ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis: Behavioural Economics

Trust and Reciprocity towards Different Genders: An Experimental Analysis of Trust Games

Sjoerd van Velzen

456903

Supervisor: Merel van Hulsen

Date final version: 13-08-2019

Abstract:

This paper uses an experimental design in order to explore the differences in trusting behaviour and reciprocity towards men and women in the Netherlands, using the trust game introduced by Berg et al. in 1995. When analyzing the results of several t-tests, Fisher's exact tests and Mann-Whitney U tests, we find no evidence of significant differences in trusting behaviour towards men and women, nor do we find any proof of significant reciprocity differences towards men and women. Even though the differences were not significant, differences did present themselves. This leads to the conclusion that further research should be conducted, which could present significant results on the matter.

Table of contents

1. Introduction

In economics, when facing certain decision making problems, we often use models that implement perfect rationality, where the players are all assumed to be *homo economicus*, which Keynes describes as 'an economic man, whose activities are determined solely by the desire for wealth' (Keynes, 1890). However, actual strategic decision making deviates from the original theory in a way where trust and altruism is involved (Gintis, 2000).

When standard economic models fail in explaining people's behaviour, it is often because it disregards people's altruistic characteristics and their concerns for fairness (Gintis, 2000). In order to account for these characteristics, economists have developed models of social preferences, where altruism and reciprocity is involved. These models are useful in various situations such as consumer response to price changes, attitudes toward different tax schemes, and employee response to changes in wages and employment practices (Charness & Rabin, 2002). For instance, an employer could expect an increase in work rate if he increases his employee's wage, due to positive reciprocity.

Not only are trust and reciprocity important when dealing with situations like wage changes, but also when individuals are faced with incomplete contracts or deals (Sutter & Kocher, 2003). When dealing with incomplete contracts or deals, you are put into a position where you have to trust the other person on matters that cannot be put into contracts. A better understanding of trust can help with such contracts. Recent economic studies have revealed possible differences between men and women in trustworthiness and trusting behaviour (Eckel & Grossman, 1998). This paper will aim to expand on those differences, focusing on the differences between trust and reciprocity *towards* men and women instead.

When there is enough evidence for systematic differences between men and women when it comes to trust and altruism, this could have an important effect on current economic models regarding things like social dilemmas, charity donations, bargaining and household decision making (Andreoni & Vesterlund, 2001, Chaudhuri & Gangadharan, 2003). For instance, charitable donations could be improved with a better understanding of people's altruistic preferences towards certain groups or people.

As mentioned before, this paper will aim to explore the differences between trusting behaviour and reciprocity *towards* different genders instead of the differences *between*

different genders. This is a fairly unexplored part of this subject, which this paper will try to expand on. As such, the research question is as follows:

"Which differences exist in terms of trusting behaviour and reciprocity towards different genders?"

This paper consists of the following sections: Section 2 will start by explaining the original trust game by Berg et al. (1995), after which it will provide a brief overview of the existing literature regarding the effect of one's gender on trusting behaviour and reciprocity. Section 3 will provide the experimental design on which this research is based, and explain the data that was used and how it was changed in order to be able to perform both parametric and non-parametric tests. This section also takes a first look at the data by discussing the descriptive statistics. After that, Section 4 will provide the methodology of this paper, and explain the relevant statistical tests. Section 5 will lay out the results of the experiment and the statistical tests, after which Section 6 will provide a conclusion and discuss potential limitations and improvements of this research. Section 7 will consist of the list of all the references used in this paper, and finally Section 8 will contain all the Appendices.

2. Literature Review

This chapter will begin to explain the trust game and how it works. It will then aim to provide a brief review of the existing literature on the relationship of gender and trust/reciprocity in trust games, ultimatum games and dictator games. Subsequently, this will provide the information on which I based my hypotheses and explain them.

2.1.1 The Trust Game

This paper mentions and makes use of the original trust game introduced by Berg et al. in 1995, which was designed to measure trust and reciprocity in an investment setting. The game is designed as follows: for each game, there are two players; player A and player B. These are also known as the proposer and the responder. For the sake of this explanation, we will use the latter two.

First off, the proposer is given an initial endowment of €10. He or she gets the option to decide how much of that money they want to send to their counterpart, the responder. They are informed that the money that they send will be multiplied by X before it gets to the responder. The amount sent will be denoted by M. Thus, the amount that the responder receives will be X*M. The responder then gets the option to decide how much of the money that they received (X*M) they would like keep and how much they want to send back to their counterpart. The amount that they send back will be denoted by R. The final payoff for the proposer will then be €10 - M + R. The final payoff for the responder will be X*M - R.

In recent times, this trust game has been used for various purposes. For instance, it has been used in neuroscientific research. Kosfeld et al. (2005) used the trust game in order to measure the effects of an oxytocin increase on trusting behaviour, where they found a substantial increase in trust among humans. Other researchers made use of the trust game to see if prior social and moral information about potential trading partners affected the part of the brain that guides and adjusts future behavior on the basis of reward feedback (Delgado, Frank & Phelps, 2005). They found that such prior information diminishes the use of those brain areas that are involved in trial-and-error reward learning.

Trust games have also been used for genetic research, in order to see if individual differences in cooperative behaviour result from genetic variation. Cesarini et al. (2008) found that both willingness to invest and reciprocity in the trust game are influenced by a person's genetic endowment.

In this paper, the trust game will be used in order to check for differences in trusting behaviour towards different genders, where the amount sent (M) and the amount returned (R) are indicators of trust and trustworthiness, respectively (Brülhart & Usunier, 2012).

2.2.1 Relationship between responder's gender and the proposer's offer

In 1999, Croson & Buchan conducted an experiment where they tried to look for gender differences in the trust game. They found that women on average returned 37.4% to their counterpart, whereas men on average only returned 28.6%. This lead to their conclusion that women reciprocate significantly higher amounts compared to men. There is also evidence that trusting behaviour (i.e. the amount sent by the proposer) is driven strongly by expectations of reciprocation (Chaudhuri & Gangadharan, 2003), and that women are believed to be more cooperative (Gneezy, Niederle & Rustichini, 2003). Combining these findings suggests that people, both male and female, will send more money to women than to men, because they expect to receive more money in return.

However, other research indicates that there is no significant difference between offers sent to men and offers sent to women when playing the trust game. Buchan, Croson & Solnick (2008) used a TOBIT analysis to explore potential differences between offers sent to women and offers sent to men, but they only found an insignificant difference of one cent ($7.69 for men, $7.68 for women). This paper will be based on their experimental design, but adds to Buchan's paper by exploring other statistical analyses and investigating other target groups.
The paper by Buchan, Croson & Solnick also explored potential interaction effects, meaning potential differences between offers from women to women or women to men, or between offers from men to men or men to women. They did not find any interaction effects, leading to their conclusion that there is no difference in offers towards the same sex and the opposite sex.

In their paper on discrimination at Case Western Reserve University, Slonim & Guillen (2010) ran two types of trust games, one where the proposer could select their partner, and one where they could not. They found significant evidence of gender discrimination in the games where the subjects could choose their partner. Interestingly, they found that when given the option, proposers selected a partner of the opposite sex more often than the same sex. Men chose a female partner 75% of the time, and women opted to play with a male partner in 61% of the cases. This is in contrast to other research that found that people put

more trust in female partners due to higher expectations of reciprocity (Chaudhuri & Gangadharan, 2003; Gneezy, Niederle & Rustichini, 2003).

