Exp Econ (2010) 13: 299–308
DOI 10.1007/s10683-010-9242-8

Gender and generosity: does degree of anonymity or group gender composition matter?

C. Bram Cadsby · Maroš Servátka · Fei Song

Received: 9 November 2009 / Accepted: 5 May 2010 / Published online: 18 May 2010

Abstract Employing a two-by-two factorial design that manipulates whether dictator groups are single or mixed-sex and whether procedures are single or double-blind, we examine gender effects in a standard dictator game. No gender effects were found in any of the experimental treatments for the mean or median levels of giving, or for the propensity to give nothing. However, females chose to give away half of their endowments with greater frequency than males in the pooled single-sex treatments.

Keywords Other-regarding · Selfish · Altruism · Gender · Dictator · Anonymity · Experiment · Priming

JEL Classification C91 · D64

1 Introduction

Numerous experimental and theoretical studies have examined and analyzed other-regarding versus self-interested behavior (e.g., Bolton and Ockenfels 2000; Camerer 2003; Fehr and Schmidt 1999; Rabin 1993). To date, extensive evidence suggests that people are not indifferent to the welfare of others, even to anonymous

Electronic supplementary material The online version of this article (doi:10.1007/s10683-010-9242-8) contains supplementary material, which is available to authorized users.

C.B. Cadsby (✉)
Department of Economics, University of Guelph, Guelph, ON N1G 2W1, Canada
e-mail: bcadsby@uoguelph.ca

M. Servátka
Department of Economics and Finance, University of Canterbury, Christchurch, New Zealand

F. Song
Ted Rogers School of Business Management, Ryerson University, Toronto, ON, Canada

strangers, particularly when their choices directly determine other people's outcomes (e.g., Charness 2000; Charness and Rabin 2002; for an excellent recent survey on other-regarding behavior see Cooper and Kagel 2009). A well-known and widely employed experimental framework that highlights the role of distributional other-regarding preferences is the dictator game (e.g., Forsythe et al. 1994; Kahneman et al. 1986). In this game, one of the two parties is randomly assigned to be the "dictator" and endowed with a sum of money, $\$X$. The dictator is allowed to give the other party, the "recipient," any amount of money, $\$Y$, where $0 \leq Y \leq X$, at the dictator's own choosing. The recipient makes no decision, but simply receives whatever is given to him/her. Since the recipient has no power to influence the dictator's decision, this game effectively removes any strategic motivation to behave in an apparently altruistic manner.

Within the large dictator-game literature, gender has been studied to determine whether one sex is more generous than the other. The most dramatic and impactful finding comes from Eckel and Grossman (1998) (EG hereafter). In a carefully designed study, they find that under double-blind procedures, women give away twice as much as men (16% versus 8% of the endowment). This result, which supports the conventional wisdom that women are more generous than men, has received much attention from both economists and other social scientists and has been cited widely since its publication.[1] In contrast, Bolton and Katok (1995) (BK hereafter) find no gender effects (12% versus 11% of the endowment for women and men respectively).[2] Other studies also provide mixed and ambiguous results (e.g. Andreoni and Vesterlund 2001; Cox 2002; Cox and Deck 2006; Ben-Ner et al. 2004; Dufwenberg and Muren 2006; Song et al. 2004), though the experimental settings in these papers are somewhat different from the standard dictator-game setting used in EG and BK. Two recent survey articles summarize the large and growing literature on gender differences in economic behavior (Croson and Gneezy 2009; Eckel and Grossman 2008).

A closer look at EG and BK reveals that there are notable methodological differences between these two papers. Specifically, while EG used single-sex groups in its dictator sessions and adopted a double-blind procedure, BK used mixed-sex dictator groups and a single-blind procedure. Thus, it is unclear whether the source of the different gender results was the single- versus mixed-gender dictator groups, the single- versus double-blind procedures, or some other factor. In this study, we intended to identify the source of this discrepancy by means of a two-by-two factorial design that manipulates whether dictator groups are single or mixed-sex (henceforth *ss* or *ms*) and whether procedures are single- or double-blind (henceforth *sb* or *db*).[3] In light of the results, we ultimately chose to run three of the four treatments that make up the complete factorial design: *ss:sb*, *ms:sb*, and *ss:db*.

