

Sustainability - is it contagious?

Fairness in an intergenerational three-person dictator game with social interaction

Inaugural-Dissertation

zur Erlangung des akademischen Grades eines Doktors
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Diplom-Sozialökonom Gunter Bahr
aus Kiel

Kiel, 2009

Gedruckt mit Genehmigung der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Dekan:

Prof. Dr. Thomas Lux

Erstberichterstattender:

Prof. Dr. Till Requate

Zweitberichterstattender:

Prof. Dr. Roman Liesenfeld

Tag der Abgabe der Arbeit:

19. Mai 2009

Tag der mündlichen Prüfung:

18. September 2009

Acknowledgements

First and foremost I would like to thank Till Requate. Without his support this thesis would not have been written. I would like to thank Wilfried Rickels for proofreading and very helpful comments. In addition, I am grateful to Roman Liesenfeld, Eva Camacho, Helmut Herwartz, Joachim Weimann, Reinhard Tietz, and participants at conferences in Graz, Zürich, and Rome for helpful discussions and hints. I would also like to thank all the assistants who supported me in conducting the experiments.

Contents

1	Introduction: The predicament of externalities between different generations	1
2	The evolution of the dictator game - an odyssey from the very beginnings to state-of-art adaptations	6
2.1	The ultimatum game - father of the dictator game	6
2.1.1	The ultimatum game: From the first steps to the standard setups	7
2.1.2	The influence of culture on allocation behavior	8
2.1.3	The relation between neurological responses and allocation behavior	10
2.1.4	The genetic influence on allocation behavior	11
2.1.5	An ultimatum game with two recipients where one recipient has no influence on the allocation	12
2.2	The dictator game - fundament of this experimental study . . .	13
2.2.1	The starting point of the dictator game	13
2.2.2	The first ‘real’ dictator game	14
2.2.3	Paying subjects randomly	15
2.2.4	Increasing the degree of anonymity: The double blind procedure	15
2.2.5	Introducing allocation menus	19
2.2.6	Getting to know each other	20
2.2.7	Is there really a recipient?	21
2.2.8	Varying the degree of social distance	22
2.2.9	Varying the degree of social distance within a known network structure	23
2.2.10	Manipulating the degree of privacy or the instructions . .	25
2.2.11	Varying the ‘perceived deservingness’ of the recipient . .	26
2.2.12	Using workers as subjects	29
2.2.13	Deriving preferences from experimental results	29

2.2.14	Experiments that have some similarities with our setup	34
2.2.15	Hiding behind a random process	35
2.2.16	Altruism in dictator games - an experimental artefact?	37
2.3	The impact of gender in dictator games	38
2.4	The impact of study subject on the propensity to be altruistic	40
2.5	The role of intentions in sequential allocation behavior	41
2.5.1	Intentional action triggers reciprocity in a common pool resource game	41
2.5.2	Intentional action triggers reciprocity in an ultimatum game	42
2.5.3	Mainly negative intentions trigger reciprocity	43
2.6	Common pool resource experiments	45
2.6.1	An intergenerational approach	45
2.6.2	Intergenerational generosity: Triggered by the desire to give or to achieve equity?	46
2.6.3	Investigating intergenerational altruism by eliciting expectations	48
3	Statistical methodology	50
3.1	Mann-Whitney test	50
3.2	Kolmogorov-Smirnov test	51
3.3	Cramér-von Mises test	51
3.4	Anderson-Darling test	52
3.5	Epps-Singleton test	53
3.6	Justification for the choice of tests for the data analysis	54
4	Experimental design	56
5	Analysis of the first generation's allocation choice	60
5.1	Treatment comparison entire sample	60
5.1.1	Descriptive analysis	60
5.1.2	Comparison of distributions	63
5.1.3	Comparison of proportions passing zero	69
5.2	Treatment comparison, including only <i>As</i> passing on a positive amount	71
5.3	Regression analysis of <i>A</i> 's choice	75
5.3.1	Presentation of the different income adjustment methods	75
5.3.2	Presentation of the categories for study subject	79
5.3.3	Regression analysis - results	79

6	Analysis of the second generation's allocation choice	83
6.1	Treatment comparison entire sample	83
6.1.1	Descriptive analysis	83
6.1.2	Comparison of distributions	84
6.1.3	Comparison of proportions passing zero	89
6.2	Treatment comparison, including only B s passing a positive amount	90
6.3	Regression analysis of B 's choice	91
7	Analysis of subjects' expectations	95
8	Analysis of subjects' motivations for their allocation choice	97
8.1	The first generation's explanations	97
8.2	The second generation's explanations	100
8.3	A note on the validity of subjects' answers	102
9	Suggestions for future experimental studies	104
10	Conclusion: Intergenerational solidarity - it is feasible!	106
11	Appendix	111
11.1	Instructions	111
11.2	Program codes	118
11.2.1	Anderson-Darling test - my Matlab code	118
11.2.2	Anderson-Darling test - GAUSS code by Martin Sefton	120
11.2.3	Epps-Singleton test - my Matlab code	120
11.3	Overview experiments	122
11.3.1	Ultimatum games	123
11.3.2	Dictator games	124
11.4	Some treatment comparison results for B	132
11.5	Results for the subsample of B passing on a positive amount	134
11.6	Presentation of the categories for the qualitative analysis to explain allocation behavior	138
11.6.1	A 's explanations	138
11.6.2	B 's explanations	142
11.6.3	A 's explanations in TDG	143
	Bibliography	146

Chapter 1

Introduction: The predicament of externalities between different generations

The Brundtland Report (1987) defines sustainable development as ‘meeting the needs of the present generation without compromising the ability of future generations to meet their needs’. Hence, sustainable development is an attempt to solve the dynamic problem of intergenerational fairness. Issues that directly concern this intergenerational fairness are environmental and resource problems. These problems have a particular feature in common: The present generation faces an intertemporal negative externality. That means e.g. that this generation, when it is only concerned about its own welfare, has an incentive to exploit resources in a way that there is not much left for the future generations or that this generation does not care about pollution stocks where the impact manifests when this generation is not present any longer. A prominent example is the greenhouse gas effect. Economists have been quite successful in developing instruments to deal with externalities. By introducing standards, a Pigouvian tax, or by issuing tradable permits governments cope with these external effects by internalizing them. If the implementor has sufficient knowledge about costs and benefits, an optimal solution is feasible. The problem with intergenerational externalities is the dynamic aspect of the externality: The benefits of the detrimental behavior accrue to the present generation while the future generations have to bear the burden of the costs. Given that people’s preferences are represented by the classical rational agent who cares only about his consumption possibilities, there is no one in the current generation who has to pay the burden. Hence, at the moment there is no one who is interested in the introduction of instruments that could control

a negative externality. Thus, if sustainable development should be induced, other motivations but financial ones e.g. intergenerational altruism have to be addressed.

In order to assess the existence and the possibility of using these additional, non-monetary motivations, experimental economics can provide valuable insights. Experimental evidence from the last two decades has challenged the classical economic paradigm of economic agents to be solely motivated by monetary incentives in decision making.

In various situations people frequently exhibit behavior suggesting that non-monetary considerations must have influenced their decisions. The dictator game is one of the most famous experiments showing that non-pecuniary motivations are inherent in human nature. In that game, one subject, called the dictator or allocator, is free to divide a certain amount of money between the allocator and a second person, called the recipient, whose role is completely passive. No matter how the allocator decides, the division is solely made according to the allocator's decision. Although game theory predicts that in a subgame perfect equilibrium rational allocators will take the whole pie for themselves - thus leaving nothing for the recipient -, such an outcome is not necessarily observable. Depending on the experimental set-up, it may even be the exception rather than the rule.

Forsythe *et al.* (1994, hereafter, FHSS) were the first to conduct systematic experiments on the dictatorial division of an amount of money. In their seminal study they find that, on average, recipients obtain 21% of the pie. Since the dictator game does not give rise to any strategic considerations, something else must influence the allocator's decision, for example the desire to be fair to the recipient. After FHSS many studies, summarized below, tried to shed light on this puzzle. Hoffman *et al.* (1994, 1996, hereafter, HMSS) control for subject-experimenter anonymity, while Bohnet and Frey (1999), Charness and Gneezy (2003), and Burnham (2003) use different degrees of social distance. Goeree *et al.* (2007) and Leider *et al.* (2007) use a network structure to elicit subjects' allocation behavior. Andreoni and Bernheim (2007) introduce a random element in the procedure.

