensity.

We chose to use the EET for several reasons. First, the EET provides information at the individual level and is non-parametric. This test has additional advantages over other methods, (such as a dictator game). For instance, it allows us to classify subjects in types, such as altruistic, selfish, inequity averse, and others. Moreover, this test provides two measures of social preferences, the x-score and the y-score, revealing information about individuals' benevolence (or malevolence) in the domains of disadvantageous and advantageous inequality, respectively. Finally, we conduct a within-subjects design which has some limitations, i. e. demand effect, but has also other advantages such as individual comparison. Furthermore, all the decisions were made simultaneously. In Section 2.1, we provide a further description of the EET.

The literature on social identity is well established going back to Tajfel and Turner (1979) in social psychology and Akerlof and Kranton (2000) in economics. In economics, gender differences have been studied in many contexts including risk preferences, social preferences and competitive preferences.[3] Given the characteristics of the game employed in our experiment, i. e. income allocations, we are mainly interested in the literature related to dictator games and/or giving behavior in the context of gender differences. The results of the literature are mixed but a general finding is that women seem to be more sensitive to the context than men (Croson and Gneezy, 2009). For instance, it has been shown that in a full anonymous dictator game, women give almost twice as much as men (Eckel and Grossman, 1998) whereas no gender differences in generosity are found in a setting with lower anonymity (Bolton and Katok, 1995).

3 Croson and Gneezy (2009) review the literature on gender differences in the aforementioned three contexts, i. e. risk preferences, social preferences, and competitive preferences.

In the context of allocation decisions, women are found to show systematically more inequality aversion preferences (Andreoni and Vesterlund, 2001; Dickinson and Tiefenthaler, 2002; Selten and Ockenfels, 1998; Müller, 2017). Other studies closer to ours focus on the giving behavior when the gender of the recipient is known. Dufwenberg and Muren (2006) find no differences in giving behavior regardless of the gender of the recipient, although overall men receive less than women do. Ben-Ner et al. (2004) find no differences when the gender of the recipient is unknown but find women give significantly less to another woman than to a man or to an unknown gender, while they find no differences for men. In a meta-analysis study on discrimination, Lane (2016) concludes that the identity category of gender leads to the lowest discrimination (i. e. gender discrimination is weaker than artificial groups, ethnic, and national discrimination) and finds significant out-group favoritism in gender studies (i. e. females favor males and males favor females).

Our results show that the proportion of preference types is constant across gender and across both in-group and out-group interactions. These results suggest that the EET test is robust across multiple environments, when in-group and out-group effects may be of a concern (or at a minimum in a laboratory environment in a gender context). These results are somewhat different from earlier experiments on group identity and social preferences (Chen and Li, 2009). However, in Chen and Li (2009), group identity was induced and the test was different so the results cannot be directly compared. There is a large volume of social identity literature that examines the impact of natural identities and finds positive results for intergroup discrimination (Lane, 2016; Chen et al., 2014). On the other hand, when we analyze the levels of benevolence (or malevolence) captured in the x-score and y-score of this test, we find a negative (positive) bias for women (men) when interacting with both genders (women).

The rest of the paper proceeds as follows. Section 2 overviews the experimental design, Section 3 discuss the results, and Section 4 offers concluding comments.

2 Design

Prior to the EET task subjects performed several rounds of two real-effort tasks (for details see Baier et al. 2018). Baier et al. (2018) focused on entry decisions and performance in a competitive environment and this note examines differences in distributional preferences in detail. Subjects completed two sets of the four-decision

Table 1: Equality Equivalence Test (EET).