[1] According to the Web of Science, Eckel and Grossman (1998) has been cited 96 times in published articles. Google Scholar lists 289 citations in published and unpublished work.

[2] According to the Web of Science, Bolton and Katok (1995) has been cited 41 times in published articles. Google Scholar lists 107 citations in published and unpublished work.

[3] For a discussion on the importance of anonymity and different experimental procedures, see Hoffman et al. (1996, 1994).

In economics, a double-blind treatment is one in which the decisions and payoffs of the subjects are anonymous with respect both to the other subjects and to the experimenters (Hoffman et al. 1994). In contrast, in a single-blind treatment, decisions and payoffs are anonymous with respect to other subjects, but not necessarily to the experimenters. EG stressed the importance of their double-blind setting, arguing that it "removes risk, the possibility of gender-related subject interactions, and the experimenter effect" (p. 732).

We conjectured that sitting in a room in which all other decision-makers were of the same gender might also have important effects through gender priming. There is a large psychological literature on gender priming. For example, Shih et al. (1999) demonstrate that Asian-American women perform better on a mathematics test when their ethnic identity is activated and worse when their gender identity is activated than a control group. They argue that the activation of cultural and gender stereotypes is responsible for these results. In Shih et al. (1999), the priming occurs through questionnaires that ask either gender- or ethnic-related questions. We hypothesized that sitting in a room with others, who are all of the same gender, might have a similar effect, especially as this is not likely to occur through random chance.

To our surprise, no gender effect was found in any of the experimental treatments for mean or median donation, or for the propensity to donate a positive versus a zero amount. Moreover, neither single- versus mixed-sex groups nor level of anonymity had any impact on male relative to female behavior along any of these dimensions. The only manifestation of a gender effect in our data was that in single-sex groups females had a higher propensity than males to donate half of their endowments, the maximum permitted in our design, to an anonymous recipient.

2 Experimental design and procedures

Twenty four experimental sessions took place at the University of Canterbury in New Zealand in 2009 with 764 undergraduate students serving as subjects. The recruited dictators, but not the recipients, had never previously participated in an economics experiment. Dictator sessions and recipient sessions were run separately. The recipients were recruited for sessions which took place after the dictator sessions in a different classroom. This was done to ensure that the dictators and recipients did not run into each other on the way to the experimental sessions. Each experimental session lasted about 30 minutes including the initial instruction period and the payment of subjects. The subjects earned on average 15 New Zealand Dollars (NZD), part of which was a 5 NZD show-up fee.[4]

At the beginning of a session, upon entering the classroom, subjects were free to choose any seat. Each subject was provided with a copy of the instructions, which were identical for all subjects in the session. The instructions were then read aloud by the male experimenter who was the only non-subject person present in the classroom. Subjects were informed that the size of the pie was 20 NZD and dictators were instructed that they could choose one of six options from giving 0 to 10 NZD to the

[4]The adult minimum wage in New Zealand at the time of the experiment was 12 NZD per hour.

recipient in 2 NZD increments.[5] The instructions were also projected on a screen. Any questions were asked and answered privately.

In the two *sb* treatments, the experimenter distributed decision sheets to dictators along with large manila envelopes. Once the dictator made a decision by circling one of the six available options, s/he was asked to place the decision sheet in the envelope provided to keep his/her decision private. Then the dictators were asked one by one to approach the payment desk outside the classroom where they were privately paid their experiment earnings plus the show-up fee. In the *ms:sb* treatment, the experimenter noted the dictator's gender on a separate sheet of paper which was not visible to the subject being paid. In the recipient sessions, the large manila envelopes were placed in a box. The recipients were asked to approach the payment desk one-by-one, and to draw a manila envelope from the box. They were then paid according to the decision indicated on the decision sheet inside the envelope.