To explain the deviation from the subgame perfect equilibrium in these experiments, several authors such as Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Andreoni and Miller (2002), and Charness and Rabin (2002) employ different models of social preferences. The most prominent explanations are altruism, reciprocity, and inequality aversion. By contrast, Bardsley (2005) and List (2007) interpret all findings where subjects care about somebody else's wealth as pure experimental artefacts.

In this paper, I draw on altruistic or reciprocal behavior models and ask whether friendly behavior can be bequeathed or transmitted to others. In other words, I investigate whether, if a person B has been treated nicely by a person A , then B will ‘reciprocate’ with such friendly behavior to a third person C . For this purpose, I enhance the traditional dictator game by a third ‘generation’, giving the set-up a sequential character. I therefore refer to this game as the *intergenerational three-person dictator game*. The name *sequential dictator game* is already occupied by Cason and Mui (1998), who had two dictator games played one after another but with new endowments on each occasion. In particular, I investigate the influence of social interaction, the role of reciprocity in the intergenerational set-up, and introducing uncertainty if A was a human subject or a random process.

To be more precise, I set up four different treatments with three subjects each, A , B , and C . In each treatment, subject A is asked to divide a certain amount of money Π into an amount $\alpha\Pi$ that she can keep for herself and an amount $(1 - \alpha)\Pi$ to be passed on to subject B . Subject B in her turn is asked to divide $(1 - \alpha)\Pi$ into an amount $\beta(1 - \alpha)\Pi$ that she can keep for herself and an amount $(1 - \beta)(1 - \alpha)\Pi$ to be passed on to subject C . In all treatments I ensure (almost total) anonymity between subject A and subject C .

My baseline treatment 1 is just a sequence of the usual two person dictator game. In treatments 2 and 3, I introduce some kind of social interaction both between A and B and between B and C , thus contributing to the studies where social distance is varied. The social interaction consists of a simple cooperative form of a scrabble game, where additional money can be earned. In treatment 2, the social interaction takes place after A or B respectively have divided the pie, while in treatment 3 the order is reversed. Here, social interaction takes places first between A and B (or B and C), and then A (or B) makes the decision how to divide the pie. The final treatment 4 is similar to treatment 1, but I introduce a random element that has some parallels to the one proposed by Andreoni and Bernheim (2007). In this treatment there is no social interaction. The toss of a coin decides whether subjects A or a random mechanism divides the pie.¹

My main findings are, first, that even in the normal intergenerational dictator game (treatment 1), subject A passes on an amount to B and C that is significantly higher than in the two-person dictator game. Second, I observe more generosity on the part of players A and B in the two treatments with social interaction. Third, I do not find any significant differences between be-

¹I am grateful for John List’s ideas on the set-up of treatment 4.

havior in treatments 2 and 3, although the motivation for being more generous than in treatment 1 may be different in both treatments. In treatment 2, the reason may be that subjects A (or B) want to avoid a situation where they have to spend some embarrassing time with subjects B (or C). In treatment 3, by contrast, the driving force for greater generosity may be increased affection towards B (or C) induced by the previous social interaction. Comparing treatments 1 and 4, I do not find any significant differences concerning A 's behavior. But, to my surprise, there was a highly significant difference between the allocation behavior of the subjects B in treatment 1 and 4.

In addition, I controlled for several socioeconomic data pertaining to the participants. For the first generation (A), I find that income is positively related to generosity. Gender also matters, since females are less parsimonious than males. For the second generation (subjects B), the dominant factor determining their allocation choice is the first generations' choice and the introduction of the random element, while other factors are not significant.

To the best of my knowledge, there has so far been only one set-up that is remotely similar to this experiment, but this set-up lacks the sequential character. Karni *et al.* (2001) study the behavior of a dictator who assigns probabilities of winning a lottery to two recipients. My set-up can also be seen in line with the literature on common-pool resources, where other-regarding behavior in intergenerational set-ups is investigated. The studies closest to my are the ones by Chermak and Krause (2002), Sadrieh (2003), and Fischer *et al.* (2004). The problem with regard to these approaches is that they confront subjects with complicated strategic considerations that make it difficult to interpret their decisions in terms of other-regarding behavior. Subjects usually are faced with a growing resource that can be extracted. Overall welfare is maximized if extraction takes place in the last period possible. But there is always an individual incentive to preempt the others in extracting. This strategic trade-off makes it difficult to assess the optimal behavior for altruistic subjects and also to evaluate if subjects try to be altruistic. My approach, by contrast, is so simple that the subjects' altruistic intentions are obvious. Hence, the intergenerational three person dictator game can be perceived as an ideal type in the sense of Max Weber² in order to capture problems with intertemporal externalities like environmental and resource problems.

The thesis is organized as follows. In the second chapter I survey the present

²Max Weber developed so-called ideal-types that are simplified models of the reality where redundant aspects are neglected while only focussing on important features. These ideal types are no one-to-one translation from reality but still have informative value, think e.g. of a map or of most economic models.

literature that is relevant for the intergenerational dictator game. Therefore, I briefly sketch the father of the dictator game, the ultimatum game, and some experimental studies of the ultimatum game. Then, I summarize all relevant studies of the dictator game. In this summary, I will focus on procedural aspects and results. Experimental aspects like show-up fee, amount to be allocated, numbers of observations, applied statistical methods and some of the results can be seen in an overview in the appendix (see section 11.3.2). The reason for including these information is to show which setup-features are standard and which are not and to later contrast these with my experiment. In the third chapter I present the methodology that is mainly used to analyze dictator games and I explain why I chose particular methods. In the fourth chapter I present the experimental setup of my study in detail. In chapter five, I begin to investigate the results from the intergenerational dictator game. I start by analyzing the allocation behavior of subjects A . First, I use non-parametric methods to compare the different treatments. Then I continue with a regression analysis where I explain allocation behavior using personal data, in addition. In chapter six, I proceed with the allocation behavior of subjects B in the same way done for the allocation behavior by A as in chapter five. In chapter seven I examine the expectations of subjects B and C with regard to allocation behavior. In chapter eight, I qualitatively analyze the justification of subjects A and B for their allocation behavior. In chapter nine, I mention the lessons I have learned from doing experimental studies with respect to the setup and procedure. In chapter ten, I will summarize the experimental studies from the literature review and I will relate this summary to the experimental study of this thesis. Finally, I will recapitulate the results from the analysis of the intergenerational dictator game and conclude by reviewing and extending these results.

Chapter 2

The evolution of the dictator game - an odyssey from the very beginnings to state-of-art adaptations

2.1 The ultimatum game - father of the dictator game

The introduction of the ultimatum game has been a cornerstone for the evolution of experimental economics. Ståhl (1972) developed this experiment theoretically in a work about bargaining behavior and ever since this experimental concept has been widely applied in various experimental setups and modifications. The dictator game, on which this experimental study is built, can be seen in the line of the offsprings of the ultimatum game. I will present the dictator game in section 2.2. The original ultimatum game matches two people, a ‘proposer’ and a ‘respondent’. The proposer is in charge of dividing a certain amount of money between him and the respondent while the respondent accepts or rejects his offer. If the offer is accepted the money is divided according to the proposed distribution. If the offer is rejected both receive nothing. Under the classic assumption of a subject motivated solely by monetary payoff the ultimatum game has a unique subgame-perfect equilibrium. Using backward induction we obtain the result that the proposer should allocate the minimum amount possible to the respondent. The respondent should accept this amount since the respondent is better off taking the money than receiving nothing.

Unless mentioned otherwise, the experimental studies on both the ultimatum and the dictator game, summarized below, share a couple of common design features: Subjects are taken from a student pool. The experimental setup is complete information to all subjects. The position of being proposer/allocator and being respondent/recipient is assigned randomly. ‘Between-subject anonymity’ is ensured which means that subjects do not get to know personal information from each other during the experiment. Usually, subjects are paid for all decisions they make and payment is done in private so that only the subjects themselves and the experimenter know about the subjects’ allocations. In some experiments, only some subjects are paid randomly, I will only mention this specifically if it is of importance.

2.1.1 The ultimatum game: From the first steps to the standard setups

Güth *et al.* (1982) were the first to challenge the concept of the rational agent by actually conducting an experimental study on the ultimatum game. In their experiment, between-subject anonymity seemed not to be ensured since subjects were recruited from a seminar. Surprisingly, Güth *et al.* found that a considerable share of proposers offered much more than the amount of money which should be offered according to game theoretic predictions. The average share offered was 0.35, a share of 0.33 of the proposers chose a fair split. A share of 0.1 of the respondents refused the allocation by the proposer (in the first case the proposer had taken all, in the second case the proposer took 0.8). Güth *et al.* conclude, that the classic economic bargaining theory is of little help understanding ultimatum game results.