Benevolence Behind			
Option A		Option B	
You receive	Your partner receives	You receive	Your partner receives
1.5 €	3 €	2 €	2 €
2.5 €	3 €	2 €	2 €

Benevolence Ahead			
Option A		Option B	
You receive	Your partner receives	You receive	Your partner receives
1.5 €	1 €	2 €	2 €
2.5 €	1 €	2 €	2 €

version of the EET (see Table 1). The decisions were made simultaneously meaning that on the same screen subjects made EET choices when the partner was from the same gender and from the other gender. Following the EET test, subjects completed a questionnaire and were paid for the entire experiment. The experiment was run at the EconLab at the University of Innsbruck with 360 subjects, which were equally split between men and women. The mean age in our sample is 23 years (st. dev. 3.81), and approximately two thirds of the participants are undergraduate students and one third are graduate students. 47.8 % of the subjects are enrolled in economics and business administration, 35.8 % in natural sciences, and 16.4 % in humanities. The experiment was run with the z-Tree software[4] (Fischbacher, 2007), and subjects were recruited by HROOT (Bock et al., 2014).[5]

2.1 The Equality-Equivalence Test (EET)

The Equality Equivalence Test (Kerschbamer, 2015) elicits distributional preference types at the individual level. Each subject has to make a series of binary

4 See the full set of instructions of the experiment on Appendix A.1, and screenshot of this EET task on Appendix A.2.

5 It is important to note that the data were collected at the end of the sessions from the experiment run in Baier et al. (2018). This data comes from three different treatments. We checked for possible differences between treatments and we did not find any significant differences. Therefore, henceforth we pool all our data. See Appendix A.3 for more details on the distributions of distributional preferences types by treatment and by gender pair. All the comparisons had a p-value > 0.1.

choices between allocations that both involve an own payoff for the decision maker and a payoff for a randomly matched anonymous second subject. We use the four binary choice version of the EET. In each of the four binary decision problems, one of the two allocations is symmetric – i. e., egalitarian, giving the same payoff to each person – while the other leads to unequal payoffs for the two subjects. The four choices are shown in Table 1, which breaks down into a disadvantageous inequality block (benevolence behind) and an advantageous inequality block (benevolence ahead), depending on whether inequality is to the advantage or disadvantage of the decision maker. Based on the choices in this task, each subject reveals information about his or her benevolence in the domains of disadvantageous and advantageous inequality. We use a simple coding procedure that assigns values from –1 to +1 (in steps of 1) to reveal benevolence in each domain, with positive values corresponding to higher benevolence. Benevolence in the domain of disadvantageous inequality is measured by the so-called 'x-score,', and benevolence in the domain of advantageous inequality by the so-called 'y-score'. Based on these values, we can classify each subject into one of nine distinct behavioral types. In particular, these types are: spiteful, altruist, inequality averse, inequality loving, kiss-up, envious, maxi-min, kick-down, and selfish. For the sake of simplicity in our analysis and of comparability to previous research, we will classify subjects in 4 types: Altruistic (ALT) which merges altruist and maxi-min types, Inequity Averse (IAV) corresponding to inequity averse subjects, selfish (SEL) corresponding to selfish subjects, and others (OTHERS) which merges spiteful, inequity loving, kiss-up, envious, and kick-down subjects (for further explanations on the EET test, see Kerschbamer, 2015).

3 Findings

We first use the results from Kerschbamer et al. (2019) as a benchmark to analyze whether our results are similar to previous studies using the EET in the lab. In line with that paper, we divide our total sample into the 4 aforementioned types (shown in Table 2): ALT, IAV, SEL, and OTHERS.[6] Table 2 shows that the distribution of types in our sample is close to the one in Kerschbamer et al. (2019) with the exception that in our sample there are significantly less inequity averse (IAV) subjects and more subjects falling in the category of OTHERS. We use this table

6 Note that our data comes from three different treatments (as mentioned in Section 2) and Kerschbamer et al. (2019)'s data comes also from three different treatments.

Table 2: Comparison of distributional preference type in our sample versus other sample.

	Our sample (N = 697)	Kerschbamer et al. (2019) (N = 344)	Chi-square test, p-value
Altruistic	31.56 % (N = 220)	38.08 % (N = 131)	P = 0.04*
Inequality Averse	11.62 % (N = 81)	10.75 % (N = 37)	P = 0.75
Selfish	46.91 % (N = 327)	45.93 % (N = 158)	P = 0.81
Other	9.91 % (N = 69)	5.24 % (N = 18)	P = 0.02*

Note: Chi-square test is used to compare the proportion of subjects of each type between our sample and the sample of Kerschbamer et al. (2019). 23 observations are excluded from our sample for inconsistency in the decisions.