Additionally, Slonim & Guillen (2010) found that proposers did not only select the opposite gender more often, but also sent significantly more money on average to the opposite gender. Male proposers sent an average of $16.4 to female partners and an average of $10.4 to male partners. For women, these differences were smaller, having sent on average $11.3 to male partners and $9.9 to female partners.
These findings led to their conclusion that there is in fact a positive interaction effect of gender on amount sent.

Besides the trust game, there are also other games that can be used in order to test trust, such as the ultimatum game. In the ultimatum game, the money sent by the proposer is not tripled and the responder does not get the opportunity to return any money. The responder only gets the choice between accepting the offer, or declining it (leaving both the proposer and responder with a payout of zero).

In her research using the ultimatum game, Solnick (2001) found evidence that initial offers are influenced by the gender of the responder, which is in contrast to the findings of her paper using the trust game (Buchan, Croson & Solnick, 2008). She found evidence that men are actually the ones that receive higher offers, instead of women. Specifically, women on average received an offer of $4.37, whereas men on average received $4.73 from male proposers and $5.13 from female proposers. The results of Solnick's experiments showed that male responders tend to receive higher offers than female responders, and even higher offers from female proposers (Solnick, 2001). Additionally, studies show that women are expected to be of lower economic status than men due to stereotyping, which causes people to believe that women are content with lower offers than men are, leading to women receiving lower offers overall (Ball et al., 2001, Schwieren, 2003). This suggests both a direct effect as well as an interaction effect of gender on offers made between men and women.

Even though there is no unanimous answer by the existing literature, the consensus seems to be that there exists an interaction effect between the gender of the proposer and the gender of the respondent. This leads to the following hypothesis:

Hypothesis 1: On average, proposers send more money to responders of the opposite sex than to their own sex.

2.2.2 Relationship between proposer's gender and the responder's offer

In their paper on chivalry and solidarity, Eckel & Grossman (2001) conducted a series of ultimatum games where the gender of the opposing player was known. They found that men, on average, sent exactly $1.828 to both men and women. However, men rejected 18.8% of those offers by men whereas women only rejected 17.2%. For the offers made by women, the difference was even greater. Even though average offers by women were higher to men than to women ($1.992 compared to $1.981), men still rejected 9.4% of those offers, whereas women only rejected those offers in 3.1% of the cases.
This led to their conclusion that women respondents are more likely to accept a certain offer, and that offers that came from women are more likely to be accepted as well. This could mean that when playing the trust game, responders show greater reciprocity towards women because they 'accept' a low offer more if it comes from a woman rather than a man.

In contrast, there is also evidence that both men and women are actually less lenient towards women in the ultimatum game. In her paper on gender differences in ultimatum games, Solnick (2001) found that on average, responders accepted a $3 offer from men, but not from women. On average, responders had a minimum acceptable offer of $2.59 for male proposers and $3.73 for female proposers. This difference was explained by the difference in expectations that people had from women compared to men, where people on average expected an offer of $4.15 from men and $4.25 from women. This could mean that, in contrast to Eckel & Grossman's (2001) findings, responders would actually show greater reciprocity towards men instead of women when playing the trust game.

Evidence from a series of trust games conducted at the University of Melbourne showed that on average, men showed more trust (i.e. sent more money to the responder) than women, and women exhibited more reciprocity (i.e. sent more money back to the proposer) than men (Chaudhuri & Gangadharan, 2003). To be precise, men on average sent $5.30 to the responder whereas women only sent $3.47 on average. On the contrary, men on average only returned 14.7% of their endowment compared to women's return percentage of 19.8%. This could suggest that because men send more on average, and women reciprocate good behaviour more, women will send more back to men than to women.

In their paper on human generosity as a mating signal, Iredale, van Vugt & Dunbar (2008) conducted a series of experiments where they tested the generosity of people in the presence of potential mating partners, compared to when they were alone. The participants had to play a number of experimental games where they could earn money. After the

games, the participants were given the option to donate a certain amount of their winnings to charity. This amount donated to the charity was used as a measure of generosity. During the donation, they were either observed by someone or not. Their results showed that whilst women's donations did not change between when they were observed or not, men's donations were significantly higher when in presence of a female observer compared to when the observer was male or there was no observer at all (57% when observed by a female compared to 34% with no observer and 28.67% with a male observer). This could mean that in a trust game scenario, male responders will send back more money to women than to men.

The current literature shows evidence for different conclusions. Some say that reciprocity is higher towards women than towards men, and others say exactly the opposite. Most studies are unable to find a significant effect of the opposing player's gender on reciprocity, leading to the following hypothesis:

Hypothesis 2: There is no difference in reciprocity towards women than towards men.

3. Experimental Design

In total, 70 people participated in this experiment. Of the 70 participants, 36 were male and 34 were female. Ages ranged from 19 to 57 years old, with a mean age of 23.4 years. They were all either students or working people, and almost all had a Dutch nationality, with exceptions of one Romanian, one Italian and one Greek.

3.1 Experimental Design

This study will use a between-subject design where participants play the trust game only once, against a partner. The main advantage that a between-subject design has over a within-subject design is that between-subject designs are less susceptible to the 'demand effect', due to there not being any repetition of the experiment (Charness, Gneezy & Kuhn, 2012). The 'demand effect' is the effect where participants of an experiment tend to try to 'satisfy' the experimenter by providing answers that they think will confirm their hypothesis (Nichols & Maner, 2008). Using the between-subject design, combined with not giving the participants any information on the purpose of the experiment eliminates this bias.
Another advantage of between-subject designs is that when individuals only face a single decision, the external validity of a between-subject design might be better than that of a within-subject design, because it better represents the reality of single-decision situations (Charness, Gneezy & Kuhn, 2012).
Furthermore, because participants get exposed to multiple treatments in within-subject designs, the experimenter has to worry about the order in which each individual gets exposed to those treatments. To avoid the bias induced by individuals learning from repetition of the experiment, treatment orders have to be randomized by the experimenter. This is not the case for between-subject designs, because participants of a between-subject design only get one treatment.

There are also some disadvantages of using a between-subject design. To begin, the internal validity of a between-subject design depends on random assignment. This is not the case for a within-subject design, because the people with and without treatment are one and the same, meaning they do not differ intrinsically.
Furthermore, in many situations, a within-subject design has greater statistical power than a between-subject design (Charness, Gneezy & Kuhn, 2012).

Participants for this experiment were acquired through an online form that was spread on social media, asking them if they were willing to partake in a brief online experiment. In this

form, they had to fill in their full name, phone number and gender. This data was then collected and used to make “teams”. To make the teams, an online random team generator[1] was used in order to make the teams truly random. Each time, the first name on the team was selected as the proposer and the second name as the responder. These teams were then put into excel, after which all the proposers were sent a message containing the instructions to the game, and a set of questions they had to answer. The proposers were not specifically told what the gender of their responder was, in order to avoid them knowing about the intentions of this experiment which could cause a bias in the results. Instead, for proposers with a female responder, the words “she” and “female responder” were deliberately used when talking about the responder, making sure they knew the gender of their partner without exposing the goal of this experiment. Vice versa, for proposers with a male responder, the words “he” and “male responder” were used to obtain the same effect.