There had been numerous replications and modifications of the ultimatum game by Güth *et al.* (1982). As an example, I present the study by Forsythe *et al.* (1994), from here on FHSS. FHSS try to elicit subjects preferences by comparing the results of ultimatum and dictator games (see section 2.2.2 for the dictator game results). Proposers were gathered in one room and respondents in another room and both did not meet during the whole experiment. Communication’ was done via forms carried from one room to another by the experimenter, thus in-between’ subjects anonymity was ensured. The allocation of the pie was done on paper and the money was paid later, seemingly in private. None of the proposers offered to pass no money, most offers ranged around the equal split. The mean share offered was 0.46. A share of 0.09 of the proposers offered a share of 0.20. Half of these offers were rejected. Even an offer of 0.4 of the share was rejected. Confirming the results by Güth *et al.*

(1982), FHSS conclude that subjects do not behave according to theory since they depart from choosing the subgame perfect Nash equilibrium.

Bolton and Zwick (1995) developed the so-called cardinal ultimatum game which allows for a higher degree of control and enables the researcher to test more specific hypotheses. In the cardinal ultimatum game, the proposer has to choose between given distributions of a pie. Then the respondent either accepts the chosen distribution or rejects it. If the respondent accepts it, both will receive their share of the pie according to the given distribution. If the respondent rejects the distribution, both will be left without any payoff. In each cardinal ultimatum game, the subject had to decide between a equal split of the pie of 4\$ and an unequal split. The unequal splits were according to Figure 2.1. Then, the respondents had to decide whether they accepted or not.

Table 2.1: Allocation menus in the dictator game by Bolton and Zwick (1995)

Trial	Allocator's amount	Recipient's amount
1	2.2	1.8
2	2.6	1.4
3	3.0	1.0
4	3.4	0.6
5	4.0	0
6	2.2	1.8
7	2.6	1.4
8	3.0	1.0
9	3.4	0.6
10	4.0	0

Source: Own design

Bolton and Zwick found that the reduced degrees of freedom with respect to the allocation choice did not alter the results compared to the 'usual' ultimatum game results (see e.g. FHSS).

2.1.2 The influence of culture on allocation behavior

Roth *et al.* (1991) conducted a series of ultimatum games with samples from different industrial countries (USA, Israel, Slovenia and Japan) in order to see if cultural differences may induce differences in allocation behavior. Mean offer for American and Japanese subjects was 0.45, while mean offer for Israeli subjects was 0.36.¹ There were significant differences between the distribution of offers for all countries. The only exception was that American and Slove-

¹The mean values are taken from the article by Henrich (2000) where Slovenia is not mentioned.

nian offers did not significantly differ. Rejection rate was 0.19 for American subjects, 0.23 for Slovenian subjects, 0.14 for Japanese subjects, and 0.13 for Israeli subjects. Where offers were low (Japan and Israel), no significantly higher rates of rejection compared to other countries could be observed. Roth *et al.* conclude that there must be different expectations on which offer is adequate triggered by cultural differences.

Henrich (2000) continued to investigate the influence of the cultural background on allocation behavior. He compared allocation behavior by comparing the ultimatum game results from the USA to ultimatum game results from the Machiguenga who live in the tropical forest and subsist on hunting, fishing and gathering. In the Machiguenga society there is little social hierarchy or political complexity, cooperation above family level is almost unknown and most Machiguenga live in 300 person communities. In the USA, only subjects from the department of Anthropology could participate to mimic the community' of the Machiguenga. Pie-size in Los Angeles was equivalent to the pie size for the Machiguenga in that respect that both would have had to work roughly the same time to earn this amount of money. Mean offer was 0.26 for the Machiguenga and 0.48 for students from Los Angeles. The difference between both samples was significant. The rejection frequency was very low among the Machiguenga with a share of 0.05. But also in Los Angeles the rejection frequency was low with a share of zero. A qualitative analysis revealed that the Machiguenga would have accepted any amount. They did not feel unfairly treated by the proposer but rather they simply felt that they had had bad luck being respondent. It became clear that the Machiguenga anticipated that the respondent would be unlikely to reject the offer. The L.A. students, however, said that they assumed that most subjects would have accepted unfair offers while only few subjects might reject an unfair offer. But in order to be sure to gain the 80\$ they split fairly. Henrich concludes that obviously, there are cultural differences. He states that the Machiguenga proposers seem to have no feeling of an obligation to split equally and that Machiguenga respondents do not expect them to do so. Subjects seem to acquire behavioral rules via social learning, hence not all human beings seem to share the same economic preferences.

Henrich *et al.* (2001) continued the study with different other societies including three foraging societies, six that practice slash-and-burn horticulture, four nomadic herding groups and three sedentary, small-scale agricultural societies. Mean offers in the sample ranged from 0.26 to 0.58. Also the rejection rates varied considerably but since these could be contingent on offers, I neglect to list them in detail. Henrich *et al.* explained the allocation behavior by

two dimensions: The payoff, a group receives from cooperation in their daily life and the extent to how much the group relies on market exchange in their daily life. He constructed two indices and uses these indices as independent variables in a regression where the offers were the dependent variable. Both regressors were positive and had a substantial impact. Also R^2 was considerably high with 0.68. It is worth mentioning with regard to the rejection of offers that in one society even hyperfair offers (offering a share above 0.50) were sometimes rejected. This society was a gift exchange society where unsolicited gifts were expected to be reciprocated in the future and rendered one in a subordinant position. Henrich *et al.* state, that there is a close link between bargaining behavior and subjects' daily life. They conclude that long-run evolutionary processes both with regard to the distribution of genes as well as cultural practices could predispose subjects' behavior in certain situations.

2.1.3 The relation between neurological responses and allocation behavior

Sanfey *et al.* (2003) conducted an ultimatum study where subjects' brains were scanned during their decision in order to see if neurological patterns can explain allocation behavior. Subjects were only participating as respondent. Subjects were placed in a MRI² and played the ultimatum game via a computer screen. Each respondent was played ten rounds of the ultimatum game with a human proposer, ten rounds of the ultimatum game with a computer as proposer and ten rounds where subjects received money for pressing a button as control treatment. Subjects knew whether they faced a human proposer or a computer as proposer. Both human and computer offers (in contrast to what has been said to subjects) were predetermined: Half of the offers were fair, the rest was increasingly unfair.

Participants accepted all fair offers while acceptance rates were declining for unfair offers. Comparing the rejection rate of unfair offers made by the computer to the unfair ones by humans yielded a significant difference. Using the MRI, Sanfey *et al.* found that there were certain brain areas that responded stronger to unfair than to fair offers from human partners. These particular areas are usually associated with negative emotional states. In particular, comparing human to computer offers, or human to control offers respectively, yielded a significantly stronger magnitude in activation. There was also a significant correlation between rejection of an offer and the activation

²A MRI (Magnetic resonance imaging) measures signal changes in the brain caused by a change in neural activity.

of particular areas. Sanfey *et al.* also measure the difference in activation between the region with the negative emotional representation and another region which is associated with cognitive tasks. They assume that the region associated with cognitive tasks is used to ensure the achievement of the goal to earn as much money as possible. Comparing the described difference for rejected and accepted offers yielded a significant difference. The difference between activation in the region for emotions and the one for cognitive tasks is assumed to represent the trade-off of between obtaining a maximum of money and retributing for unfair offers. Sanfey *et al.* conclude that there are emotional factors influencing economic-decision making and that the neurological substrates of these emotional states have predictive power in decision making. This result is also confirmed by Knoch *et al.* (2006) who used magnets to disrupt the neural activity in one area of subjects' brains where emotion impulses are thought to be processed. In this study, subjects accepted offers in an ultimatum game that are normally perceived as intentionally unfair and hence were rejected in comparable studies (see e.g. FHSS).

2.1.4 The genetic influence on allocation behavior

Wallace *et al.* (2007) conducted an ultimatum study with twins to investigate the genetic influence on allocation behavior. They compared the behavior of monozygotic (MZ, with the same genes) to the behavior of dizygotic (DZ, whose genes are imperfectly correlated) twins. Subjects had to state in advance for multiples of shares of 0.1 if they accepted the offer or not. Comparing acceptance thresholds between the group of MZ twins and the group of DZ twins yielded no significant difference. However, Spearman's rank correlation coefficient for acceptance thresholds was 0.39 for pairs of MZ twins and -0.04 for pairs of DZ twins. This difference was significant. Hence, there is a strong claim for a genetic component in ultimatum game behavior, since the allocations of a MZ twin correlated considerably while in case of DZ twins there was almost not correlation.

