Heinz, Matthias; Juranek, Steffen; Rau, Holger A.

Working Paper
Do women behave more reciprocally than men? Gender differences in real effort dictator games

DICE Discussion Paper, No. 24

Provided in Cooperation with:
Düsseldorf Institute for Competition Economics (DICE)

Suggested Citation: Heinz, Matthias; Juranek, Steffen; Rau, Holger A. (2011) : Do women behave more reciprocally than men? Gender differences in real effort dictator games, DICE Discussion Paper, No. 24, Düsseldorf Institute for Competition Economics (DICE), Düsseldorf

This Version is available at:
http://hdl.handle.net/10419/48600

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Do Women Behave More Reciprocally than Men? Gender Differences in Real Effort Dictator Games

Matthias Heinz, Steffen Juranek, Holger A. Rau

June 2011
The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors’ own opinions and do not necessarily reflect those of the editor.
Do Women behave more Reciprocally than Men?
Gender Differences in Real Effort Dictator Games*

Matthias Heinz†, Steffen Juranek†
Department of Management and Microeconomics
Goethe University Frankfurt, Germany

Holger A. Rau§
Düsseldorf Institute for Competition Economics (DICE)
Heinrich-Heine University Düsseldorf, Germany

Abstract
We analyze dictator allocation decisions in an experiment where the recipients have to earn the pot to be divided with a real-effort task. As the recipients move before the dictators, their effort decisions resemble the first move in a trust game. Depending on the recipients’ performance, the size of the pot is either high or low. We compare this real-effort treatment to a baseline treatment where the pot is a windfall gain and where a lottery determines the pot size. In the baseline treatment, reciprocity cannot play a role. We find that female dictators show reciprocity and decrease their taking-rates significantly in the real-effort treatment. This treatment effect is larger when female dictators make a decision on recipients who successfully generated a large pot compared to the case where the recipients performed poorly. By contrast, there is no treatment effect with male dictators, who generally exhibit more selfish behavior.

JEL Classification numbers: C72, C91.

Keywords: Gender, Reciprocity, Dictator Game, Real Effort.

*We are specially grateful to Hans-Theo Normann for his input at various stages of this project. We are also grateful to Robert Dur, Catherine Eckel, Tore Ellingsen, Armin Falk, Guido Friebel, Daniel Hausser, Ben Jacobsen, Michael Kosfeld, Kevin McCabe, Jörg Oechsler, Christiane Schwieren, Uwe Walz, and conference and seminar audiences at the Economic Science Association World Meeting Copenhagen 2010 and the Goethe University Workshop in Riedern 2010, for helpful comments. We also want to thank Volker Benndorf for supporting us with the survey analysis. We are grateful to the Goethe University of Frankfurt, the Düsseldorf Institute for Competition Economics (DICE), the Sparkasse Bank Hessen e.G., and the Business and Economics Alumni Association of Goethe University for their financial support.

†Grüneburgplatz 1, 60325 Frankfurt, Phone: +49 69 798 34825, EMail: heinz@econ.uni-frankfurt.de
‡Grüneburgplatz 1, 60325 Frankfurt, Phone: +49 69 798 34806, EMail: juranek@econ.uni-frankfurt.de
§corresponding author: Holger Rau, Universitätsstr. 1, 40225 Düsseldorf, Phone: +49 211 81 15615, Email: rau@dice.uni-duesseldorf.de
1 Introduction

Our study analyzes gender differences in reciprocal behavior. We apply a modified dictator game based on Cherry et al.’s (2002) study where dictators had to do a real-effort task before deciding on the money amount to be dictated to recipients. In contrast to most standard dictator games (e.g. Kahneman et al. (1986), Eckel and Grossman (1996) and Dana et al. (2006)), Cherry et al. (2002) report that 95% of their dictators behaved in line with the standard neoclassical prediction of selfish maximization of their own monetary income.1