Table 3: Comparison of distributional preference type of male vs. female.

	Male	Female	Chi-square test, p-value
Altruistic	34.57 %	28.53 %	p = 0.075
Selfish	46.86 %	46.97 %	p = 0.940
Inequality Averse	9.43 %	13.83 %	p = 0.077
Other	9.14 %	10.66 %	p = 0.527

simply to show that the order of the four types from most to least prevalent is the same across the two samples.[7]

We compare the distribution of male and female based on the four categories by pooling the in-group-out-group observations. In Table 3 we see that there is a higher proportion of alturistic men than women but women are slightly more inequity averse than men, although the differences are not significant at the 5 % level. These result go in line with previous literature (Andreoni and Vesterlund, 2001; Dickinson and Tiefenthaler, 2002; Selten and Ockenfels, 1998).

Result 1. *There are no significant differences in the proportion of types between men and women.*

Henceforth we provide power calculations and sensitivity analysis (Satterthwaite's t-test assuming unequal variances). The figures present the power dynamics given the observed differences for each test indicated in the label. The y-axis

7 While fragility could be a concern, because the other sample does not have a gender domain and ours has a larger sample size, we find the latter two explanations to be a more plausible reasons for the differences, but we cannot test for this.

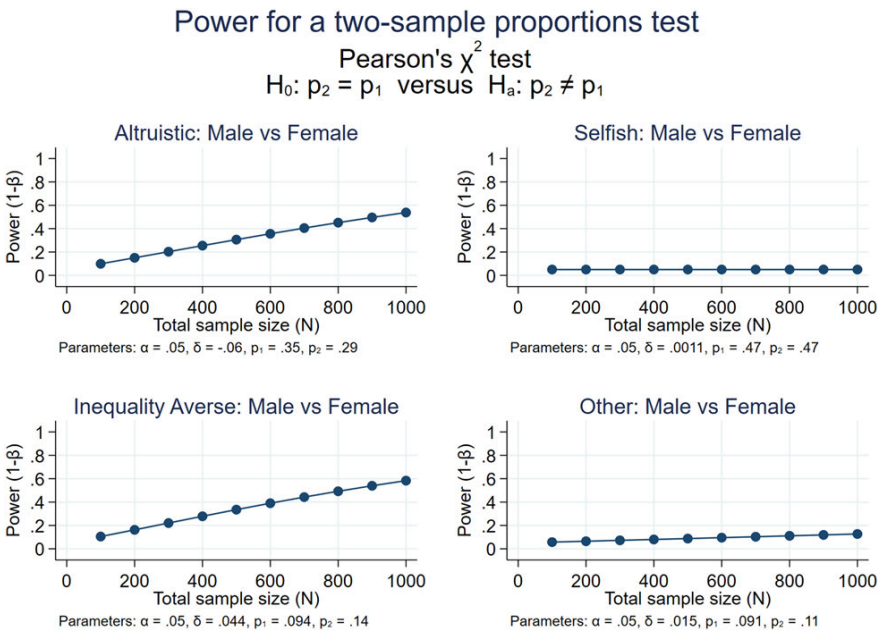


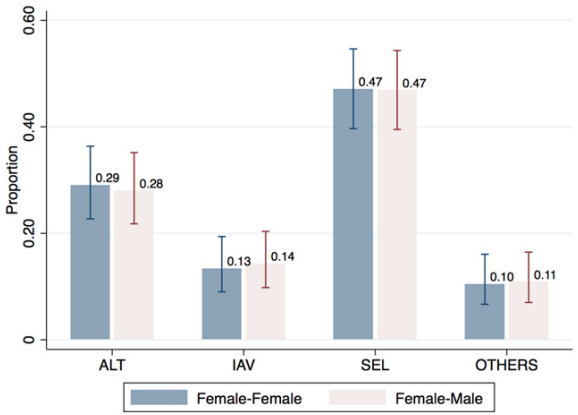
Figure 1: Power test – male vs female.

stands for the power and the x-axis for the necessary sample for each level of power to be obtained. Overall, we would need extremely large sample sizes to detect significant differences with power of 80 % or greater for the distribution of preference types by gender and by gender across both in- and out-groups. We describe our power analyses in greater detail below.