The authors used a mixed-effects ordered probit model to explain the threshold levels of acceptance of the subjects. As independent variables, they took genes, shared environment (e.g. childhood diet, schooling, parental socialization), and unshared environment (e.g. influences not shared by the co-twins. What these were in specific the authors did not mention). The authors were not explicit on how these regressors were measured, this would be particularly interesting for the genetic factor. They find that 0.42 of the variance of the thresholds of acceptance levels could be explained by genetic factors, while the

influence of the same environment was very modest. The authors claim that genetic factors have a substantial impact on the rejection decision in ultimatum game. This implies a heterogeneity among subjects' preferences which in turn explains that authors from previous studies had considerable difficulties in finding a unified, stable model for preferences.³

Sanfey *et al.* (2003) found that activity in certain areas in the brain is correlated with acceptance levels. Since the structure of the brain in this area is under a strong genetic influence, Wallace *et al.* claim that the importance of genetic factors is also backed up by the study by Sanfey *et al.*.

2.1.5 An ultimatum game with two recipients where one recipient has no influence on the allocation

Güth and van Damme (1998) conducted an experimental study in order to investigate whether subjects propensity to be generous in past ultimatum studies was due to strategic considerations or due to altruism. In their ultimatum game setup, they matched an allocator with two respondents. One of the respondents had the same function as in the traditional two-person ultimatum game while the other respondent, called 'dummy', was not allowed to decline the proposed allocation and was hence completely powerless. There were three treatments. In the first treatment, called 'full information', the respondents with the power to decline received information on the entire allocation. In the second treatment, called 'essential information' the respondents with the power to decline received only information on their share of the pie. In the last treatment, called 'irrelevant information', the respondents with the power to decline received only information on the share of the dummy.

In the treatments 'full' and 'essential information', the proposer demanded an average share of one half while leaving an average share of one third to the powerful respondent. In the 'irrelevant' treatment, the proposer took almost the entire amount possible, leaving both respondents with a very low amount of the initial pie.

Güth and van Damme (1998) claim that when reciprocating is possible, proposers act strategically, making their offers appear fair. However, they state that subjects do not have a strong intrinsic motivation for fairness due to the low amounts proposed to the dummy in all treatments and due to the low amounts proposed to both respondents when retaliating was not possible for the powerful respondents.

³See later in section 2.2.13

2.2 The dictator game - fundament of this experimental study

The dictator game is a very simple form of the ultimatum game reducing the interaction of the game to the offer of the proposer. The respondent has no opportunity to reject the offer. Hence, we have no longer a bargaining game but a mere allocation task. Thus, this experimental setup enables us to abstract from strategic considerations and allows us more precisely to pinpoint the motivations behind the allocation behavior. In the dictator game, I will refer to the proposer as 'allocator' and to the respondent as 'recipient' to account for the passive nature of the second player. Under the classic assumption of a subject motivated solely by monetary payoff there is a unique Nash-equilibrium for the dictator game: The allocator takes all and leaves nothing for the recipient.

2.2.1 The starting point of the dictator game

Kahneman *et al.* (1986) were among the first to conduct dictator experiments. Their aim was to challenge the traditional model of the rational agent by investigating the impact of fairness considerations on subjects' allocation behavior. They carried out two experiments: In 'experiment 1', they asked subjects to choose between two divisions of a pie of 20\$. They could choose between splitting equally or an allocation where the dictator received 18\$ while the recipient obtained 2\$. A share of 0.76 of the subjects preferred to allocate an equal split. In 'experiment 2', subjects could either choose to divide 10\$ evenly with an allocator which had been 'fair' in experiment 1' or divide 12\$ evenly with a formerly 'unfair' dictator from 'experiment 1'. A share of 0.74 of the students chose to split with the 'fair' allocator although they had to sacrifice 1\$ for it. Separating the subjects of 'experiment 2' into the ones which had been 'fair' in 'experiment 1' and the others which had been 'unfair' in 'experiment 1' reveals, that there is a strong relation between the results of the two experiments. Considering the subjects that had been 'fair' in 'experiment 1' reveals that a share of 0.88 of these subjects would like to reward the other subjects that had been 'fair' in 'experiment 1'. While considering the subjects that had been 'unfair' in 'experiment 1' reveals that only a share of 0.31 of the 'unfair' allocators would like to reward the subjects that had been 'fair' in 'experiment 1'. Kahneman *et al.* conclude that obviously subjects both care about being treated fairly and treating others fairly.

2.2.2 The first ‘real’ dictator game

Forsythe *et al.* (1994, FHSS) were the first to gather empirical data on the standard dictator game, even though the publication in the year 1994 was preceded by e.g. Sefton (1992). But since they conducted their experiment in 1988, they were the first to run a standard dictator game. In their study they investigated if the generosity observed in ultimatum games stems from strategic or from altruistic concerns. Most experimental features (recruitment, payment,...) were similar to the ultimatum game by FHSS (see section 2.1.1). Surprisingly at that time, FHSS found that the empirical results differed from the ones theoretically predicted, the recipient obtained a mean share of 0.21 of the pie. FHSS claim that incentives for a proposer in the ultimatum game and an allocator in a dictator game could be different. In the ultimatum game, the proposer could pass a share that is larger than the smallest amount possible due to two reasons: Either the proposer is altruistic or anticipates that the respondent might reciprocate on behavior that the respondent deems unfair. Hence, there is a bargaining component involved. In the dictator game, only altruistic motives could drive the allocator’s generosity. FHSS checked if fairness considerations were the only explaining factor in the ultimatum game. If this fairness’ hypothesis was true, there should be no significant differences between the results of the dictator and the ultimatum game. The comparison of the results of the distribution for the money passed on in the dictator and the ultimatum game revealed that the fairness hypothesis could not be sustained. There was a significant difference between both games, hence there were strategic considerations in the ultimatum game. Interestingly, the distribution of the assigned shares for the dictator game was bimodal. That means that there was a ‘cloud’ of observations close to the fair outcome where the allocator passed half of the pie and a ‘cloud’ of observations where the allocator passed zero. FHSS conclude that there are two types of players, one of them solely driven by monetary incentives and another type which must have preferences which do not only incorporate monetary considerations but also different, other motivations.

Finally, FHSS examine whether the size of the pie, which had to be allocated, mattered. So far, pie-size had been 5\$. Now, they conducted additional treatments of the dictator and the ultimatum game with a pie size of 10\$. None of the distributions were significantly changed if the size of the pie was increased, also the fairness hypothesis was still rejected significantly. FHSS conclude that the pie size (at least if it is not extremely small or large) does not matter.

2.2.3 Paying subjects randomly

The study by FHSS was carried out with real payments, however, the authors conducted a series of additional experiments with hypothetical payments. They investigated a question, which they called the ‘pay hypothesis’: They checked whether the results of the experiments were altered if the subjects allocate hypothetical money without real payments. They found that there was a significant difference between the dictator game with real and virtual money. FHSS conclude that using real’ money is mandatory to obtain useful results.

Sefton (1992) conducted a series of dictator games where he tested the hypothesis that it does not matter if you pay all subjects in a dictator game or if you only pay some subjects randomly. Sefton applied the experimental procedure by FHSS. There were two treatments. In the first treatment called ‘full payment’ all allocators were paid. Sefton finds no significant difference between the results by FHSS and his replication and hence he pools the data. In the second treatment called ‘random pay’ everything was left unchanged except that only a share of 0.25 of the subjects was paid randomly. Comparing the distribution of the offers of ‘random payment’ and pooled ‘full payment’ yielded a significant difference between both samples. The mean share of offers in ‘random pay’ was 0.37 while that of the pooled ‘full pay’ sample was 0.22. Comparing the results of the no pay experiments by FHSS (mean share of offers was 0.38) to the results of ‘random pay’ by Sefton, yielded no statistically significant difference. Sefton concludes that random payment could distort dictator game results severely.

2.2.4 Increasing the degree of anonymity: The double blind procedure

Hoffman *et al.* (1994, HMSS), investigated the role of anonymity with regard to subjects’ allocation behavior and tested if a higher degree of anonymity compared to the FHSS-setup leads to results which restore the paradigm of the rational agent. They introduced ‘subject-experimenter’ anonymity in one of their treatments. That means that not only between-subject anonymity (see section 2.1) was guaranteed but that also the experimenter was not able to link allocations to specific subjects The conducted experiments had four different treatments and differed from the experiments by FHSS in several aspects. In the first two treatments, all subjects were gathered in one room. The allocators and recipients stayed in the same room together but it was

ensured that subjects did not get to know who was allocator and recipient and who was matched with whom. Between-subject anonymity was ensured, payment was done individually and privately.