In the message to the proposers (see Appendix A1.1 and A1.2), in which the explanation of the experiment was given, the proposers could fill in how much money they wanted to send to their responder, as well as some background information including age, nationality, field of study and highest degree obtained. As soon as the proposer had sent their response, their answers were put into an excel sheet. Then, their respective responder was sent a message (see Appendix A2.1 and A2.2), again containing instructions for the experiment, which told them the amount that was sent to them by their proposer, as well as their following endowment which was three times the amount sent by the proposer. As in the message to the proposers, the message to the responders did not specifically mention the gender of the proposer, but made it clear which gender they were playing against by using the right pronouns and other gender-specific words. When the responder sent back their information and how much they wanted to return to the proposer, this was again put into the excel sheet in order to keep everything organized.

It is important to note that all participants were informed that at the end of the experiment, when all the data was collected, one of the pairs would be selected randomly, and those players would receive their individual payouts. This provides incentive for the players to actually choose what they think will give them the highest payout, because there is a chance that they will actually get that money.

[1] https://www.randomlists.com/team-generator

3.2 Data

The data for this study was retrieved from individuals living in the Netherlands. The data consists of a number of demographic variables such as age, gender, nationality, field of study and highest degree obtained, as well as the results of a trust game experiment in the form of money sent and received by each individual. These trust games were anonymous, where the players only knew the gender of the opposing player. In order to perform statistical analyses, the program StataMP was used.

Before analysing the data, some modifications were made, as well as some variables added. The variables '*gender*' and '*nationality*' were transformed into dummy variables. The dummy variable for gender takes on a value of 0 for men and 1 for women. The dummy variable for nationality takes on a value of 0 if the individual was Dutch, and 1 if the individual was not Dutch. This was done because of the 70 participants, only 3 people were not Dutch. Those three people were Romanian, Italian and Greek.

The original data for the variables '*field of study*' and '*highest degree obtaine*d' consisted of string data, so these were converted to categorical variables. For '*field of study*', the variable took on values between 0 and 5. This was based on the frequency of certain fields of study. This variable took on 0 for Economics, 1 for Medicine (including Dentistry and Movement Sciences), 2 for Law, 3 for Social Studies, 4 for Psychology and 5 for other fields of study.
The categorical variable for highest degree obtained ranges from 0 to 2, taking on 0 for people that completed high school and/or an MBO education, 1 for people that completed their Bachelor's degree, and 2 for people with a Master's degree.

Finally, some additional variables were created. First, a variable '*percentage returned*' was created which indicates what percentage of their total funds the responders had returned. This was calculated by dividing the amount returned by three times the amount that the proposer had sent, because the money that was sent to them got tripled. This variable was read by Stata as a string variable, so before analysing the data this was changed to numeric values using a simple command in Stata.

Secondly, two variables were created to show the final payoffs for both the proposer and the responder, called '*Payout P*' and '*Payout R*', respectively. The payout for the proposer was the sum of what they initially kept for themselves, plus the amount that the responder sent back to them. The payout for the responder was three times the amount that the proposer sent to them, minus the amount of money that they returned.

Third, a categorical variable called '*Gender Combination*' was created to indicate which combination of genders each game consisted of. This variable takes on value 0 for a male proposer and responder, value 1 for a male proposer and female responder, value 2 for a female proposer and responder and value 3 for a female proposer and male responder.

Additionally, one last variable '*oppositesex*' was created in order to be able to test the second hypothesis. This was done by taking the dummy values of '*Gender Proposer*' and '*Gender Responder*' and creating a formula that took on the right values. This formula is as follows: '*oppositesex'* = |*'Gender Proposer'* - *'Gender Responder'*|. This variable takes on a value of 0 when the game has two players of the same gender, and a value of 1 if the two players are of opposite genders.

Finally, some other variables were created in order to be able to perform Fisher's exact test correctly. For Fisher's exact test, it is useful to use a categorical variable instead of a continuous variable. This is why the variables '*OfferSize*', '*ReturnedSize*' and '*PercentageGroup*' were created from '*Eurossent*', '*Eurosreturned*' and '*percentage returned*', respectively. Each of these variables takes on values 0-3 depending on which group they are in. The variables were created in such a way that the total number of observations in each group was as equal as possible.

3.3 Descriptive Statistics

Before discussing the methodology of this research, we will briefly look at some descriptive statistics of the data to see if there are some interesting numbers at first glance. As can be seen in Table 1 below, it stands out that out of all the female proposers, not a single one decided to keep their endowment of €10 and send nothing (Table 1). Furthermore, you can see that the average amount sent by men is substantially lower than the average amount sent by women (€6.50 for men compared to €8.21 for women). This is interesting, because it is in contrast with earlier findings of men showing more trust and higher offers than women (Chaudhuri & Gangadharan, 2003; Slonim & Guillen, 2010).

Table 1: Euros sent, by gender combination, in €

Gender Proposer	Gender Combination	Obs.	Mean	Std. Dev.	Min.	Max.
Male	MM	9	6.39	3.2766	0	10
	MF	9	6.61	3.7398	0	10
	Total	18	6.50	3.4128	0	10
Female	FF	8	7.75	2.6049	3	10
	FM	9	8.61	1.4954	6	10
	Total	17	8.21	2.0696	3	10

When looking at the descriptive statistics of the final payouts of proposers and responders separately (Table 2), you can see that even though some responders ended up with a final payout of €0, the average payout for responders was higher than that of proposers, especially for males. This difference between men and women could be explained by the finding of Croson & Buchan (1999) that women tend to reciprocate more than men, resulting in more equal payouts for women than for men. The overall difference is not surprising, considering the fact that responders get to 'choose' the final payouts of both players, which leads to them choosing to have a slightly higher payout that the proposer.

Table 2: Final payouts, by gender, in €

	Gender	Obs.	Mean	Std. Dev.	Min.	Max.
Proposer	Male	18	11.56	2.5315	8	15
	Female	17	12.03	2.6485	6.5	15
	Total	35	11.79	2.5618	6.5	15
Responder	Male	18	13.28	4.3999	0	20
	Female	17	12.44	4.7033	0	20
	Total	35	12.87	4.5022	0	20

When looking at the difference between percentages returned to men and women in Table 3, the results look virtually identical, apart from a 2.09% difference between average percentages returned. For both male and female proposers, the maximum percentage returned was 50%, which is not surprising since any return percentage above 50% would mean that the responder was content with the proposer ending up with more money than him/herself. The lowest percentage returned to both men and women was 11.11%, which were both offers of €3 that were rewarded with only €1 in return, which is the 'fair' response since both players end up with €8 in this scenario.

Another thing that stands out is that all the female proposers got at least some money back, whereas there were some male proposers that received nothing (Table 3). This is due to the fact that there were some male proposers that sent nothing (and kept the full €10), and thus automatically got nothing in return. For both males and females, there were no responders that received money but did not return any money at all.