In the *ss:db* treatment, the dictators were asked to go one by one to a private decision room to make their decision in complete privacy. On the way to the decision room, each dictator was given a small envelope containing ten two-dollar coins and a large manila envelope.[6] The subjects were instructed to take out between five and ten two-dollar coins (thus having six options) and place them in their wallets, pockets, purses etc., leaving the rest for an anonymously paired recipient. They were then asked to put the small envelope in a larger manila envelope and place it in the box outside the decision room. This procedure ensured that neither the experimenter nor anybody else learned what decisions were made by individual subjects. All subjects received a show-up fee after they exited the decision room. As the *db* sessions were either all-male or all-female, there was no need to note the gender of the dictators. In the recipient sessions, the large manila envelopes containing small envelopes with the two-dollar coins (if any) were placed in a box as in the *sb* sessions. The recipients were asked to approach the payment desk one-by-one and to draw a manila envelope from the box. It was stressed by the experimenter that recipients could not touch more than one envelope when making their selection. This was to avoid having subjects choose envelopes based on their weight. Once the recipients selected an envelope, they kept the two-dollar coins (if any) that were in the small envelope and received a show-up fee.

3 Results

Table 1 summarizes the experimental results. Panel A shows the average amount of money given away by dictators and the sample size in each of the six categories. Panel B provides the distribution of dictators' decisions in each treatment by men and women.

As summarized in Table 2, whether using Mann-Whitney ranking tests, Kolmogorov-Smirnov tests for equal distributions, Mood's Chi-Squared tests for

[5] In EG, subjects were endowed with $10 US and chose from 11 options in $1 increments. In BK, subjects were also endowed with $10 US, but chose from six options comparable to those in our design.

[6] We could not implement the ingenious design used by EG, following Hoffman et al. (1994), because neither $1 nor $2 bills are used in New Zealand.

Table 1 Experiment summary

	Men	Women
Panel A: Average amount of money given away by dictators		
Double blind, single sex	3.2982 ($n = 57$)	3.5600 ($n = 59$)
Single blind, mixed sex	2.9355 ($n = 62$)	3.4545 ($n = 55$)
Single blind, single sex	3.0344 ($n = 58$)	3.6949 ($n = 59$)
All treatments	3.08 ($n = 177$)	3.57 ($n = 173$)

Level	Double-blind		Single-blind			
	Single-sex		Single-sex		Mixed-sex	
	Men	Women	Men	Women	Men	Women
Panel B: Distribution of amounts given away by dictators (percentage in parentheses)						
\$0	21 (37%)	24 (41%)	32 (55%)	29 (49%)	36 (58%)	26 (47%)
\$2	12 (21%)	9 (15%)	4 (7%)	5 (8%)	3 (5%)	6 (11%)
\$4	9 (16%)	7 (12%)	2 (3%)	5 (8%)	5 (8%)	6 (11%)
\$6	3 (5%)	6 (10%)	8 (14%)	3 (5%)	6 (10%)	1 (2%)
\$8	5 (9%)	1 (2%)	4 (7%)	0 (0%)	0 (0%)	6 (11%)
\$10	7 (12%)	12 (20%)	8 (14%)	17 (29%)	12 (19%)	10 (18%)
Total	57 (100%)	59 (100%)	58 (100%)	59 (100%)	62 (100%)	55 (100%)

equal medians or conventional t-tests for equal means, no gender differences were found in any of the three treatments. Moreover, using a binary measure of whether a dictator gave away a positive amount or not as the dependent variable and gender as the independent variable, logit regressions did not yield any significant gender effects in any of these three treatments either. However, a logit regression using a binary measure of whether or not a dictator gave away half of his/her endowment did suggest a gender difference for the *sb:ss* treatment ($p = 0.052$) on this dimension.[7] Pooling the data first from the two single-blind treatments and then from the two single-sex treatments, and running the same tests also showed no significant differences between male and female donations, again with one exception. The probability of female dictators donating half of their endowments was significantly higher than the probability of male dictators doing so for the pooled *ss* treatments ($p = 0.029$).