In the first treatment, referred to as ‘random exchange’, HMSS used the same instructions and the same way to instruct the subjects as FHSS but depart in one aspect. FHSS used the term: ‘A sum of 5\$ has been provisionally allocated to each pair’ when describing the pie, the allocator had to divide. HMSS presented the allocator as seller and recipient as buyer. Depending on which price the seller (allocator) chose, seller and buyer (recipient) received certain shares of the pie which can be seen in the payoff chart (see Figure 2.1). Thus, the allocation problem remained completely the same compared to FHSS, only the way of presentation was changed. HMSS explained the

**Seller Chooses
PRICE**

\$0	\$1	\$2	\$3	\$4	\$5	\$6	\$7	\$8	\$9	\$10	
\$0	\$1	\$2	\$3	\$4	\$5	\$6	\$7	\$8	\$9	\$10	Seller profit
\$10	\$9	\$8	\$7	\$6	\$5	\$4	\$3	\$2	\$1	\$0	Buyer profit

Figure 2.1: Payoff chart given to subjects. Source: Hoffman *et.al.* (1994)

introduction of new instructions in ‘random exchange’ with the different impressions which were conveyed. If it was said that a sum was ‘provisionally allocated to a pair’ this could frame the subjects to believe that the pie belongs to both the allocator and the recipient and hence trigger altruism. If you act as buyer and seller, this framing effect should disappear.

The next treatment, referred to as ‘context exchange’, was done in the same way as ‘random exchange’. The only difference was that HMSS assigned the positions as allocator or recipient according to the results of general knowledge quiz. Thus, the subjects results were ordered with respect to success and the upper half became allocator and the lower half recipient. According to HMSS, replacing the randomization process deciding on subjects’ position by an assignment via the knowledge contest should have an impact. If subjects earned the right to decide, they might be inclined to perceive themselves as more deserving as it would have been the case, if they obtained this position by pure chance and hence be less generous.

In the third treatment, referred to as ‘double blind 1’, HMSS ensured

subject-experimenter anonymity. Thus, not only the subjects did not get to know each other but also the experimenter did not get to know which allocation was made by which subject. The position as allocator or recipient was assigned randomly. The allocators were seated in room A and the recipients in room B. Then a monitor was voluntarily selected from the allocators. The monitor took the role of the experimenter and received a payment of 10\$⁴. Now, each of the remaining allocators received an envelope. For most of the allocators (a share of 0.83) the envelope contained 10 blank slips and 10 bills of 1\$. For the other allocators (a share of 0.17), the envelopes contained 20 blank slips of paper. Then one allocator after another entered a large cardboard box in the same room. In this cardboard box, the allocator was invisible for the rest. Then, the allocator chose an allocation of the pie by choosing the number of bills and slips for themselves and for the recipient. The slips and bills for the recipient were put in an envelope, the remaining bills and slips were kept by the allocator. The number of slips and bills had to add up to 10 such that subjects could not infer from the thickness of the envelopes to the amount of money distributed. Then the recipients from room B were one by one called in room A and chose one of the envelopes (without knowing which allocator filled the envelope). The money left for B was recorded.

The last treatment was similar to the ‘double blind 1’ but there was no monitor and the share of envelopes which only contained slips were replaced by envelopes containing also dollar bills. This treatment was referred to as ‘double blind 2’

The idea behind the introduction of the double blind treatment was the following: HMSS claim that subjects (maybe unconsciously) build up expectations about the experimenter’s objective. They may think that their behavior in the current experiment might affect future decisions of the experimenter. Maybe the chance of being chosen for a future experiment or the chance of being chosen for a particular beneficial experiment could be affected by the current behavior. The double blind treatment rules out such considerations. Another reason to introduce the double blind treatment may be that subjects do not like to appear greedy or selfish in the eyes of the experimenter. Handing out some envelopes with no bills ensured that recipients and the experimenter could not know if the allocator had been greedy or if she simply had not had the possibility to pass some money (in case all other allocators had taken the whole amount). Still there were some possible influences left. The fact that the monitor gained an amount equivalent to the whole pie might have induced the

⁴Which is the same amount a dictator receives if the dictator takes all.

subjects to believe that they were supposed to take the entire pie as well. Giving a certain share of envelopes to subjects which only contained slips might evoke the impression that leaving no money was all right. These influences were eliminated in ‘double blind 2’ by omitting the monitor and the envelopes that only contained slips.

Comparing the ‘random exchange’ to FHSS yielded no statistically significant difference, while comparing the ‘contest exchange’ to FHSS lead to a significant difference. HMSS found that the ‘contest exchange’ lead to significantly lower results than the ‘random exchange’. ‘Double blind 1’ was not statistically different from ‘double blind 2’. But there was a significant difference between the double blind treatments and all other treatments. The double blind procedures lead to much lower offers than the other treatments.

HMSS claim that the results from the double blind procedure are a hint that the evidence from other experiments that have so far been interpreted in a way that fairness should be incorporated into the subjects utility function could be a mere artefact. Subjects consider the regard of others of themselves as an important issue but not the particular welfare of the others.⁵ In line of this argument the idea of the selfish agent can be sustained. If there is social monitoring, subjects will incur costs to be ‘nice’ while hoping that they would be treated nicely in return since people usually reciprocate and discriminate against ‘unkindly’ people. Hence, HMSS claim that subjects simply apply the theory of repeated games in the laboratory experiments: Subjects in real life are frequently confronted with repeated game environments where being friendly is useful to build up a reputation as ‘nice guy’ and hence reap the profits of positive reciprocity from this reputation.

In a sequent article Hoffman *et al.* (1999), from here on HMS, added two more treatments to the treatments by HMSS. One treatment, called ‘single blind 1’ was similar to ‘double blind 2’ except that there was no experimenter-subject anonymity any longer but only between-subject anonymity. After allocating the money while being alone in a cardboard box, the subject returned to the experimenter and the allocation was recorded. Then the subject left with his money. Social distance in ‘single blind 1’ was narrowed with respect to the double blind treatment.

The other treatment was called ‘single blind 2’. Here, the social distance was reduced even more. Now, subjects used a decision form instead of real dollar bills. HMS predicted, that there should be a negative relation between social distance and generosity: Subjects will be more generous if there is less

⁵As noted above, the chance of being chosen for a future experiment could be affected by being fair. Or subjects do not like to appear selfish in the eyes of the experimenter.

social interaction between the subjects. Using a non-parametric order test, the hypothesis that the offers were the same across all treatments was significantly rejected. However, pairwise comparisons of the treatment generally did not lead to significant differences. HMS interpret the results as a confirmation for their results from HMSS.

2.2.5 Introducing allocation menus

Bolton and Zwick (1995) developed a modification of the dictator game, which was derived from the cardinal ultimatum game (see section 2.1.1). The allocator had (again) a menu of possible allocations from which he could choose, e.g. choosing between passing nothing or half of the pie. By using this kind of game, Bolton and Zwick tried to enhance the control of the experiment by filtering certain aspects while directly testing certain other aspects. The menu for unequal splits followed the ultimatum game by Bolton and Zwick (1995) (see section 2.1.1, Figure 2.1). Overall, a share of 0.98 of all choices was unequal. Bolton and Zwick conclude that subjects behaved in line with classical economic theory.