Table 3: average % and amount returned to men and women, in €

	Gender Responder	Obs.	Mean	Std. Dev.	Min.	Max.
% returned	To men	16	37.61%	.1248	11.11%	50%
	To women	17	39.70%	.1180	11.11%	50%
	Total	33[2]	38.69%	.1191	11.11%	50%
Amount returned	To men	18	8.06	5.5780	0	15
	To women	17	10.24	4.2980	1	15
	Total	35	9.11	5.0470	0	15

[2] 33 observations in total (instead of 35), due to the two proposers that sent nothing, which causes the variable '% returned' to be a missing value (because divided by 0)

4. Methodology

In order to analyse our data, several statistical tests will be performed. In this section, these tests and their assumptions will be discussed, as well as some information on significance levels and statistical power.

4.1 Two-sample t-test

The main and most basic test that will be used is the two-sample t-test. This test compares the means of two separate samples to see if they differ from each other significantly. If the p-value of the test statistic is smaller than 5% or 10% (depending on which significance level you use), the means differ significantly from each other (Cressie & Whitford, 1986).

This test will be used in order to compare several means. It will provide a basic first look at whether or not certain means differ from each other significantly. These include mean amounts sent to/by men and women, mean amounts/percentages returned to men and women and mean amounts sent to players of the opposite sex compared to the same sex.

The two-sample t-test relies on different assumptions. Firstly, this test assumes that the samples are selected randomly from their population. The second assumption is the normality assumption, which implies that both sample groups should follow a normal distribution. The third assumption is the homogeneity assumption, which assumes equal variances for both groups. Finally, it is assumed that the data is continuous (not discrete) and that the two groups are independent from each other.

The first assumption of this test is met because the samples consist of random people who signed up to participate in this experiment, and the selection of proposers and responders within this sample was done through a random number generator[3].

The normality assumption can be tested by computing the Shapiro-Wilk W test for normality. Running this test shows that the variable '*Eurosreturned*' follows a normal distribution (p-value = 0.6308), but variables '*Eurossent*' and '*Percentage Returned*' follow a non-normal distribution (p-value = 0.0002 and 0.0005, respectively) (see Appendix B1.1). However, it is argued that when sample sizes are equal, relatively large and when using a two-tailed test instead of a one-tailed test, t-tests are reasonably robust to non-satisfaction of the normality

[3] https://www.randomlists.com/team-generator

assumption (Sawilowsky & Blair, 1992). Our sample is consistent with all of these conditions, so we will continue as if this assumption is satisfied.

The homogeneity assumption can easily be tested in Stata using the test for equal variances, which rejects the null hypothesis of equal variances if the p-value<0.05. Results of these tests show that all groups that are compared have equal variances, with all tests producing non-significant p-values (see Appendices B2.1-5).

The final two assumptions are also satisfied, because the data set is continuous and participants did not know each other's answers, so observations are independent.

4.2 Fisher's Exact test

Fisher's Exact test is a non-parametric test, used to test if the observed sample outcomes of two groups differ from random (=equal) assignment of the classes. This test is chosen because in our data set we are dealing with small samples, which Fisher's exact test is robust to.

In this research, Fisher's exact test was computed using two groups and four classes. These groups were either determined by the gender of the proposer (male or female) or by whether or not the opposing player was of the opposite sex (yes or no). The classes were created as such that the four classes were as equal in size as possible. The variables that were used as classes were '*OfferSize*', '*ReturnedSize*' and '*PercentageGroup*'.
For '*OfferSize*', the classes were €0 to (but not including) €6, €6 to (but not including) €8, €8 to (but not including) €10 and €10. For '*ReturnedSize*', these classes were €0-€5, €5-€9, €9-€15 and €15. Finally, for '*PercentageGroup*', the classes were 11.11%-33.33%, 33.33%-38.89%, 38.89%-50% and 50%.

As with nearly all statistical tests, this test relies on certain assumptions. First off, this test assumes that the samples are selected randomly from their population. Additionally, it also assumes independence, meaning that observations are not affected by other observations. Finally, this test assumes mutual exclusivity, which implies that observed values should only fall in one group and class.

The first two assumptions are satisfied, as explained in section 4.1. The final assumption is also met, because observations can only fall in one group (male or female) and one class

(classes are made by several hard boundaries, which implies that one observation cannot belong in two classes at the same time).

Fisher's exact test outputs a one-tailed p-value indicating the likelihood of obtaining the observed values when distribution is random (=equal). When the p-value is small (<0.10), the distributions differ from each other significantly.

4.3 Mann-Whitney U test (Wilcoxon Rank-sum test)

The Mann-Whitney U test is a non-parametric test that tests if two independent groups differ from each other significantly. This test ranks all the data and then compares the sum of all the ranks of each group to see if they are different or not. The null hypothesis of the Mann-Whitney U test states that it is equally likely that a random observation from one group is greater than or smaller than a random observation from the other group. When the p-value of this test is smaller than 10%, it means that the null hypothesis can be rejected and there are significant differences between the groups. When the distributions of the two groups have a different shape, the Mann-Whitney U test is used to test if the distributions are different. If the distributions have the same shape, the Mann-Whitney U test is used to test if the medians are different. This test is applicable for our data because it provides a non-parametric alternative to the two sample t-test. We will use this test to test for significant differences in amounts sent to the opposite sex compared to the same sex, and for amount/percentage returned to men and women.

Unlike the two-sample t-test, the Mann-Whitney U test does not assume normal distributions. However, when testing for differences in distributions, it does assume that the distributions from both groups are the same, except for location. Furthermore, this test assumes that the samples are selected randomly from their population. The third assumption of this test is the independence assumption, which implies that the two groups are independent of each other and of other observations. Finally, the test also requires a continuous dependent variable (not discrete).

As can be seen in Appendices B3.1-3, all compared groups have roughly the same distributions, so we can use the Mann-Whitney U test to see if the medians are different.
As explained in section 4.1, the random sample assumption and independence assumption are also satisfied.
Finally, the dependent variable of interest is a continuous variable each time, so the last assumption is also met.

4.4 Significance levels and statistical power

The rule of thumb when performing statistical analyses is that there is a significant difference when the p-value<0.05. However, because the data set is split into four groups (one for every gender combination), we are working with relatively small samples. When dealing with small sample sizes, significance levels of 10% are also accepted (Lindley & Scott, 1995). So in this paper, p-values lower than 10% will be seen as significant.

Statistical power is defined as "the probability that the H_0 will be rejected when it is false, that is, the probability of obtaining a statistically significant result" (Cohen, 1992). In other words, it is the probability that an effect will be detected when there is an effect there to be detected. Statistical power is inversely related to the probability of making a Type II error, which is the probability of detecting an effect when in reality there is no effect to be detected. Vice versa, the probability of making a Type I error is the probability of rejecting the null hypothesis when the null hypothesis in reality is true (Banerjee et al., 2009).
Power is affected by four parameters: the sample size (N), the effect size, the significance level used (α) and the statistical power (β). These four parameters are mathematically related, meaning that if you know three of them, you can calculate the fourth one.