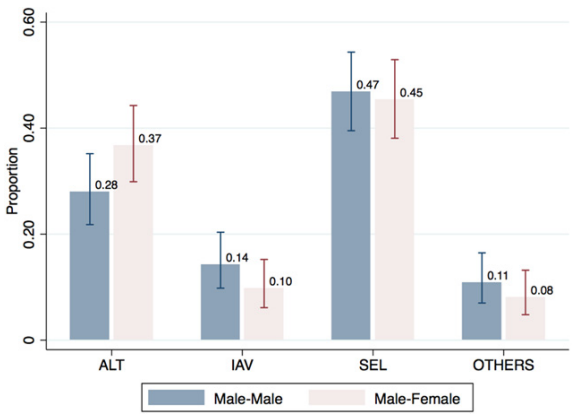
In Figure 1, for the given difference in altruistic subjects between treatments Male and Female to be significant in a 5 % level with a power of 80 % (probability of type II error) a sample of N = 1,858 observations would be required for both samples (929 per group). Similarly, for Selfish a sample of N = 6,461,966 observations would be required (3,230,983 per group). For Inequality Averse, the necessary sample would have to be N = 1,666, and for Others the sample would have to be N = 12,120.

In Figure 2 we break subject's distributional preferences into the four categories.[8] Our four within-subject comparisons are female-female when a woman is playing as decision maker (DM) with a woman as passive receiver (PS), female-

8 In Appendix 4 we also provide a table with the distribution of social preferences into the 9 categories.



(a) Female distributional types



(b) Male distributional types

Figure 2: Proportion of Distributional Preference Type by Gender and in-out-group.

male (a woman is playing as DM with a male PS), male-male (a man is playing DM with a man PS), and male-female (a man playing as DM with a female PS). Figure 2, Panel (a) shows the distribution of female types when matched with another female and when matched with a male. Panel (b) shows the distribution of male types when matched with another male and when matched with a female. Most of the subjects are selfish (around 47 %). Overall, we find no significant differences in the distribution of types between male and female regardless of whether they are matched with someone from the same or the opposite gender (none of the p-values are below 0.48, McNemar chi2). In Panel (a), we observe that around 29 %

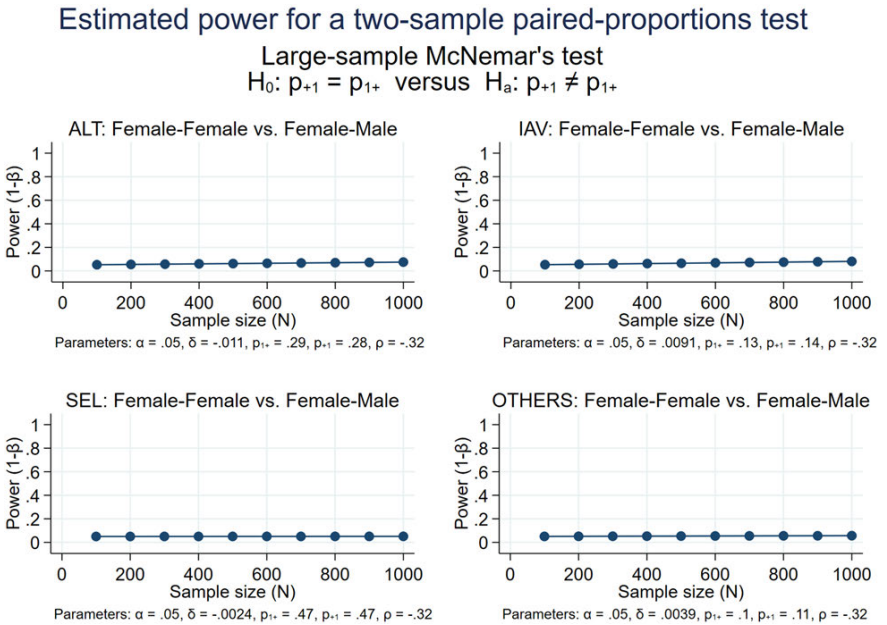


Figure 3: Power test – Female-Female vs Female-Male.