Bolton *et al.* (1998) conducted a series of experiments which was based on the experiments by Bolton and Zwick and hence used menus. In each treatment (see Figure 2.2), the overall amount of money to be allocated was 10\$. In the so called ‘1Game6Card’ treatment, the allocator had to make one

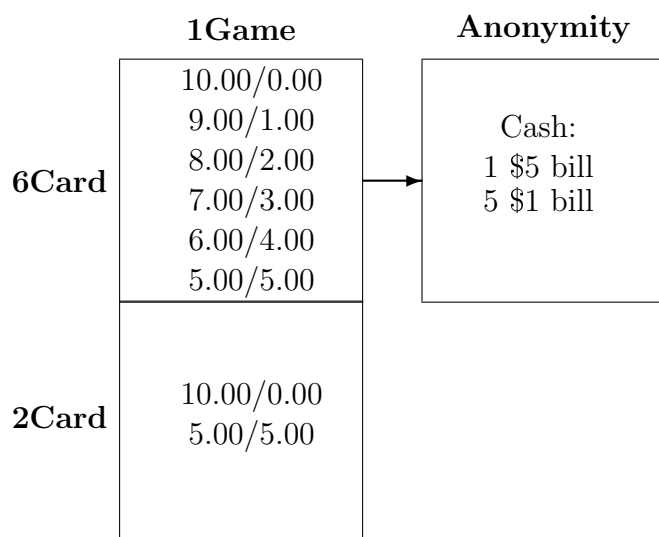


Figure 2.2: Dictator game treatments. Source: Bolton *et al.* (1998), slightly altered compared to original.

decision between different amounts of money which he would keep, starting

with 5\$ up to 10\$ with increments of 1\$. In the so called ‘1Game2Card’ only two choices were provided : Either the allocator took all or she split equally. Then there was another treatment, the anonymity treatment. This treatment was similar to the ‘1Game6Card’ treatment but ensured subject-experimenter anonymity and allocators received the cash directly instead of using paper and being paid, later. The menu in the ‘2Card’ treatment put the allocators in a position where they faced two possibilities when it came to their decisions: Either one of the splits offered in the menu matched their preferences (that means they would have taken all or divided equally) or they would like to choose a split in between these two extreme values which they now could not accomplish. In the latter case, the subjects had to decide which split fitted their preferences best. They were confronted with the decision to either deviate from their original preference to an allocation in their favor or to one in favor of the recipient. Comparing the proportion of subjects leaving 0 in the ‘1Game2Card’ (0.53) to this proportion in the ‘1Game6Card’ (0.93) yielded a significant difference. Bolton *et al.* (1998) conclude that in case the menu does not allow to choose the desired allocation, subjects switch to the allocation which favors them.

Comparing the ‘1Game6Card’ and ‘anonymity’ yielded no significance difference, hence, in this case increased anonymity has no impact on allocation behavior. Bolton *et al.* claim that the effect of being more greedy in the double blind treatments by HMSS compared to their baseline treatment (or compared to FHSS) is not caused by the increased anonymity but by framing via the instructions and the setup. They state that using the phrase ‘provisionally allocated’ or the idea of a market exchange generates a notion of being obliged to pass on money while putting money directly into envelopes conveys a different moral surrounding possibly not evoking this sense of obligation. Bolton *et al.* conclude that the effect of decreased generosity in presence of experimenter-subject anonymity does not show up in their study because they held their game frame constant.

2.2.6 Getting to know each other

Bohnet and Frey (1999) used the experimental study by HMS as point of departure. They investigated if allocators became more generous, given that they get to know the recipient closer. There were four treatments. The first treatment, ‘anonymity’ was designed to be close to ‘double blind 2’ by HMSS. In the second treatment, ‘one-way identification’, allocators could take a glance at their respective recipient, while the recipients did not know their allocators.

In the third treatment, ‘one-way identification with information’, recipients were also asked to tell their name, where they came from, their study major and their hobbies. In the last treatment ‘two way identification’, allocators and recipients both saw each other for a short while. In all treatments, subject-experimenter and between-subject anonymity was guaranteed (for allocators).

The distribution of money passed on in the one-way identification with information and the two way identification were significantly different from the anonymity treatment, while the one-way identification without information was not significantly different from the anonymity treatment. Comparing the proportion of allocators taking the whole pie yields the same results as the comparisons of the distributions of the share passed. In contrast to HMSS, Bohnet and Frey conclude that the social interaction ‘transforms anonymous, faceless entities into visible, specified human beings’ (Bohnet and Frey, 1999, p. 339) and hence triggers altruism.

2.2.7 Is there really a recipient?

Frohlich *et al.* (2001) investigated if increased anonymity as in HMSS could cast doubts among the subjects if there really was a recipient. They ran two treatments. The first treatment, ‘two-rooms’, was similar to the ‘double blind 1’ treatment by HMSS, there was only one thing changed: After allocating the money, the subjects had to fill out a questionnaire about their beliefs concerning the existence of a recipient and if this recipient would receive all the money.⁶ The second treatment, ‘one-room’, was similar to ‘two-rooms’ except that allocators and recipients were together in one room, now (hence eliminating doubts about the actual existence of the recipient).

Mean share passed in ‘two-rooms’ was 0.18, mean share passed in ‘one-room’ was 0.25. Unfortunately, there was no statistical comparison of the distributions of the amount of money passed in the two treatments. Instead Frohlich *et al.* define a binary variable: The first category, ‘selfish’ contained the allocators passing nothing or 1\$. The other category, ‘other-regarding’ contained the allocators who passed 4\$ or more. Note, that pie size was 5\$. The observations, where an allocator passed 2\$ or 3\$, were discarded, leaving overall 41 observations.⁷ Comparing the beliefs (which were measured on a scale) of the ‘selfish’ and the ‘other-regarding’ type with regard to the existence of the recipient yielded a significant difference. The selfish type had much higher

⁶This was implemented in a way such that ‘double blind’ was still preserved

⁷The reason for discarding these observations was possibly to form clear cut categories, since it is difficult to attribute choices of 2\$ or 3\$ to the categories ‘selfish’ and ‘other-regarding’.

doubts about the veracity of the procedure. Frohlich *et al.* also ran a probit regression with a binary dependent variable for the allocator type being either ‘selfish’ or ‘other regarding’. Regressors were a dummy variable for allocators who doubted that the money was passed to a recipient, a dummy variable for allocators who perceived the experiment as game⁸ and the product of both regressors. All regressors were significant.⁹ Frohlich *et al.* conclude that the double blind treatment could cast substantial doubts on the the existence of the recipients hence rendering allocators purely selfish. They note that the one-room design does decrease the degree of anonymity and thus could also increase generosity through this channel.

2.2.8 Varying the degree of social distance

There had been many studies that investigated the impact of varying the social distance in dictator games. All of these studies had a baseline treatment in the spirit of the traditional dictator game, most of these studies used a replication of the ‘double blind 1’ procedure by HMSS.

Johannesson and Persson (2000) used a treatment where recipients were not present but were randomly chosen from the general population in Sweden. These recipients did not participate actively in the experiment and did not know about being chosen. Comparing both mean share passed and proportion of subjects keeping all of this treatments to the baseline treatment yielded no significant difference.

Charness and Gneezy (2003) used a treatment where the allocators knew the name of the recipient. Telling the private name increased the mean passed share by about 50% compared to the baseline treatment, the difference in distributions was significant.

Burnham (2003) studied the impact of showing photos. In one treatment, called ‘dictator photo’, a picture of the allocator was enclosed in the envelope that was passed to the recipient. In another treatment, called ‘recipient photo’, a picture of the recipient was given to the allocators with the money slips. The pictures were taken upon arrival of the subjects to the experiment and shredded after having been viewed. Comparing the distributions of all treatments pairwise yielded no significant differences. Using the proportion of subjects passing on half reveals a significant difference between both treatments where a photo is shown and the baseline treatment. Burnham consequently states that showing photos (some information) makes a difference. He claims that for

⁸This was also an information asked in the questionnaire.

⁹Note, that the authors used one-tailed test statistics which is unusual.

ancestral humans social information was normally accompanied by likely future social interaction which made it reasonable to be generous to induce reciprocal behavior. In addition, he claims that this process is partly unconscious and evolutionary engrained in our genes. He concludes that in our modern world social interaction does not necessarily correlate with future interaction. Nonetheless, social information triggers enhanced generosity because of our genetic heritage.

Brañas (2006) investigated the impact of using friends as subjects. In a treatment called ‘friends’, allocators got to know if the recipient was a friend or someone unknown before they allocated. Comparing the two treatments revealed that allocators were significantly more generous to friends than to strangers.

2.2.9 Varying the degree of social distance within a known network structure

Unlike the studies in the previous subsection, the following studies do not vary social distance by giving e.g. one additional information (a rather ‘discrete’ way) but by changing the social distance within a network structure (in a rather ‘continuous’ way).

Goeree *et al.* (2007) investigate the impact of having different social distances between allocator and recipient. Unlike usually subjects were recruited from an all girls school from 5th and 6th grade. First, a survey was conducted to investigate the network structure with respect to friendship relations among the girls and further data on personal characteristics were collected. Then each subject had to play ten dictator games, where the name of the recipient was given. In three dictator games subjects had to allocate to friends, in three dictator games they had to allocate to friends of friends and in the remaining games they had to allocate to others. Regressing the share passed on social distance while also including some social features (race, height, shyness,...) revealed that social distance was highly significant. The only other regressor that was significant was the shyness of the allocator. Subjects with an infinite distance (strangers) would receive 0.17 of the pie, 2nd degree friends an additional 0.2 and 1st degree friends an additional 0.16. Shyness reduced the share passed by 0.03. The authors conclude that for adolescents social distance is an important factor influencing the allocation choices.