Unfortunately, for any of our tests, the statistical power does not exceed 0.6, which is not high enough (rule of thumb dictates that statistical power should be at least 0.8). This can be explained by the fact that we are not only dealing with small sample sizes, but also with small effect sizes. This is a limitation of this paper, and should be mitigated in future research by increasing the sample size.

5. Results

5.1 Amounts sent by men and women

As mentioned in Section 3.3, when looking at the descriptive statistics, the difference between money sent by men and women is notable (€6.50 for men compared to €8.21 for women). Results of a two sample t-test (see Appendix C1.1) show that this difference is statistically significant on the 10% level (two-tailed p-value = 0.0851<0.10). Fisher's exact test shows a different result, with a one-tailed p-value of 0.168 (See Appendix C1.2), meaning that the differences in distributions of amount sent by men and women are not statistically significant. This is in contrast to what Eckel & Grossman found in 2001, where they found that offers from women were higher than those from men.

5.2 Amount sent to men and women

In order to test our first hypothesis, we first look at the differences between the amount of money sent to men and women. On average, proposers sent €7.50 to male responders and €7.15 to female responders. Results of a two sample t-test show that this difference is not statistically significant (two-tailed p-value = 0.7274). When looking at male and female proposers separately, two sample t-tests give the same result with one-tailed p-values of 0.4475 and 0.2047 for male and female proposers, respectively[4] (See Appendix C2.1).

The data shows that there is a difference of €0.58 (€7.61 compared to €7.03) between average money sent to the opposite sex compared to the same sex. This supports the first hypothesis, where we predicted that people would send more money to players of the opposite sex compared to the same sex. A two sample t-test (See Appendix C2.2) shows that this difference is not statistically significant (one-tailed p-value = 0.2825)[5]. It has to be noted that with small sample sizes such as this one (N=35), it is much harder to obtain significant differences than with bigger sample sizes (Lindley & Scott, 1995). This will further be discussed in the discussions section.

Results of a Fisher's exact test also show that there is no significant difference in distributions of amount sent to the opposite sex compared to the same sex (Fisher's exact p-

[4] Here, one-tailed p-value is used because we are testing for a positive difference between money sent to the opposite sex versus the same sex.
[5] Again, one-tailed p-value is used because we are looking for a positive difference between money sent to the opposite sex versus the same sex.

value = 0.828). The Mann-Whitney U test produces the same results, showing a non-significant difference (p-value = 0.4655) (see Appendices C2.3 and C2.4)
All tests show a non-significant difference, so we must reject our first hypothesis of a positive difference between money sent to the opposite sex compared to the same sex. To put it in economic terms, this means that there is no influence of the opposing player being from the same sex when it comes to offers made in a trust game scenario.

5.3.1 Amount returned to men and women

There are two ways to look at reciprocity in this trust game. Either you look at the actual amount returned, or the percentage returned (the percentage of their total endowment, which depends on how much the proposer sent them). Both ways will be analysed, but using the percentage returned method has some benefits over the actual amount returned method. This is because in the cases where proposers decided to send nothing, the responders did not have a choice to send anything back, meaning that the value for amount returned was 0 by default.
The percentage returned method calculates the percentage returned based on the amount that the proposer sent, meaning that for those cases where proposers sent nothing, the variable percentage returned would contain a missing value. When performing statistical tests, these missing values are not included, giving a better perspective on actual reciprocity in the game.

When looking at the results of the amount returned to men and women (see Section 3.3, Table 3), there is a noticeable difference of €2.18 between the average amount returned to women compared to men (€10.24 to women compared to €8.06 to men). To check the statistical significance of this difference, a two-sample t-test was performed (see Appendix C3.1). This showed that this difference was not statistically significant (two-tailed p-value = 0.2062).
Furthermore, when checking for a difference between amount returned to men and women by men and women separately, no significant differences were found (two-tailed p-value = 0.2984 and 0.4920 for men and women, respectively) (see Appendix C3.1).

When looking at the results of Fisher's exact test and the Mann-Whitney U test, both show that the distributions of the amount returned to men and women do not differ significantly from each other (Fisher's exact p-value = 0.394, Mann-Whitney U p-value = 0.2285) (See Appendices C3.2 and C3.3).

The two sample t-test, Fisher's exact test and the Mann-Whitney U test all show no significant difference between amount returned to men and women. Thus, we cannot reject our second hypothesis of no significant difference between reciprocity towards men and women. In other words, people do not punish or reward men and women differently in trust game scenarios.

5.3.2 Percentage returned to men and women

A better way to look at reciprocity might be to look at the percentage returned instead of the amount returned, because it accounts for possible differences in the amount that the responders received from the proposers. Also, it accounts for the cases where the proposers decided to send nothing. In these cases, the percentage returned would be a missing value (because you cannot divide by 0), so these cases are omitted from the statistical tests, as explained in Section 5.3.1.

When running a two-sample t-test on the difference in percentage returned to men and women (see Appendix C4.1), we find that the difference is not significant (two-tailed p-value = 0.6251). Neither was the difference for male and female responders separately (two-tailed p-value = 0.7824 and 0.7045 for male and female responders, respectively).
Results of Fisher's exact test and the Mann-Whitney U test show a p-value of 1.000 and 0.5175 respectively, meaning that the distributions of percentage returned to men and women do not differ from each other significantly (see Appendix C4.2 and C4.3).

All tests suggest that the percentages returned to men and women do not differ significantly from each other, so we cannot reject our second hypothesis of no significant difference. This again shows that people do not punish or reward men and women differently in trust game scenarios.

6. Conclusion and discussion

This part will aim to summarize the paper and its conclusions. With these conclusions, we will try to answer the research question "*Which differences exist in terms of trusting behaviour and reciprocity towards different genders?*". Finally, we will discuss some possible improvements to this study in order to accommodate further research.

6.1 Conclusion

This paper aimed to explore the effects of the gender of the opposite sex in trust games, in order to see if any differences exist in terms of trust and reciprocity towards different genders. Using an experimental design based on the original trust game (Berg et al., 1995), we tested the hypotheses in order to formulate an answer to our research question. When testing our first hypothesis, we came to the conclusion that there was no significant difference between money sent to responders of the opposite sex compared to those of the same sex. This is in accordance with the findings of Buchan, Croson & Solnick (2008), who concluded that men and women are equally skilled when it comes to predicting how much they will get in return, but that men sent more because their predicted returns were higher.

Furthermore, we did not find any difference in reciprocity towards men and women, confirming our second hypothesis. This is in contrast with the findings of Chaudhuri & Gangadharan (2003), who found that women exhibit higher reciprocity than men. To further understand these differences, future studies could look further into the expectations that people have of other people, which could help explain the at times ambiguous outcomes of these experiments.

To conclude this section, we will provide an answer to the research question: *"Which differences exist in terms of trusting behaviour and reciprocity towards different genders?"*. Based on the data, we found small differences in offers sent to men and women, and somewhat bigger differences in reciprocity towards men and women. These differences could have implications on how people should approach scenarios like wage changes and contract negotiations, as discussed in the introduction. However, we did not find any proof of those differences in trusting behaviour and reciprocity towards men and women being statistically significant, suggesting that it should not make a difference whether you are dealing with a man or a woman in these scenarios. However, because of several limitations of this research (which will be discussed in the following section), this answer is not

definitive, and thus should not be interpreted as fact, but instead as a base for future research.