of women are classified as ALT, 14 % as IAV, 47 % as SEL, and 11 % as OTHERS. In Panel (b), around 32 % of men are classified as ALT, 12 % as IAV, 46 % as SEL, and 10 % as OTHERS.

Result 2. *There is no difference in the proportion of distributional preference types across in-group-out-group interactions for both men and women.*

We further examine if individuals switch their preference type depending if they are matched with an individual of the same or opposite gender. We find very few men (5.38 %) and women (8.48 %) switch types, and the difference is not significant ($p = 0.41$, chi–squared-test).

We now turn to the power for the tests comparing female and male behaviour towards the in-group and out-group gender, when women are decision makers. In Figure 3, we perform a power analysis for the difference between female-female and female-male for the different distributional preferences types. For the given differences to be significant in a 5 % level with power of 80 % the following samples would be necessary: For Altruistic (ALT) the number of observations would have to be N = 36,854, for Inequality Averse (IAV) the number of observations

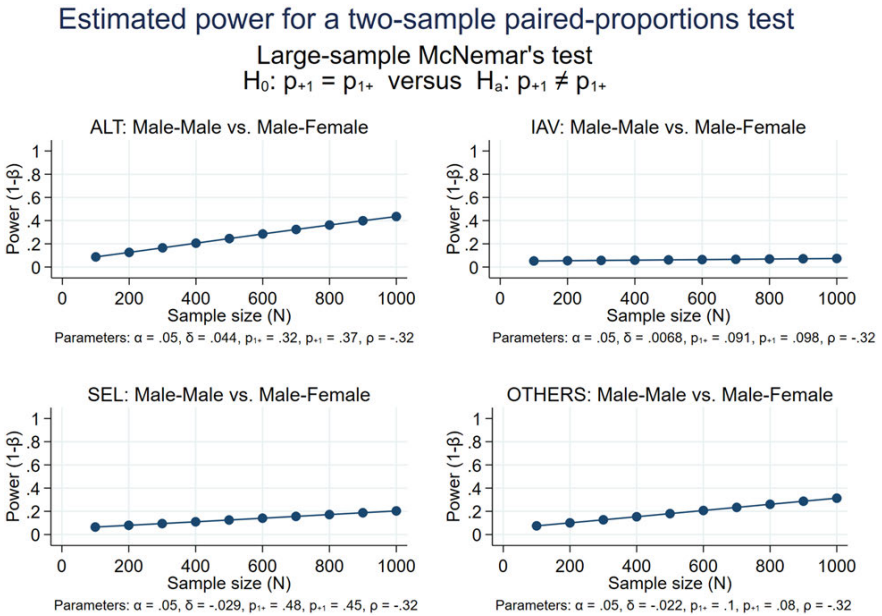


Figure 4: Power test – Male-Male vs Male-Female.

would have to be N = 29,524, and for Selfish (SEL) and OTHERS the samples would have to be N = 925,876 and N = 128,175, respectively.

In Figure 4, we perform a power analysis for the difference between male-male and male-female for the different distributional preferences types. For the given differences to be significant in a 5 % level with power of 80 % the following samples would be necessary: For Altruistic (ALT) the number of observations would have to be N = 2,429, for Inequality Averse (IAV) the number of observations would have to be N = 38,393, and for Selfish (SEL) and OTHERS the samples would have to be N = 6,170 and N = 3,614, respectively.