Our study modifies the Cherry et al. (2002) setup in that we make the recipients (rather than the dictators) conduct the real-effort task. Since recipients move before dictators, their effort decisions resemble the first move in a trust game (Berg et al. (1995)). That is: if the recipients do not trust dictators, they should not invest effort. Note here a crucial difference to the trust game is that our game does not include an outside option, i.e. all first-movers have to send their money to the second-movers.2 According to our setup’s resemblance to a trust game, we expect gender differences in terms of dictators’ trustworthiness in our setup. We therefore hypothesize that female dictators will show increased reciprocity to recipients who have worked. We expect them to reciprocate a good working performance by strongly reducing their taking-rates if the recipients performed well. This is motivated by the trust game literature on gender differences, which reports that female second-movers more often reciprocate first-movers’ offers by sending back positive amounts (e.g. Croson and Buchan (1999), Chaudhuri and Gangadharan (2003), Snijders and Keren (2004), and Buchan et al. (2008)).3

Our setup is most closely related to Rue’s (1998) study on tipping behavior in the ultimatum game and in the dictator game. Here, the recipients also endogenized the money to be divided by doing a real-effort task, before dictators decided about the split. The main difference in our study is that we focus on gender differences in reciprocal behavior.4 Related is also Oxoby and Spraggon (2008), who demonstrate that dictators make significantly lower offers when recipients did a real-effort task. However, there are two crucial differences: The authors used a standard dictator game as a control treatment. In contrast, we use a baseline treatment, where dictators can take money from recipients in a windfall environment and, furthermore, we focus on gender differences. Our results not only successfully replicate Rue’s (1998) and Oxoby and Spraggon’s (2008) findings; we also establish that female dictators are affected by reciprocity and show different magnitudes of reciprocal behavior depending on recipients’ working performance.

---

1 Further, some dictator games report that dictators decrease their taking-rates due to increased anonymity (e.g. Hoffman et al. (1996), Koch and Normann (2008)).

2 That setup ensures that dictators have to decide about both: successful and low performing recipients.

3 There also exist papers which report gender effects in standard dictator games (e.g. Eckel and Grossman (1998), Bolton and Katok (1995), Andreoni and Vesterlund (2001) and Dufwenberg and Muren (2006)). For a complete survey see Croson and Gneezy (2009).

4 Carlsson et al. (2010) also study gender effects in a dictator game with a real-effort task. However in this setup the dictators do the real-effort task. The authors do not find a significant gender difference.
2 Experimental Design

We use a modified dictator game with two differences to the standard game. The first difference is that the size of the pot which the dictators decide about is not constant. The second difference is that there are two stages in both the Windfall and the Real Effort treatment. As in Cherry et al. (2002), there is a first stage (money-generation stage) in which the size of the pot is determined. Dictators only decide in the second stage (allocation stage) about the allocation of the money.

In the Real Effort treatment the participants were randomly assigned to two groups of equal size and split between separate rooms, rooms A and B. In the money-generation stage, subjects in room A (the recipients) had the opportunity to take a quiz which consisted of 20 questions taken from the Graduate Record Examination test. Depending on their results, we allocated money to the subjects such that subjects who answered at least 13 questions correctly were given 10€, otherwise they received 5€. Subjects knew that they had 20 minutes to complete the quiz. As we corrected the tests, dictators in room B had to wait for approximately 30 minutes and we provided them with coffee and cake. The allocation stage of the Real Effort treatment randomly matched subjects in room A with those in room B. Neither subjects in room A nor subjects in room B knew the identities of their partners. Furthermore subjects in room A were informed about the amount of money we allocated, which depended on their results in the quiz. Individuals in room B were not told about the exact result of their interaction partner. They only learned whether the recipient generated 5€ or 10€. Every subject in room B dictated in a one-shot dictator game a split of the wealth to the recipients in room A. Subjects in room A were informed by the experimenter about the allocation decision and the final earnings which they received according to the dictator’s decision. Afterwards both the dictators and the recipients had to complete a short survey. Finally subjects A and B were paid out at the end of the experiment.