6.2 Discussion

In this section, we will discuss the limitations of this paper and propose possible improvements for further research on the topic.

First off, in order to correctly investigate the effects of the other player's gender in trust games, you need a sufficiently large sample. Because we are exploring trust games, we need two players for each game, so two individuals for only one observation. Furthermore, because we explore gender effects, we also have to split our games in four groups: M-M, M-F, F-F and F-M. This meant that with the 70 individuals that played the game, we only had 8 or 9 observations for each gender combination. This makes it hard to find significant results (Lindley & Scott, 1995), and also discredits the results that we do find. With more funding and more time, this could be resolved in future research.

Also, not only was the sample size small, but it could also suffer from selection bias. Because almost all of the subjects were either friends or family of mine, they automatically share some characteristics (both intrinsically and extrinsically). This showed in the demographics. Out of the 70 individuals, 67 were Dutch, 32 studied in the field of Economics and 57 only had their high school diploma. Because most individuals were highly educated and/or studying in the field of Economics, it is likely that they were aware of the 'mistakes' people make in trust games. They most likely knew how to optimally play the game (which is to send everything), causing a positive selection bias, because uneducated people might have sent less because they do not know the optimal strategy. For further research, individuals should be more randomly selected in order to avoid selection bias.

A third limitation of this research is the way in which the messages to the subjects were worded. I decided to not specifically mention that the subject was playing against a male or female, because that could have influenced their answers if they had deduced that this experiment focused on gender differences. Instead, I decided to use gender-specific words like him/her, he/she and male/female. This might have ensured that the subject did not know the purpose of this research, but on the other hand might have meant that the subjects disregarded the gender of the opposite player entirely, impacting the internal validity of the experiment. For further research, this could be mitigated by having a control group that does

not know the other's gender at all, and then comparing that to the group that does. This, however, requires a very big sample size, which is only attainable with sufficient funding and time.

Lastly, because the funds for this research were severely limited, it was impossible to give every single subject their final payout. In order to provide at least some incentive besides possible intrinsic motivations, the subjects were informed that in the end, one of the pairs would be selected at random, and they would receive their payout. This way, the players knew they had a chance to win and thus were incentivized to choose the amount that they thought would give them the highest final payout. However, some players might have disregarded this lottery because they believed their chances to win to be too slim. This could have caused them to not care about the outcome enough, to where they did not play the (in their mind) optimal strategy. Considering that most individuals were highly educated and studying in the field of Economics, it could have caused them to send less than €10, which is the optimal strategy. This could have created a negative bias in the measures of trust that was put into opposing players.

To conclude, further research could be improved by increasing the sample size and ensuring its randomness, and providing sufficient incentives in order to ensure the internal validity of the experiment.

7. References

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

Ball, S., Eckel, C., Grossman, P. J., & Zame, W. (2001). Status in markets. *The Quarterly Journal of Economics*, *116*(1), 161-188.

Banerjee, A., Chitnis, U. B., Jadhav, S. L., Bhawalkar, J. S., & Chaudhury, S. (2009). Hypothesis testing, type I and type II errors. *Industrial psychiatry journal*, *18*(2), 127.

Berg, J., Dickhaut, J., & Mccabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior, 10*(1), 122-142. doi:10.1006/game.1995.1027

Brülhart, M., & Usunier, J. (2012). Does the trust game measure trust? *Economics Letters,115*(1), 20-23. doi:10.1016/j.econlet.2011.11.039

Buchan, N. R., Croson, R. T., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization*, *68*(3-4), 466-476.

Cesarini, D., Dawes, C. T., Fowler, J. H., Johannesson, M., Lichtenstein, P., & Wallace, B. (2008). Heritability of cooperative behavior in the trust game. *Proceedings of the National Academy of sciences*, *105*(10), 3721-3726.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, *81*(1), 1-8.

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

Chaudhuri, A., & Gangadharan, L. (2003). Gender differences in trust and reciprocity. *Department of Economics Working Paper Series 248*

Cressie, N. A. C., & Whitford, H. J. (1986). How to use the two sample t- test. *Biometrical Journal*, *28*(2), 131-148.

Croson, R., & Buchan, N. (1999). Gender and culture: International experimental evidence from trust games. *American Economic Review*, *89*(2), 386-391.

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature neuroscience*, *8*(11), 1611.

Eckel, C. C., & Grossman, P. J. (1998). Are women less selfish than men?: Evidence from dictator experiments. *The economic journal*, *108*(448), 726-735

Eckel, C. C., & Grossman, P. J. (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry*, *39*(2), 171-188

Gintis, H. (2000). Beyond Homo economicus: evidence from experimental economics.

Ecological economics, *35*(3), 311-322

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, *118*(3), 1049-1074.

Iredale, W., Van Vugt, M., & Dunbar, R. (2008). Showing off in humans: Male generosity as a mating signal. *Evolutionary Psychology*, *6*(3), 147470490800600302.

Keynes, J. N. (1890). The scope and method of political economy. *History of Economic Thought Books*.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*(7042), 673.

Lindley, D. V., & Scott, W. F. (1995). *New Cambridge statistical tables*. Cambridge University Press.

Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: Investigating participant demand characteristics. *The Journal of general psychology*, *135*(2), 151-166.

Sawilowsky, S. S., & Blair, R. C. (1992). A more realistic look at the robustness and type II error properties of the t test to departures from population normality. *Psychological bulletin*, *111*(2), 352.

Schwieren, C. A. A. (2003). The gender wage gap - due to differences in efficiency wage effects or discrimination? (METEOR Research Memorandum; No. 046). Maastricht: METEOR, Maastricht University School of Business and Economics.

Slonim, R., & Guillen, P. (2010). Gender selection discrimination: Evidence from a Trust game. *Journal of Economic Behavior & Organization*, *76*(2), 385–405. https://doi.org/10.1016/j.jebo.2010.06.016

Solnick, S. J. (2001). Gender differences in the ultimatum game. *Economic Inquiry*, *39*(2), 189-200.

Sutter, M., & Kocher, M. G. (2003). Age and the development of trust and reciprocity. *Available at SSRN 480184*.

8. Appendix

Appendix A: Messages to proposers and responders

A1.1 Message sent to proposers with male responders

Hello, thank you for participating in this experiment! It will only take a minute or two.

With this experiment, there is a chance for you to win some money. The experiment is played in pairs. In the end, one pair of players will win the amount of money that they ended up with after the game. This pair will be selected randomly.

The experiment goes as follows:

In this game you have two players, **player A** and **player B**, who will remain anonymous to each other. Player A is given an initial endowment of €10. He or she has the option to send some (or all, or nothing) of that money to Player B, which is tripled before it gets there. Player B then gets the same option, to return a certain portion of the money that they received to player A (this **does not** get tripled again). In the end, the final payoff for player A will be what they kept of their initial endowment plus the amount player B returned. For player B, the final payoff will be what they kept of the money that was given to them by player A.

You were randomly selected as **player A**.

You now have the choice of how much of your endowment of €10 you will send to your male partner.