The power calculations show that our results would require overall extremely large samples sizes ranging from sample sizes of N = 1,666 to N = 6,461,966 to find significant results at the 5 % level with 80 % power. Given those results we believe that our results are robust.

Our regressions show more evidence of an effect of gender and in-group-out-group bias on choices in the EET than evidenced by the non-parametric tests. A possible explanation is that in Table 4 the dependent variable is not the individual type, but the level, at the intensive margin, of benevolence (or malevolence) towards another individual as measured by the x-score (specification 1) and the

Table 4: OLS regressions.

	(1) x-score	(2) y-score
Female-Female	−0.14*	−0.01
	(0.06)	(0.06)
Female-Male	−0.16*	−0.01
	(0.06)	(0.06)
Male-Female	0.02	0.06
	(0.06)	(0.05)
Constant	0.06	0.37**
	(0.04)	(0.04)
N	697	697

Note: The two specifications are Ordinary Least Square. Standard errors in parentheses, clustered by session. 23 observations are excluded from the sample for inconsistency in the decisions. *, **, indicates statistical significance at the 5 %, 1 % level, respectively.

y-score (specification 2). This method of analysis allows for pooling of all types along with documenting and exploiting changes in revealed benevolence that are not large enough to result in changes in type.

We perform two Ordinary Least Square regression specifications (Table 4) to analyze the gender differences depending on the recipient of the monetary allocation on the x-score (column 1) and on the y-score (column 2). Both *Female-Female* and *Female-Male* have a significant and negative effect on x-score (column 1). When estimating the x-score, women as decisions makers are predicted to have significantly different x-scores than when men are decision makers. But this result does depend on the gender of the inactive person. Turning to estimation of the y-score in the second column, we find that only the constant is significant, indicating that gender does is not a significant predictor for the y-score.[9]

Result 3a. *Women exhibit an in-group-out-group bias in the domain of disadvantageous inequality: they are less benevolent towards both other women and men.*

Result 3b. *Gender appears not to have a significant effect in the domain of advantageous inequality.*

Finally, we measure the size of the intergroup discrimination: We generate a variable called x-score-difference (y-score-difference) which measures, at

9 Given that in our experiment subjects participated in seven previous stages of a competition experiment, we provide a complete robustness check analysis in appendix A.5 in which we check for endogeneity issues.

Table 5: Comparison of x and y-score-difference by gender.

	Female	Male	Mann–Whitney test-statistic, p-value
Mean x-score-difference	0.023	−0.023	$z = -1.927$, $p = 0.054$
Mean y-score-difference	−0.012	−0.063	$z = -2.502$, $p = 0.012$*

*, **, indicates statistical significance at the 5 %, 1 % level, respectively.

the individual level, the difference between the x-score (y-score) when matched with the in-group and the x-score (y-score) when matched with the out-group. Therefore, we can classify subjects in 5 categories: a x-score-difference (y-score-difference) of 2 means high in-group favoritism, a score of 1 means in-group favoritism, 0 means no-favoritism, –1 means out-group favoritism, and –2 means high out-group favoritism. Table 5 shows the mean x-score-difference and the mean y-score-difference for female and male. The mean x-score-difference is negative for men while positive for women and this difference is not significant (MW, $p = 0.054$). Both genders have a negative mean y-score-difference, although the absolute value is significantly higher for men than for women (MW, $p<0.05$). Men show a higher out-group favoritism than women in terms of benevolence when ahead.

Result 4. *Overall, men are more benevolent towards women when playing in the advantageous block.*

4 Conclusion and discussion

In this paper, we examined how distributional preferences vary when individuals interact with a member of the same gender and a member of the opposite gender. We show the distribution of types is robust to one form of in and out-group effects as individuals rarely switch their preference type across the in and out group. Overall, we find no difference in the proportion of types when comparing between men and women. Further, we find no differences in the proportion of types by gender across in-group and out-group interactions. We argue these findings are important as they attest to the robustness of the EET, at least in the context of gender in a laboratory environment. On the other hand, we find a gender difference, but not an in-group-out-group bias, but in terms of benevolence as women display more malevolence than men. Finally, we find that men exhibit greater intergroup favoritism than women when measuring the size of the intergroup discrimination.