Leider *et al.* (2007) conducted a study in the spirit of Goeree *et al.* (2007), where they also tested allocation behavior with different social distances in a network. The experiments were done via internet. There was a prelim-

inary experiment in order to elicit the social network of the participants. Then, two experiments with dictator games followed. The first experiment was called ‘nameless’. Here, the recipient was anonymous but from the same dormitory. Decisions should be made for two treatments: In the ‘anonymous’ treatment, the allocators’ and recipients’ names remained confidential. In the ‘non-anonymous’ treatment, at the end of the experiment, both players were informed about each other’s identity via e-mail. Analogously to Andreoni and Miller 2002 (see later section 2.2.13), the dictator game was done with different hold and pass values of tokens (the tokens were converted into real money after the experiment). To explain the concept of hold and pass values I use the following examples: Consider e.g. a hold value of one for a token and a pass value of three. That would mean that transferring one token raises the recipient’s payoff by three tokens. Hence, the price of giving is one third. If e.g. the hold value is one and the pass value is one too, the price of giving is one and we have the usual dictator game. If e.g. the hold value is three and the pass value is one, the allocator has to sacrifice three tokens in order to pass one token to the recipient, hence the price of giving is three. In the experiment by Leider *et al.* there were three decisions. In the first decision, giving was ‘efficient’: The hold value of a token was one while the pass value was three. In the second decision, hold and pass value were both two, hence giving was ‘neutral’ and in the third decision, hold and pass value from the first decision were reversed, hence giving was ‘inefficient’.

Some days after the first experiment, the second experiment was conducted, called ‘social distance’. Now, procedure was the same as in the ‘nameless’ experiment (hence also both treatments ‘anonymous’ and ‘non-anonymous’ were conducted) except that the allocator played the games five times with five different recipients where social distance was widened in the following way: Direct friend - friend of a friend - friend of a friend of a friend - someone from the same floor - someone from the dormitory who was not in the former categories.

For both experiments and for any social distance (in the second experiment) the share given was significantly higher in the ‘non-anonymity’ treatment than in the ‘anonymity’ treatment. In ‘social distance’, friends received significantly higher shares than subjects at lower social distances in all treatments. Leider *et al.* (2007) used a Tobit regression with the amount given in the ‘social distance’ experiment in the ‘anonymity’ treatment as dependent variable. They used different models including different regressors, e.g. the social distance, amount given in the ‘nameless’ experiment, gender and other factors. They found that subjects who tended to give more in ‘nameless’ experiment also gave more in

the ‘social distance’ experiment. Gender, both for allocator and recipient was not significant. Also, allocations to friends were higher than to subjects with larger social distance and were decreasing with social distance, though mostly not significantly.

In a second Tobit regression, Leider *et al.* used the amount given in the ‘social distance’ experiment in the ‘non-anonymity’ treatment as dependent variable. Now, in addition to the regressors from the first regression they included a dummy for the ‘anonymity’ treatment. They found that there was a significant positive effect of being non-anonymous. If giving was efficient¹⁰ being matched with a friend had a significant positive impact, this effect also showed up for ‘neutral’ while it vanished in ‘inefficient’. Leider *et al.* conclude that directed altruism (giving to a friend) makes subjects more generous.

2.2.10 Manipulating the degree of privacy or the instructions

Haley *et al.* (2005) tested if subjects react to cues not directly related to the experiment. The first treatment was called ‘baseline’. This treatment was a usual dictator game but conducted using a computer. The second treatment called ‘eyespot’ was similar to ‘baseline’ except that there was a picture of a pair of eyespots in the background of the screen. There was a significant difference between both treatments, both in means and proportions of subjects keeping the all. The authors conclude that generosity could be influenced by many types of inputs and that the degree to which subjects react to these cues is different, too. They claim that the presence of eyes facing in one’s direction triggers the impression of being watched. Altering behavior if you were being watched was evolutionary useful and this behavior seems still to manifest in modern life. Haley *et al.* emphasize the importance of possible cues in experiments that could lead subjects to specific assumptions with regard to the social context in which their behavior takes place.

Brañas and Morales (2005) varied the instructions of their experimental study with suggestive terms to see in how far subjects respond to such manipulations. Among others there were two treatments. The first treatment, called ‘baseline’ was a dictator game, where the allocator divided points for the exam of the course where the subjects were taken from.¹¹ In the second treatment, called ‘helping’, the sentence ‘Note that your recipient relies on

¹⁰A pass value >1

¹¹Actually, I have difficulties believing that researchers put students’ grades at stake in experiments, but maybe they have another view on that matter in Spain.

you’ was added to the instructions. There was a significant difference between both treatments. The authors conclude that subjects can be induced to pay increased attention to helplessness via instructions.

Dufwenberg and Muren (2006) varied the experimental procedure with regard to the way, subjects were paid. The objective was to investigate the impact of making the allocators decision public. Hence, they check if subjects respond in their allocation behavior to the possible reputation effect of their allocating behavior. Actually, only two pairs of all 352 observations were randomly drawn to be payed, one pair for each treatment. Dufwenberg and Muren used two treatments, ‘private payment’, where subjects were paid privately, and ‘open stage’, where subjects were paid on stage in a lecture hall where all participants were invited. Comparing distribution of ‘private payment’ and ‘open stage’ yielded a significant difference. In this case, decreased anonymity, unlike in the other literature, reduces generosity. Dufwenberg and Muren suggest that group or audience effects may have outweighed the impact of the reduced anonymity, here. They claim that subjects that were recruited from a pool of economists¹² may deviate from the choice they would make in private in order to behave in accordance with the expectation of the group. This expectation may be to behave as rational, selfish agent if you were beginning to study economics. Complying to this expectation could be seen as another reward apart from the monetary gain.

2.2.11 Varying the ‘perceived deservingness’ of the recipient

Eckel and Grossman (1998) matched the allocator with a charity organization as recipient. They investigated whether a recipient who could be perceived as particularly deserving like a charity organization would receive a higher share compared to usual recipients. There were two treatments, the first treatment was a replication session of the double blind treatment by HMSS, called ‘replication’. Comparing the results from ‘replication’ to the results by HMSS yielded no significant difference, hence Eckel and Grossman pooled the data. The second treatment, called ‘charity’ was similar to ‘replication’ but instead of matching the dictator with another subject, the recipient was replaced by a charity organization (the American Red Cross). ‘Replication’ (pooled) was significantly different from ‘charity’. Eckel and Grossman conclude that altruism is part of subjects preferences and manifests if the recipient is perceived as deserving even in double blind conditions, hence increased anonymity does

¹²Which was the case in this experiment.

not necessarily extinguish altruism.

Ruffle (1998) conducted a series of dictator games where the size of the pie was fixed endogenously according to the recipients skill. The idea behind this was to test if the proposer changes his attitude towards the recipients deservingness with respect to the outcome of the contest. Anonymity was ensured according to the ‘double blind 1’ treatment by HMSS. There were two treatments. In the first treatment, ‘skill’, the size of the pie is determined by a skill contest. The contest had to be done by the recipients. The recipients were ordered according to the skill-test, the upper half of the candidates of the session gained 10\$ which could be allocated by the dictator while the rest received 4\$ to be allocated by the dictator. The second treatment was called ‘coin’. Here, the pie-size (also 10\$ or 4\$) was determined by tossing a coin in presence of the recipient. Both recipient and allocator learned about the outcome of the pie size. For the subsample of winners (the allocators deciding on 10\$), comparing the distributions of money passed in ‘contest’ to ‘coin’ yielded a significant difference. There was a reward of merit, since the winners of the contest received a share of 0.45 on average compared to the share of 0.34 which the lucky subjects received. For the 4\$ pie, there seemed to be a mild punishment of losers: The losing recipients in ‘skill’ received a mean share of 0.23, while the unlucky recipients in ‘coin’ received a mean share of 0.31. Comparing the distributions of money passed in ‘contest’ to ‘coin’ for a pie-size of 4\$ yielded no significant difference. Ruffle (1998) claims that the proposer faced a trade-off between the monetary benefit he received if he took a high share and the moral distress he had to incur by perceiving himself as unjust (since he chose an inequitable distribution). The allocators understood that being able to distribute a large pie was due to the effort or competence of the recipient. In order to maintain a positive self-image the allocators passed on more than the classical strategic Nash-equilibrium would have suggested. Of course the losing recipients were less deserving from the point of view of the allocators and hence received less than the successful ones. Ruffle’s results contrast with the double blind treatment by HMSS where the offers were much lower on average. Ruffle rejects the conclusion by HMSS who claim that fairness is primarily explained by expectations. If someone appears deserving (and for the skill winners this fact seems to be obvious) then other-regarding behavior is the normal reaction of subjects. Ruffle concludes that ability and effort are rewarded, lack of skill or effort are moderately punished.