Altogether we conducted five sessions of the Real Effort treatment. Four of these sessions had the structure explained above. In the remaining session we checked whether dictators are sensitive to overconfidence. We therefore provided them with a copy of the exam questions taken by those in group A (which may have induced positive or negative reciprocity, depending on the dictator’s self-assessment of his or her own ability). Dictators were given the chance to have a look at the exam questions for 10 minutes, before they were asked to estimate the number of questions they would have solved correctly. After making their allocation decision we asked them to do the test, to check whether they had overestimated their own performance. However, we did not find significant differences in dictator-takings, if we compare this session with the other four Real Effort sessions (Mann-Whitney test, p-value = 0.980). Thus we pooled the data from

---

5 These questions are based on basic arithmetic concepts (e.g. algebra, geometry and data analysis).
6 The threshold of 13 correct answers was calibrated based on a pilot session of the GRE test among the undergraduate students of a seminar at Frankfurt University.
7 Before making their decision dictators were also told that the recipients knew (before they started the real-effort task) that a dictator will decide on the allocation.
8 Statistical analysis of this data revealed that only gender and age were significant variables, i.e. older people and women are more likely to take lower amounts. In contrast cultural differences (e.g. people’s religion) were not significant at all.
this session with the data from the other four Real Effort sessions.

The Windfall treatment was identical to the Real Effort treatment except that the recipients did not have the opportunity to take the quiz. Instead, the pot size was determined randomly. Subjects had to draw a lottery ticket worth either 5€ or 10€, and they had a 50% chance of winning either a low or high stake size. In order to keep both treatments comparable, dictators also had to wait for 30 minutes and we provided them with coffee and cake.

We used ORSEE (Greiner, 2004) to recruit the subjects among the undergraduate students at Frankfurt. A total of 352 subjects attended the experiment. We ran five sessions of each treatment. A session lasted about 75 minutes and on average subjects earned 8.75€ including a 5€ show-up fee. To maintain transparency, all subjects were informed about the whole procedure of the experiment.

3 Hypotheses

Our Real Effort treatment resembles a trust game, thus dictators benefit from recipients investments in effort. We hypothesize that this triggers reciprocal behavior. Since Croson and Buchan (1999) report that female second-movers are highly sensitive to reciprocal behavior, we expect female dictators in Real Effort to take lower amounts compared to the Windfall treatment. The trust game literature only reports increased reciprocity for female second-movers, i.e. we do not expect that male dictators are as sensitive to reciprocal behavior as female dictators.

HYPOTHESIS 1: Female dictator-taking will be lower in the Real Effort treatment compared to the Windfall treatment. In contrast, we do not expect male dictators to decrease their taking-rates by the same amount.

However, there exists a second level at which reciprocity may play a role. Depending on recipients’ performance, the size of the pot is either high or low in the Real Effort treatment, whereas, in the baseline treatment, the size of the pot is randomly determined. Dictators know in the Real Effort treatment that successful recipients will receive 10€ whereas less well-performing recipients get only 5€. Since the gender literature emphasizes reciprocity for women, we expect that female dictators especially reward successful recipients, i.e. if recipients generate 10€, dictators will take lower amounts.

HYPOTHESIS 2: In contrast to male dictators, female dictators in the Real Effort treatment take less from recipients who generate the large pot, i.e. in Real Effort there will be a stake size effect.

From Hypothesis 1 and 2 it follows that female dictators care about recipients’ performance. Thus positive reciprocity should additionally matter, when recipients worked and successfully generated a large pot. Furthermore, we expect that this will lead to a greater difference between female dictators’ taking-rates when they decide about high-pot-size recipients in the Real Effort compared to the Windfall treatment. In contrast, negative reciprocity might reduce their
generosity if recipients fail to get a high pot. That is: the difference in taking-rates between female dictators (in Windfall and Real Effort) will be larger when deciding about 10€ recipients compared to the case of 5€ recipients. Based on the evidence from gender trust games, men care much less about reciprocity. Thus we do not expect that men will show different magnitudes of reciprocity.