(1) How much money do you want to send to your male partner?

Total endowment: €10

You keep: €(1a)

You will send him: €(1b)

(2) How old are you?

(3) What is your gender?

(4) What is your field of study?

(5) Highest degree obtained?

(6) What is your nationality?

Thank you for participating! As soon as your partner made his choice, I will let you know how much he returned.

A1.2 Message sent to proposers with female responders

Hello, thank you for participating in this experiment! It will only take a minute or two.
With this experiment, there is a chance for you to win some money. The experiment is played in pairs. In the end, one pair of players will win the amount of money that they ended up with after the game. This pair will be selected randomly.

The experiment goes as follows:
In this game you have two players, **player A** and **player B**, who will remain anonymous to each other. Player A is given an initial endowment of €10. He or she has the option to send some (or all, or nothing) of that money to Player B, which is tripled before it gets there. Player B then gets the same option, to return a certain portion of the money that they received to player A (this **does not** get tripled again). In the end, the final payoff for player A will be what they kept of their initial endowment plus the amount player B returned. For player B, the final payoff will be what they kept of the money that was given to them by player A.

You were randomly selected as **player A**.
You now have the choice of how much of your endowment of €10 you will send to your female partner.

(1) How much money do you want to send to your female partner?

Total endowment: €10
You keep: €(1a)
You will send her: €(1b)

(2) How old are you?
(3) What is your gender?
(4) What is your field of study?
(5) Highest degree obtained?
(6) What is your nationality?

Thank you for participating! As soon as your partner made her choice, I will let you know how much she returned.

A2.1 Message sent to responders with male proposers

Hello, thank you for participating in this experiment! It will only take a minute or two.
With this experiment, there is a chance for you to win some money. The experiment is played in pairs. In the end, one pair of players will win the amount of money that they ended up with after the game. This pair will be selected randomly.

The experiment goes as follows:
In this game you have two players, **player A** and **player B**, who will remain anonymous to each other. Player A is given an initial endowment of €10. He has the option to send some (or all, or nothing) of that money to Player B, which is tripled before it gets there. Player B then gets the same option, to return a certain portion of the money that they received to player A (this **does not** get tripled again). In the end, the final payoff for player A will be what they kept of their initial endowment plus the amount player B returned. For player B, the final payoff will be what they kept of the money that was given to them by player A.

You were randomly selected as **player B**.
Your male partner was given the initial endowment of €10. He decided to send €? to you, meaning you get €(3*?) to work with in total.

(1) How much money do you want to send back to your male partner?

Total endowment:	€(3*?)
You keep:	€(1a)
You will send him:	€(1b)

(2) How old are you?
(3) What is your gender?
(4) What is your field of study?
(5) Highest degree obtained?
(6) What is your nationality?

Thank you for participating! I will contact you soon if you have won the lottery to send you your payout!

A2.2 Message sent to responders with female proposers

Hello, thank you for participating in this experiment! It will only take a minute or two.
With this experiment, there is a chance for you to win some money. The experiment is played in pairs. In the end, one pair of players will win the amount of money that they ended up with after the game. This pair will be selected randomly.

The experiment goes as follows:
In this game you have two players, **player A** and **player B**, who will remain anonymous to each other. Player A is given an initial endowment of €10. She has the option to send some (or all, or nothing) of that money to Player B, which is tripled before it gets there. Player B then gets the same option, to return a certain portion of the money that they received to player A (this **does not** get tripled again). In the end, the final payoff for player A will be what they kept of their initial endowment plus the amount player B returned. For player B, the final payoff will be what they kept of the money that was given to them by player A.

You were randomly selected as **player B**.
Your female partner was given the initial endowment of €10. She decided to send €? to you, meaning you get €(3*?) to work with in total.

(1) How much money do you want to send back to your female partner?

Total endowment:	€(3*?)
You keep:	€(1a)
You will send her:	€(1b)

(2) How old are you?
(3) What is your gender?
(4) What is your field of study?
(5) Highest degree obtained?
(6) What is your nationality?

Thank you for participating! I will contact you soon if you have won the lottery to send you your payout!

Appendix B: Assumptions

B1 Shapiro-Wilk W tests for normality

Table B1.1: Results of Shapiro-Wilk W tests for variables '*Eurossent*', '*Eurosreturned*' and '*Percentage Returned*'

Variable	Obs.	W	V	z	P-value
Euro's sent	35	0.85064	5.331	3.493	0.00024
Euro's returned	35	0.97612	0.852	-0.334	0.63078
% Returned	33	0.85655	4.897	3.304	0.00048

B2 Standard deviation tests for homogeneity

Table B2.1: Results of test for homogeneity of variances of '*Eurossent*' by male and female proposers

	Gender Proposer						95% CI for Combined Mean			
	Male			Female						
	M	SD	n	M	SD	n		f	p-value	df
Euro's sent	6.50	3.41	18	8.21	2.07	17	6.32, 8.34	2.72	0.052	17,16

Notes: M=Mean, SD=Standard Deviation, n=# of observations

Table B2.2: Results of test for homogeneity of variances of '*Eurossent*' to male and female responders

	Gender Responder						95% CI for Combined Mean			
	Male			Female						
	M	SD	n	M	SD	n		f	p-value	df
Euro's sent	7.50	2.72	18	7.15	3.21	17	6.32, 8.34	0.72	0.507	17,16

Notes: M=Mean, SD=Standard Deviation, n=# of observations

Table B2.3: Results of test for homogeneity of variances of '*Eurossent*' to same sex and opposite sex

	Opposite Sex						95% CI for Combined Mean			
	No			Yes						
	M	SD	n	M	SD	n		f	p-value	df
Euro's sent	7.03	2.97	17	7.61	2.95	18	6.32, 8.34	1.02	0.972	16,17

Notes: M=Mean, SD=Standard Deviation, n=# of observations

Table B2.4: Results of test for homogeneity of variances of '*Eurosreturned*' to male and female proposers

	Gender Proposer						95% CI for Combined Mean			
	Male			Female						
	M	SD	n	M	SD	n		f	p-value	df
Euro's returned	8.06	5.58	18	10.24	4.30	17	7.38, 10.85	1.68	0.303	17,16

Notes: M=Mean, SD=Standard Deviation, n=# of observations

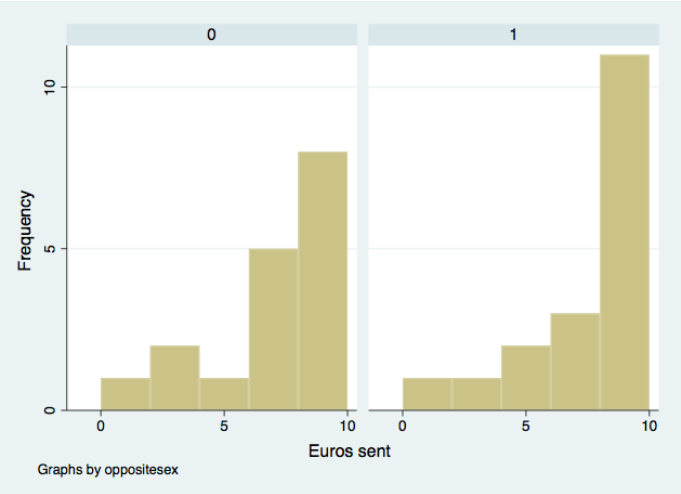
Table B2.5: Results of test for homogeneity of variances of '*Percentage Returned*' to male and female proposers