Next, we examined whether gender-priming influenced the other-regarding behavior of men relative to women. To do so, we first ran an OLS regression with donation amount as the dependent variable and gender (female = 0, male = 1), a treatment dummy for gender priming ($ss = 0$, $ms = 1$) and an interaction between the two as independent variables. As indicated in Table 3, Panel A, it yielded no significant effects. We then ran a logit regression with whether or not a dictator donated a positive amount as the dependent variable, using the same independent variables as the OLS

[7] We thank an anonymous referee for suggesting that we examine the effect of gender on the propensity to give away half of one's endowment.

Table 2 Gender effects on giving (p-values in parentheses)

Treatment	Nonparametric tests			Parametric t-test of difference in means*	Logit regression on positive amount given away*	Logit regression on half endowment given away*
	Mann-Whitney	KS	X^2 test of difference in medians			
db:ss	0.774 (0.439)	0.081 (0.992)	0.045 (0.831)	−0.378 (0.706)	−0.162 (0.672)	−0.601 (0.548)
sb:ms	0.900 (0.368)	0.108 (0.886)	1.363 (0.243)	−0.697 (0.487)	0.435 (0.244)	−0.077 (0.871)
sb:ss	0.152 (0.879)	0.150 (0.524)	0.424 (0.515)	−0.866 (0.388)	0.242 (0.515)	0.928 (0.052)
sb (*ms* and *ss* pooled)	1.248 (0.212)	0.081 (0.992)	0.065 (0.772)	−1.120 (0.264)	0.198 (0.713)	−0.439 (0.182)
ss (*db* and *sb* pooled)	0.900 (0.368)	0.812 (0.524)	0.425 (0.515)	−0.865 (0.389)	0.242 (0.515)	−0.776 (0.029)
All treatments	1.107 (0.268)	0.073 (0.741)	0.521 (0.471)	−1.156 (0.248)	0.185 (0.387)	N/A (see text)

*Dummy variables are coded as follows: Male = 1, Double-blind = 1, and Mixed-sex = 1

Table 3 Gender and treatment effects on giving (p-values in parentheses)

	OLS regression on amount given away	Logit regression on positive amount given away	Logit regression on half endowment given away
Panel A: The effects of gender and gender-priming treatment*			
Male dummy	−0.546 (0.271)	−0.053 (0.835)	−0.837 (0.015)
Mixed-sex dummy	−0.304 (0.614)	−0.103 (0.735)	−0.467 (0.243)
Interaction	0.217 (0.800)	−0.412 (0.346)	1.104 (0.055)
Constant	3.699 (0.000)	0.207 (0.254)	−1.099 (0.000)
Panel B: The effects of gender and anonymity treatment*			
Male dummy	−0.596 (0.250)	−0.338 (0.198)	−0.439 (0.182)
Double-blind dummy	−0.020 (0.975)	0.307 (0.344)	−0.195 (0.618)
Interaction	0.335 (0.710)	0.500 (0.280)	−0.162 (0.792)
Constant	3.579 (0.000)	0.070 (0.708)	1.170 (0.000)

*Dummy variables are coded as follows: Male = 1, Double-blind = 1, and Mixed-sex = 1

regression discussed above. Again, nothing was significant. Finally, we ran another logit regression with the same independent variables, but using whether or not the dictator gave away half of his/her endowment as the dependent variable. Consistent with our previous results, the main effect of gender was significant ($p = 0.015$) for the pooled *ss* treatments, while the interaction effect between gender and the gender-priming treatment variable was very close to the conventional 0.05 significance level ($p = 0.055$). The latter result implies that there is a near significant difference in gender effects on the propensity to give away half of one's endowment between the pooled *ss* and *ms* treatments.

Analogous OLS and logit regressions with level of anonymity as the treatment variable showed no significant effects. The failure to reject the null hypothesis of no interaction effect in these three regressions implies that anonymity level had no significant effect on the difference between male and female contributions.

Given that no treatment effects were found except as regards the relative propensity of males versus females to donate half of their endowments, we pooled the data from all three treatments together to allow for greater statistical power in analyzing the other possible manifestations of a gender effect. The average amount given away was $3.08 (15%) for men and $3.57 (18%) for women. Once again, Mann-Whitney ranking tests, Kolmogorov-Smirnov tests for equal distributions, Mood's Chi-Squared tests for equal medians and conventional t-tests for equal means, reported in the last line of Table 2, all failed to show any significant gender differences. Moreover, a binary logistic regression of the probability of giving away any positive amount showed that gender did not have any significant impact either. We did not pool the data across treatments to examine the propensity to donate half of one's endowment because of the existence of the treatment effects discussed above.