A possible explanation of this result is the social norm of chivalry (Eckel and Grossman, 2001); however, we only offer this as one potential explanation and further research is needed to confirm these findings. These results provide a new contribution to the literature on social preferences and gender, specifically on the stability of distributional preferences in the context of gender intergroup bias. Future research would need to evaluate the motivations for these results and examine if these results hold true to other types of group identifications, including in minimal group paradigm environments, or using other group delineations such as national, linguistic, or ethnic identities, as well as for other distributional types or economic environments

Acknowledgment: We are thankful to Daniel Müller and Daniel Neururer for their helpful comments. This research was approved by the Board for Ethical Questions in Science at the University of Innsbruck.

Funding: Funding from Hans Böckler Foundation Grant Project-Nr. 2017-194-3 is greatly acknowledged.

References

Akerlof, G., and R. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115(3): 715–753.

Andersen, S., G. Harrison, M. Lau, and E. Elisabet Rutström. 2008. "Lost in State Space: Are Preferences Stable?" *International Economic Review* 49(3): 1091–1112.

Andreoni, J., and L. Vesterlund. 2001. "Which is the Fair Sex? Gender Differences in Altruism." *Quarterly Journal of Economics* 116:293–312.

Baier, A., B. Davis, T. Jaber-Lopez, and M. Seidl. 2018. Gender, Competition and the Effect of Feedback and Task: An Experiment. *Working Paper Forschungsförderung, 62. Band, Düsseldorf: Hans-Böckler-Stiftung*, 1–35.

Balafoutas, L., R. Kerschbamer, and G. Oexl. 2018. "Distributional Preferences and ego Depletion." *Journal of Neuroscience, Psychology, and Economics* 11(3): 147–165.

Balafoutas, L., R. Kerschbamer, and M. Sutter. 2012. "Distributional Preferences and Competitive Behavior." *Journal of Economic Behavior & Organization* 83(1): 125–135.

Ben-Ner, A., F. Kong, and L. Putterman. 2004. "Share and Share Alike? Gender-pairing, Personality, and Cognitive Ability as Determinants of Giving." *Journal of Economic Psychology* 25(5): 581–589.

Benz, M., and S. Meier. 2008. "Do People Behave in Experiments as in the Field? – Evidence from Donations." *Experimental Economics* 11(3): 268–281.

Bock, O., I. Baetge, and A. Nichlisch. 2014. "Hroot: Hamburg Registration and Organization Online Tool." *European Economic Review* 71:114–120.

Bolton, G. E., and E. Katok. 1995. "An Experimental Test for Gender Differences in Beneficent Behavior." *Economics Letters* 48(3–4): 287–292.

Brandts, J., A. Riedl, and F. Van Winden. 2009. "Competitive Rivalry, Social Disposition, and Subjective Well-being: An Experiment." *Journal of Public Economics* 93(11–12): 1158–1167.

Cabrales, A., R. Miniaci, M. Piovesan, and G. Ponti. 2010. "Social Preferences and Strategic Uncertainty: an Experiment on Markets and Contracts." *American Economic Review* 100(5): 2261–78.

Charness, G., and M. Rabin. 2002. "Understanding Social Preferences with Simple Tests." *The Quarterly Journal of Economics* 117(3): 817–869.

Chen, Y., and S. Li. 2009. "Group Identity and Social Preferences." *American Economic Review* 99(1): 431–57.

Chen, Y., S. Li, T. Xiao, and M. Shih. 2014. "Which Hat to Wear? Impact of Natural Identities on Coordination and Cooperation." *Games and Economic Behavior* 84(C):58–86.

Croson, R., and U. Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47(2): 448–74.

de Oliveira, A., C. Eckel, and R. Croson. 2012. "The Stability of Social Preferences in a Low-income Neighborhood." *Southern Economic Journal* 79(1): 15–45.