**HYPOTHESIS 3:** The difference in female dictators’ taking-rates between the Real-Effort and the Windfall treatments will be larger if recipients generated 10€ compared to the case where only 5€ was achieved. This effect should hold only for female dictators.

### 4 Results

This section starts with a brief outline of the average results. Afterwards, we analyze the gender differences and test our hypotheses. As our experiments are one-shot interactions, we count each participant as one observation in the statistical analysis. We report two-sided p-values and non-parametric Mann-Whitney tests throughout. Section 4.1 briefly summarizes dictator-taking in the Windfall and the Real Effort treatment (here we do not yet distinguish gender).

#### 4.1 Dictator-taking: Average results

Table 1 presents the means of dictators’ taking-rates in our two treatments separated into groups of 5€ and 10€ pot sizes. The standard deviations are in parentheses.

<table>
<thead>
<tr>
<th>Pot size</th>
<th>Windfall obs.</th>
<th>Real Effort obs.</th>
<th>Avg. obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 €</td>
<td>71.30 (22.42)</td>
<td>69.13 (24.90)</td>
<td>70.19 (23.61)</td>
</tr>
<tr>
<td>10 €</td>
<td>76.75 (21.50)</td>
<td>70.17 (24.20)</td>
<td>73.16 (23.12)</td>
</tr>
<tr>
<td>Avg.</td>
<td>73.93 (22.02)</td>
<td>69.67 (24.41)</td>
<td>71.68 (23.35)</td>
</tr>
</tbody>
</table>

Table 1: Mean of taken amounts (Windfall and Real Effort treatment)

On average, the real-effort task marginally triggers dictators’ reciprocity, i.e. dictators in the Windfall treatment take 73.93% compared to dictators in the Real Effort treatment who only take 69.67%. Nonetheless this small difference is statistically not significant (p-value = 0.223). There is also no treatment effect, if we focus on the 5€-pot (p-value = 0.709). Though if we concentrate on the 10€-pot, we find that dictators in Real Effort take 70.17% compared to dictators in the Windfall treatment who take 76.75% (one-sided p-value = 0.078). We therefore confirm the results of Ruffle (1998) and Oxoby and Spraggon (2008). Furthermore in the Windfall treatment we do not find a stake size effect (p-value = 0.238), this confirms Forsythe et al.’s (1994) findings. The same is true for the Real Effort treatment (p-value = 0.694). The brief analysis showed that dictators were prone to reciprocal behavior when recipients worked and successfully generated the large pot.
4.2 Dictator-taking: Gender effects

We now analyze dictator-taking, and separate the choices by gender in order to test Hypotheses 1-3. Table 2 presents male and female dictator taking-rates in the *Windfall* and *Real Effort* treatment.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Stake size</th>
<th>Windfall obs.</th>
<th>Real Effort obs.</th>
<th>Avg. obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>males</td>
<td>5€</td>
<td>68.73 (24.40)</td>
<td>74.21 (27.73)</td>
<td>72.10 (26.31)</td>
</tr>
<tr>
<td>males</td>
<td>10€</td>
<td>77.27 (21.62)</td>
<td>76.52 (25.65)</td>
<td>76.87 (23.38)</td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>73.81 (22.56)</td>
<td>75.39 (26.44)</td>
<td>74.71 (24.72)</td>
</tr>
<tr>
<td>females</td>
<td>5€</td>
<td>72.68 (21.62)</td>
<td>63.33 (20.33)</td>
<td>68.67 (21.38)</td>
</tr>
<tr>
<td>females</td>
<td>10€</td>
<td>76.11 (22.59)</td>
<td>63.26 (20.92)</td>
<td>68.90 (22.35)</td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>74.02 (21.82)</td>
<td>63.30 (20.40)</td>
<td>68.68 (21.70)</td>
</tr>
</tbody>
</table>

Table 2: Mean of taken amounts in our two treatments, split by gender and both stake sizes

In order to test Hypothesis 1, we compare females’ taking-rates in the *Windfall* and *Real Effort* treatment. Since reciprocity plays an important role in the *Real Effort* treatment, Hypothesis 1 predicts that female dictators will be strongly affected by the fact that recipients have worked.