4 Discussion

In this study, we were unable to explain the reason that EG found females to be more generous than males, while BK did not. Instead, our results seem largely to support the BK finding, even in the *ss:db* treatment that most resembles the EG study. However, we did obtain one result along similar lines to those obtained by EG. Females had a greater propensity to donate half of their endowments in the pooled *ss*, but not in the *ms* treatment. The EG study also used single-sex groups. In their study, 15% of females compared with 5% of males donated half or more of their endowments.[8] In the BK study, which used mixed-sex groups, the comparable percentages, eyeballed from their Fig. 2, were about 10% for males and 9% for females. The comparable percentages in our three treatments are presented in Table 1, and along with our statistical results, suggest that gender priming through group composition may be responsible for the differing male versus female propensities to donate half of their endowments in EG versus BK. Working in single-sex groups appears to prime some females to make the "ultra-fair" choice of equally dividing their endowments while priming some males to do the opposite. However, any effect of such priming is too weak to result in any other observable treatment effects on the relative overall generosity of males versus females. Given the inability of gender priming to explain the major discrepancies in male versus female giving between EG and BK, we do not feel that strong conclusions about its role in affecting the relative propensities of males and females to donate resources to others are warranted. Rather, we urge that the role of gender priming on other-regarding and other economic behavior be studied further by experimentalists.

Why do the results of our study differ from those of EG? Of course, there were differences between the two studies. We use New Zealand university undergraduates rather than American undergraduates. Owing to the non-existence of low-denomination paper currency in New Zealand, we implement the double-blind treatment in a different manner. Following BK, we do not permit our dictators to give away more than half of their endowments. However, attitudes toward gender in New Zealand appear to resemble those in the United States, our *db* treatment provides complete anonymity, and in EG, only one subject, a male, gave away more than half of his endowment.

As in any replication by a different set of researchers, the experimenters were unavoidably different. As in BK, our experiments were all conducted by a male researcher.[9] EG's experiments were conducted by a female. Different experimenters, despite their best efforts, can potentially affect the behavior of subjects. In Kantowitz et al. (2008), a well-known textbook in experimental psychology, the authors address this issue as follows: "The gender, race, and ethnicity of the experimenter are also potential experimenter effects. Experimenter characteristics are more likely to bias the results of an experiment in research that focuses on issues related to these

[8] These proportions include the one male subject who donated more than half of his endowment. Such an action was not permitted either in BK or our own study.

[9] BK refers to the experimenter as the "monitor." In a private communication, Gary Bolton indicated that he was the monitor.

characteristics—for example, the race of an experimenter who is conducting an experiment concerning the effect of skin color on work performance ratings" (p. 70). It is possible that such experimenter effects influenced the results in all three papers.

Cooper and Kagel (2009) note that "results from dictator games are sensitive to a variety of seemingly innocuous variations" (p. 29). However, if the finding that women are more generous than men depends on such design features, it is difficult to maintain that it represents "a baseline difference in men and women" as EG concluded (p. 733). While this conclusion seemed appropriate and warranted given the results of their well-designed study, good science requires replication and robustness checks. Inability to replicate demands reconsideration of earlier conclusions. This is particularly important when the conclusion feeds into a common popular stereotype and has made such a deep impression, evidenced by the number of citations EG has received.

After completing this paper, we became aware of a related study by Boschini et al. (2009). Like us, they explore the role played by different experimental settings of the dictator game in eliciting apparent gender differences in generosity. They also examine gender priming, doing so in two ways: first, they ask subjects to specify their gender either on the first or last page of a questionnaire; second, some sessions are single-sex, while others are mixed-sex. Only an administrator, who was not otherwise involved in the experiment, observed the earnings of the subjects. They found women to be more generous than men only in their gender-first, mixed-sex treatment. Since EG used single-sex groups and did not prime their subjects by means of a questionnaire, these results, while interesting and provocative, cannot explain the differences between EG and BK.