	Gender Proposer						95% CI for Combined Mean			
	Male			Female						
	M	SD	n	M	SD	n		f	p-value	df
% returned	.3761	.125	16	.3970	.1180	17	.3443, .4294	1.12	0.823	15,16

Notes: M=Mean, SD=Standard Deviation, n=# of observations

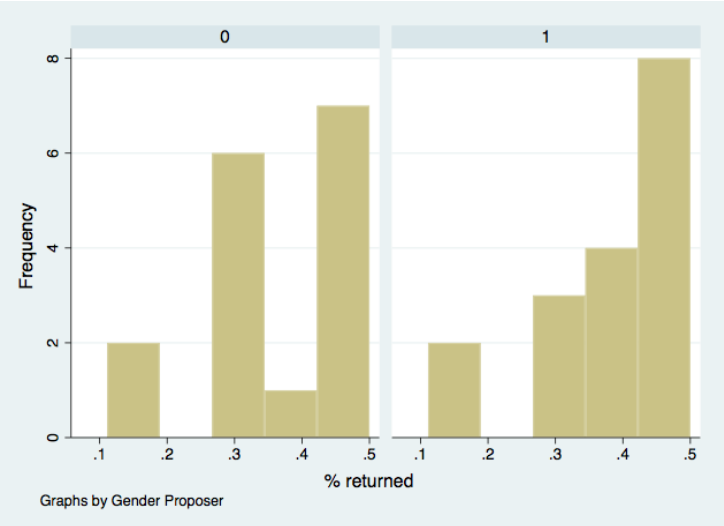
B3 Distribution graphs

Graph B3.1: Histograms of '*Eurossent*' to same sex and opposite sex



Graph B3.2: Histograms of '*Eurosreturned*' to male and female proposers

Graph B3.3: Histograms of '*Percentage Returned*' to male and female proposers



Appendix C: Results

C1 Amounts sent by men and women

Table C1.1: Results of two sample t-test and descriptive statistics for '*Eurossent*' by '*GenderProposer*'

	Gender Proposer						95% CI for Mean Difference			
	Male			Female						
	M	SD	n	M	SD	n		t	p-value	df
Euro's sent	6.50	3.41	18	8.21	2.07	17	6.32, 8.34	-1.77	0.085*	33

* $p < .10$.
Notes: M=Mean, SD=Standard Deviation, n=# of observations

Table C1.2: Results of Fisher's exact test for '*OfferSize*' by '*GenderProposer*'

	Gender Proposer		
OfferSize	0	1	Total
0-	6	2	8
6-	5	3	8
8-	1	5	6
10-	6	7	13
Total	18	17	35

Fisher's exact = 0.168

C2 Amounts sent to men and women

Table C2.1: Results of two sample t-test and descriptive statistics for '*Eurossent*' by '*GenderResponder*'

	Gender Responder						95% CI for Mean Difference			
	Male			Female						
	M	SD	n	M	SD	n		t	p-value	df
Euro's sent:										
By men	6.39	3.28	9	6.61	3.74	9	-3.736, 3.291	-0.13	0.4475	16
By women	8.61	1.50	9	7.75	2.60	8	-1.301, 3.024	0.85	0.2047	15
Total	7.5	2.72	18	7.15	3.21	17	-1.690, 2.396	0.35	0.7274	33

Notes: M=Mean, SD=Standard Deviation, n=# of observations

Table C2.2: Results of two sample t-test and descriptive statistics for '*Eurossent*' by '*oppositesex*'

	Opposite sex						95% CI for Mean Difference			
	No			Yes						
	M	SD	n	M	SD	n		t	p-value	df
Euro's sent	7.03	2.97	17	7.61	2.95	18	-2.618, 1.455	-0.58	0.2825	33

Notes: M=Mean, SD=Standard Deviation, n=# of observations

Table C2.3: Results of Fisher's exact test for '*OfferSize*' by '*oppositesex*'

	Oppositesex		
OfferSize	0	1	Total
0-	4	4	8
6-	5	3	8
8-	3	3	6
10-	5	8	13
Total	17	18	35

Fisher's exact = 0.828

Table C2.4: Results of Mann-Whitney U test for '*Eurossent*' by '*oppositesex*'

Oppositesex	Obs.	Rank sum	Expected	z	Prob>|z|
0	17	284.5	306		
1	18	345.5	324	-0.730	0.4655

C3 Amounts returned to men and women

Table C3.1: Results of two sample t-test and descriptive statistics for '*Eurosreturned*' by '*GenderProposer*'

	Gender Proposer						95% CI for Mean Difference			
	Male			Female						
	M	SD	n	M	SD	n		t	p-value	df
Euro's returned										
By men	8.00	5.55	9	10.44	3.88	9	-7.229, 2.340	-1.08	0.2948	16
By women	8.11	5.95	9	10.00	4.99	8	-7.605, 3.827	-0.70	0.4920	15
Total	8.06	5.58	18	10.24	4.30	17	-5.620, 1.260	-1.29	0.2062	33

Notes: M=Mean, SD=Standard Deviation, n=# of observations

Table C3.2: Results of Fisher's exact test for '*ReturnedSize*' by '*GenderProposer*'

	GenderProposer		
ReturnedSize	0	1	Total
0-	4	2	6
5-	6	3	9
9-	3	7	10
15-	5	5	10
Total	18	17	35

Fisher's exact = 0.394

Table C3.3: Results of Mann-Whitney U test for '*Eurosreturned*' by '*GenderProposer*'

GenderProposer	Obs.	Rank sum	Expected	z	Prob>|z|
0	18	288	324		
1	17	342	306	-1.204	0.2285

C4 Percentages returned to men and women

Table C4.1: Results of two sample t-test and descriptive statistics for '*PercentageReturned*' by '*GenderProposer*'

	Gender Proposer						95% CI for Mean Difference			
	Male			Female						
	M	SD	n	M	SD	n		t	p-value	df
% returned										
By men	.3814	.132	8	.3981	.114	9	-0.144, 0.110	-0.28	0.7824	15
By women	.3708	.127	8	.3957	.130	8	-0.162, 0.113	-0.39	0.7045	14
Total	.3761	.125	16	.3970	.118	17	-0.107, 0.065	-0.49	0.6251	31

Notes: M=Mean, SD=Standard Deviation, n=# of observations

Table C4.2: Results of Fisher's exact test for '*PercentageGroup*' by '*GenderProposer*'

	GenderProposer		
PercentageGroup	0	1	Total
11.11%-	3	2	5
33.33%-	5	6	11
38.89%-	3	3	6
50%-	5	6	11
Total	16	17	33

Fisher's exact = 1.000

Table C4.3: Results of Mann-Whitney U test for '*PercentageReturned*' by '*GenderProposer*'

GenderProposer	Obs.	Rank sum	Expected	z	Prob>|z|
0	16	254.5	272		
1	17	306.5	289	-0.647	0.5175