Testing Hypothesis 1 we find a significant treatment effect for female dictators, i.e. the average taking-rate of female dictators is 63.30% in *Real Effort* compared to 74.02% in *Windfall* (p-value = 0.021). Therefore we reject the null hypothesis that female dictators take the same amount from recipients who did a real-effort task compared to the case where recipients received a windfall gain. We thus find strong support for Hypothesis 1. If we analyze male decisions, we do not find that the real-effort task stimulates reciprocal behavior, i.e. male dictators take 73.81% in *Windfall* and 75.39% in *Real Effort* (p-value = 0.720). Interestingly male dictator-taking is very stable. They always take around 75%. Furthermore this effect holds for all of our sessions.9 This also emphasized by Figure 1 which presents diagrams comparing the cumulative distribution functions (CDF) of male and female dictators in our two treatments. The left diagram shows the *Windfall* treatment and the right diagram the *Real Effort* treatment.

---

9We also ran three sessions (with 96 subjects) of the *Real Effort* treatment where dictators did not have to wait 30 minutes, but decided immediately. This was done in order to control for possible time effects. Even though male dictators decreased their takings by 10 percent points (note there is great variance in the data: Some male dictators took 0% and others took 100%) we find no statistical support for a difference (p-value = 0.249). Furthermore our results show that female dictators take exactly the same amounts in both variants of the *Real Effort* treatment (p-value = 0.891). We thank an anonymous referee for pointing out this issue.
Figure 1: Gender effects in the Windfall (left diagram) and Real Effort treatment (right diagram)

First: male and female CDFs do not differ at all when dictators decide about windfall money (KS-Test, Max. D = 0.062, p-value = 1.000). However, there exist crucial gender differences in the CDFs when dictators decide about money which has been generated by a Real Effort task (KS-Test, Max. D = 0.326, p-value = 0.011). It is remarkable that 57% of female dictators choose the equal split decision as opposed to only 37% of male dictators. Furthermore only 16% of female dictators take the whole pot from the recipients. This stands in strong contrast to 45% of male dictators who choose the 100% taking-rate.

RESULT 1:
Comparing the Windfall with the Real Effort treatment we find significant gender differences, i.e. in Real Effort, female dictators take considerably smaller amounts compared to Windfall. In contrast, male dictators do not reduce their taking-rates at all.

Table 2 shows that in the Real Effort treatment female dictators do not care about recipients’ performance: They take 63.33% from 5€ recipients and 63.26% from 10€ recipients (p-value = 0.852). Thus we cannot reject the null hypothesis which postulates that female dictators do not care about recipients’ performance. We therefore have to discard Hypothesis 2. Focusing on male dictator-taking in the Real Effort treatment, it appears that they also do not care about recipients’ performance, i.e. they take 74.21% from recipients who generated a small pot and 76.52% from recipients who generated the large pot (p-value = 0.833).

RESULT 2:
In the Real Effort treatment, female as well as male dictators do not take smaller amounts from successful recipients who generated the large pot.

In order to test Hypothesis 3 we now compare the treatment effect generated for the 5€ pot with the treatment effect for the 10€ pot. Focusing on female dictator-taking for the 5€ pot, we find that they take 72.68% in the Windfall treatment and 63.33% in the Real Effort treatment. However, this difference is not significant (p-value = 0.154).\textsuperscript{10} If we focus on the large pot, we