The lesson in all this is that the jury is still out on gender differences in generosity. More evidence and replication is required before we can confidently distinguish between a corroborated conclusion and either type-1 or type-2 statistical error.

Acknowledgements Anne Boschini, Gary Bolton, Philip Grossman, and two anonymous reviewers provided helpful comments and suggestions. Financial support was provided by the Social Sciences and Humanities Research Council of Canada, grant #2006-410-138 (Fei Song). The Erskine Programme also facilitated this research through a Visiting Erskine Fellowship awarded to C. Bram Cadsby to visit the University of Canterbury.

References

Andreoni, J., & Vesterlund, L. (2001). Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics*, *116*(1), 293–312.

Ben-Ner, A., Kong, F., & Putterman, L. (2004). Share and share alike: gender-pairing, personality, and cognitive ability as determinants of giving. *Journal of Economic Psychology*, *25*(5), 581–589.

Bolton, G., & Katok, E. (1995). An experimental test for gender differences in beneficent behavior. *Economics Letters*, *48*(3–4), 287–292.

Bolton, G., & Ockenfels, A. (2000). ERC: a theory of equity, reciprocity, and competition. *American Economic Review*, *90*(1), 166–193.

Boschini, A., Muren, A., & Persson, M. (2009). Constructing gender in the economics lab. www.ne.su.se/paper/wp09_15.pdf. Accessed 19 April 2010.

Camerer, C. F. (2003). *Behavioral game theory: experiments on strategic interaction*. Princeton: Princeton University Press.

Charness, G. (2000). Responsibility and effort in an experimental labor market. *Journal of Economic Behavior and Organization*, *42*(3), 375–384.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, *117*(3), 817–869.

Cooper, D. J., & Kagel, J. H. (2009). Other regarding preferences: a selective survey of experimental results. http://www.econ.ohio-state.edu/kagel/Other_RegardingPreferences_survey.pdf. Accessed 19 April 2010. Forthcoming in J. H. Kagel, & A. Roth (Eds.), *The handbook of experimental economics* (Vol. *2*). Princeton University Press.

Cox, J. C. (2002). Trust, reciprocity, and other-regarding preferences: groups vs. individuals and males vs. females. In R. Zwick, & A. Rapoport (Eds.), *Experimental business research*. Norwell: Kluwer Academic.

Cox, J. C., & Deck, C. (2006). When are women more generous than men? *Economic Inquiry*, *44*(4), 587–598.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 1–27.

Dufwenberg, M., & Muren, A. (2006). Generosity, anonymity, gender. *Journal of Economic Behavior and Organization*, *61*(1), 42–49.

Eckel, C., & Grossman, P. (1998). Are women less selfish than men? Evidence from dictator experiments. *Economic Journal*, *108*(448), 726–735.

Eckel, C., & Grossman, P. (2008). Differences in the economic decisions of men and women: experimental evidence. In C. Plott, & V. Smith (Eds.), *Handbook of experimental economics results*. New York: Elsevier.

Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868.

Forsythe, R., Horowitz, J., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, *6*(3), 347–369.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preference, property rights and anonymity in bargaining games. *Games and Economics Behavior*, *7*(3), 346–380.

Hoffman, E., McCabe, K., & Smith, V. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review*, *86*(3), 653–660.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of Business*, *59*(4), 285–300.

Kantowitz, B. H., Roediger, H. L., & Elmes, D. G. (2008). *Experimental psychology* (9th edn.). Belmont: Wadsworth Cengage Learning.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, *83*(5), 1281–1302.

Shih, M., Pittinsky, T., & Ambady, N. (1999). Stereotype susceptibility: identity salience and shifts in quantitative performance. *Psychological Science*, *10*(1), 80–83.

Song, F., Cadsby, C. B., & Morris, T. (2004). Other-regarding behavior and behavioral forecasts: females versus males as individuals and group representatives. *International Journal of Conflict Management*, *15*(4), 340–363.