Dickinson, D. L., and J. Tiefenthaler. 2002. "What is Fair? Experimental Evidence." *Southern Economic Journal* 414–428.

Dufwenberg, M., and A. Muren. 2006. "Generosity, Anonymity, Gender." *Journal of Economic Behavior & Organization* 61(1): 42–49.

Eckel, C., and P. Grossman. 1998. "Are Women Less Selfish than Men? Evidence from Dictator Experiments." *The Economic Journal* 108(448): 726–735.

Eckel, C., and P. Grossman. 2001. "Chivalry and Solidarity in Ultimatum Games." *Economic Inquiry* 39:171–188.

Fischbacher, U. 2007. "Z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10(2): 171–178.

Flynn, N., C. Kah, and R. Kerschbamer. 2016. "Vickrey Auction vs. BDM: Differences in Bidding Behaviour and the Impact of Other-regarding Motives." *Journal of the Economic Science Association* 2(2): 101–108.

Goette, L., D. Huffman, and S. Meier. 2012. "The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups." *American Economic Journal: Microeconomics* 4(1): 101–15.

Griesinger, D. W., and J. W. Jr Livingston. 1973. "Toward a Model of Interpersonal Motivation in Experimental Games." *Behavioral Science* 18(3): 173–188.

Hargreaves-Heap, S., and D. Zizzo. 2009. "The Value of Groups." *American Economic Review* 99(1): 295–323.

Iriberri, N., and P. Rey-Biel. 2013. "Elicited Beliefs and Social Information in Modified Dictator Games: What do Dictators Believe Other Dictators Do?" *Quantitative Economics* 4(3): 515–547.

Kerschbamer, R. 2015. "The Geometry of Distributional Preference and a Non-parametric Identification Approach." *European Economic Review* 76:85–103.

Kerschbamer, R., and D. Muller. 2020. "Social Preferences and Political Attitudes: an Online Experiment on a Large Heterogeneous Sample." *Journal of Public Economics* 182, 104076.

Kerschbamer, R., D. Neururer, and A. Gruber. 2019. "Do the Altruists Lie Less?" *Journal of Economic Behavior and Organization* 157:560–579.

Liebrand, W. B. 1984. "The Effect of Social Motives, Communication and Group Size on Behaviour in an N-person Multi-stage Mixed-motive Game." *European Journal of Social Psychology* 14(3): 239–264.

Lane, T. 2016. “Discrimination in the Laboratory: A Meta-analysis of Economics Experiments.” *European Economic Review* 90:375–402.

Meier, S., and C. Sprenger. 2015. “Temporal Stability of Time Preferences.” *Review of Economics and Statistics* 97(20:273–286.

Müller, D. 2017. “The Anatomy of Distributional Preferences with Group Identity.” *Journal of Economic Behavior & Organization* 166:785–807.

Offerman, T., J. Sonnemans, and A. Schram. 1996. “Value Orientations, Expectations and Voluntary Contributions in Public Goods.” *The Economic Journal* 106(437): 817–845.

Selten, R., and A. Ockenfels. 1998. “An Experimental Solidarity Game.” *Journal of Economic Behavior & Organization* 34(4): 517–539.

Sutter, M. 2009. “Individual Behavior and Group Membership: Comment.” *American Economic Review* 99(5): 2247–2257.

Sutter, M., S. Haigner, and M. G. Kocher. 2010. “Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations.” *The Review of Economic Studies* 77(4): 1540–1566.

Tajfel, H., and J. C. Turner. 1979. “An Integrative Theory of Intergroup Conflict.” In *The Social Psychology of Intergroup Relations*, edited by W. G. Austin and S. Worchel, 33–47. Monterey, CA: Brooks/Cole.

Volk, S., C. Thöni, and W. Ruigrok. 2012. “Temporal Stability and Psychological Foundations of Cooperation Preferences.” *Journal of Economic Behavior & Organization* 81(2): 664–676.

Supplemental Material: The online version of this article offers supplementary material (https://doi.org/10.1515/ger-2019-0119).