\textsuperscript{10}Note there exists weak significance for a one sided p-value, i.e. p-value = 0.077.
find that female dictators in the Windfall treatment take 76.11% compared to 63.26% in the Real Effort treatment. Thus the difference in taking-rates is larger when deciding about recipients who generated a large pot. In contrast to the 5€ pot, this difference is weakly significant (p-value = 0.069).\textsuperscript{11} We therefore reject the null hypothesis that the difference in female dictators’ taking-rates between 5€ recipients and 10€ recipients is not different. Thus we find support for Hypothesis 3. Interestingly females take nearly the same amounts from 5€ and 10€ recipients in Real Effort. Therefore, it cannot be that they were influenced by negative reciprocity in the case of 5€ recipients. Nevertheless, the difference in female taking-rates is larger if recipients generated a large pot. This is due to the fact that female dictators in Windfall take a larger amount from 10€ recipients compared to 5€ recipients.\textsuperscript{12} Male dictators do not show different magnitudes of reciprocity. That is: taking-rates from 5€-recipients in Windfall and Real Effort are not different (p-value = 0.540). The same is true for the treatment difference in taking-rates for the 10€ recipients (p-value = 0.947).

RESULT 3: 
The difference in female taking-rates, caused by the real-effort task is higher if recipients generated 10€ compared to the case where only 5€ was achieved. In contrast, male dictators do not show this behavior.

5 Discussion

Do women behave more reciprocally than men? The answer is yes.

We analyzed a modified dictator game with a real-effort task (based on Cherry et al. (2002)) where dictators were asked to dictate a money amount which was generated by the recipients. Our results show that women significantly decrease taking-rates when the recipients generated the money (to be divided) by a real-effort task instead of a lottery task. Furthermore, female dictators decreased taking-rates more strongly if recipients generated a large pot compared to the opposed case, where recipients only generated a small pot. In contrast, male dictators did not show reciprocal behavior at all, i.e. they did not lower their taking-rates in the environment of the real-effort task.

If we do not focus on gender, the general results show that dictators are sensitive to the real-effort task, however, the result depends on whether the recipient generated a large pot. That is: dictators only significantly lower taking-rates when the recipient was successful and generated 10€. Thus we confirm the results of Oxoby and Spraggon (2008) who argue that dictators are influenced by property rights legitimization.\textsuperscript{13} Further, our results are in line with Ruffle (1998)

\textsuperscript{11}Note there exists a significant difference for a one sided p-value, i.e. p-value = 0.035.
\textsuperscript{12}Probably they do not grant a large pot to recipients because the money was endogenized by pure chance.
\textsuperscript{13}However we extend their framework by applying a treatment where dictators decide about windfall money
who points out that dictators treat low performing recipients as if they did not work at all. Our study emphasizes real-effort’s impacts on gender differences due to reciprocal behavior. That is: we extend Ruffle’s (1998) and Oxoby and Spraggon’s (2008) studies and give an explanation for their findings.

It is interesting that we only find a significant effect for female dictators. Thus our study sheds new light on gender differences in reciprocal behavior driven by a real-effort task. Our paper therefore provides important new insights as an explanation for Ruffle’s (1998) and Oxoby and Spraggon’s (2008) results. These findings seem to provide valuable new insights in terms of other-regarding preferences induced by a real-effort task. For the future, it seems to be promising to uncover further gender differences in setups with real effort tasks.

References


which was won by recipients in a lottery. This enables us to emphasize the role of reciprocity induced by a real-effort task. Note there also exist other studies which emphasize the importance of endowments’ origins (e.g. Mittone and Ploner (2006), Cherry and Shogren (2008)

\[14\] Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Falk and Fischbacher (2006) highlight the importance of other regarding preferences in their models.


<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Authors</th>
<th>Journal/Publication Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>Technology Licensing by Advertising Supported Media Platforms: An Application to Internet Search Engines</td>
<td>Sapi, Geza and Suleymanova, Irina</td>
<td>Published in: The B. E. Journal of Economic Analysis &amp; Policy.</td>
</tr>
<tr>
<td>15</td>
<td>How Effective is European Merger Control?</td>
<td>Duso, Tomaso, Gugler, Klaus, Yurtoglu, Burcin</td>
<td>Forthcoming in: European Economic Review.</td>
</tr>
</tbody>
</table>


Herr, Annika, Quality and Welfare in a Mixed Duopoly with Regulated Prices: The Case of a Public and a Private Hospital, September 2010. Forthcoming in: German Economic Review.