The experiment by Cherry *et al.* (2003) was in the spirit of Ruffle’s (1998) study, but here, the pie size was determined according to the allocators results

from a quiz.¹³ The objective was to see, if the allocators change their allocating behavior given that the pie-size was dependent on a previous effort by the allocators. There were three treatments. In the first treatment, ‘earning’, there were two stages. In the first stage, the allocators played a quiz. If allocators answered ten or more questions correctly they earned 40\$, otherwise they received 10\$. In the second stage, the allocator could split the earned money with the recipient, the setup was similar to FHSS. The second treatment, ‘double blind with earnings’, was identical to ‘earnings’ except that in the second step the procedure was according to ‘double blind 1’ by HMSS. The third treatment, ‘baseline’ was similar to the procedure by FHSS and involved either a pie of 10\$ or 40\$ (in order to make this treatment comparable to the two other treatments). Cherry *et al.* found that the proportions of allocators passing on zero and the distribution of offers were significantly higher in ‘baseline’ compared to the ‘earning’ treatments. Comparing ‘earning’ with ‘double blind with earning’ yielded a significant difference for the high stakes (40\$) both for proportion of nonzero offers and for the distribution while for the low stakes there were no significant results. Hence, anonymity only partially increased parsimony. Nevertheless, the striking evidence of the experiment were the results for ‘double blind 1’. A share of 0.95 of the low stake proposers and a share of 0.97 of the high stake proposers kept all of the money they earned and thus stuck to the game-theoretic prediction. The authors conclude that if the money the proposers receive is legitimized via effort and if strategic behavior is ruled out by guaranteeing anonymity, the rational, game-theoretic behavior is the norm. According to Cherry *et al.*, altruism is triggered by strategic concerns (e.g. pleasing the experimenter as suggested by HMS).

List and Cherry (2004) check if an equal pre-allocation endowment of the subjects affects generosity in dictator games and if it matters if subjects earned their pre-endowment. There were three treatments. The first treatment was called ‘symmetric’ and had two stages: The first stage was the earning stage: Here, both allocators and recipients had to answer a quiz. If they were successful, they received 100\$, if not, they received 20\$. In the second stage allocators now decided on how to split the amount of money they earned. The second treatment was called ‘asymmetric’. This treatment was similar to ‘symmetric’ except that recipients skipped the first stage and did not receive any money. The last treatment was called ‘baseline’ and was almost similar to HMSS except that stake size is altered to 20\$ or 100\$ respectively. Comparing the distribution of the samples for the stake size of 20\$ and 100\$ within

¹³Remember, in Ruffle (1998) the pie was assigned according to the recipients results from a quiz.

the treatments (hence, checking if stake size matters) yielded no significant difference. Thus, the data was pooled within the treatments. Comparing the distributions of ‘asymmetric’ and ‘baseline’ yielded no significant difference while comparing ‘symmetric’ to ‘baseline’ yielded a significant difference, also comparing ‘symmetric’ to ‘asymmetric’ yielded a significant difference. List and Cherry conclude that in contrast to the results by HMSS¹⁴, earning entitlements have at most a modest effect. Since subjects propensity to give was reduced a lot in the symmetric treatment, List and Cherry claim that subjects exhibit inequality aversion.

2.2.12 Using workers as subjects

Carpenter *et al.* (2005) ran a dictator game where they control for the social background of the individuals. There was one group (both allocator and recipient) of affluent, young students, one group of older and less affluent students and a group of less affluent workers who were older than students from both groups. Comparing the distribution of offers between groups lead to a significant difference between affluent students and workers and to a significant difference between less affluent students and workers. The difference between students was not significant. Carpenter et al. conclude that altruistic norms have a higher impact on workplace than in classroom.

2.2.13 Deriving preferences from experimental results

The aim of the experiments of the following section was to derive the subjects’ utility function via revealed preferences. Andreoni and Miller (2002) conduct a series of menu-based dictator games with tokens which have different hold and pass values (see section 2.2.9 for an explanation of hold and pass values). The different budgets and hold and pass values can be seen in Table 2.2. Subject-experimenter anonymity was guaranteed.

Andreoni and Miller compared the results of the decisions where the tokens had a pass and hold value of one with the corresponding treatment results of FHSS. They found that the results were strikingly similar, all ranging between a share passed of 0.22 to 0.25. They conclude that their results are representative with respect to the previous experiments.

Andreoni and Miller constructed a utility function that reflects subjects’ preferences. They developed three categories of preferences, the perfectly self-ish type whose utility depends only on her own pay, the inequality aversion

¹⁴Here, earning the position as allocator made a crucial difference.

Table 2.2: The different treatments by Andreoni and Miller (2002)

Treatment	Token endowment	Hold value	Pass value	Average number of tokens passed
1	40	3	1	8.0
2	40	1	3	12.8
3	60	2	1	12.7
4	60	1	2	19.4
5	75	2	1	15.5
6	75	1	2	22.7
7	60	1	1	14.6
8	100	1	1	23.0
9	80	1	1	13.5
10	40	4	1	3.4
11	40	1	4	14.8

Source: Andreoni and Miller (2002)

type with a Leontief utility function with her and the recipient's payoff as arguments and finally the type who maximizes overall utility where the payoffs for the allocator and the recipient are perfect substitutes. Every type was subdivided into either types of the pure form, where subjects' preferences match exactly the pure form of the utility function (a share of 0.43 of all subjects), and types of the weak form, where the utility function is relaxed by some parameters to better match the preferences (a share of 0.57 of all subjects). The parameters of the utility functions for the weak forms were determined for every subcategory using a Tobit estimation (since the subjects choices were limited at both ends). A CES function had the best fit with the data and hence was chosen. For every weak type the parameters selfishness and convexity of preferences were significant. Andreoni and Miller conclude that altruism can be rational if you apply the correct preferences. Using different quasi-concave utility functions for individuals reconciles the theory of the rational agent with empirical observations. Consequently, the paradigm of the rational agent is sustained but under different assumptions. The often cited altruistic components in individuals utility functions have to be taken into account but only for parts of the population.

Charness and Rabin (2002) used an approach which was quite similar to Andreoni and Miller (2002). First, they conduct a series of menu-based dictator games (see Table 2.3) where allocators had to choose between two allocations.

They found, that given that the payoff of the allocator is the same for both allocations, more than $\frac{2}{3}$ of the allocators chose allocations where the

Table 2.3: Two-person dictator game

Treatment (# of obs)	Two-person dictator game (Payoff A , payoff B)	Left	Right
Berk29 (26)	B chooses (400,400) vs. (750,400)	0.31	0.69
Barc2 (48)	B chooses (400,400) vs. (750,375)	0.52	0.48
Berk23 (36)	B chooses (800,200) vs. (0,0)	1.00	0.00
Barc8 (36)	B chooses (300,600) vs. (700,500)	0.67	0.33

Source: Charness and Rabin (2002), slightly altered compared to original.

recipient received more than the allocator instead of the same amount. Hence, they preferred Pareto-superior allocations even if this implied increased inequity at their own expense. The other third of the subjects thus exhibited Pareto-damaging preferences induced by difference aversion. If increasing the recipient's payoff was at a little expense of the dictator's payoff (hence also Pareto-damaging), the share of dictators refusing to help the recipient diminished to one-half. All dictators chose the Pareto-superior allocations, even if inequity was increased, given that the alternative would be a zero payoff for both players.

Second, Charness and Rabin introduced the so-called 'response game' (see Figure 2.3, π denotes subjects' payoff). In this game, there was a proposer (' A ', from here on) and a respondent (' B ', from here on). At the first stage, A could decide if she chose one given allocation or let B choose between two other given allocations.

If A chose the allocation ('out'), the game was over and the payoffs were paid according to the one given allocation. If A chose to let B choose ('in'), B chose one out of two allocations and then payoffs were made according to this allocation.¹⁵ Considering the different menus given in this game, there could be different motivations that had to be taken into account. First, the question was whether A chose 'in' or 'out'. Possible motivations could of course simply be about the distribution of the payoffs for A and B as in the menu-based dictator game, depending on the different menus. But this setup implied possible strategic interaction as in the ultimatum game. If A chose 'in' to let B make the decision, this could either be perceived as neutral, nice or malign

¹⁵Note, that in order to maximize observations, B was asked to decide contingent on the case that A chose 'in' without knowing if A actually chose in. From my point of view, this could distort outcomes.

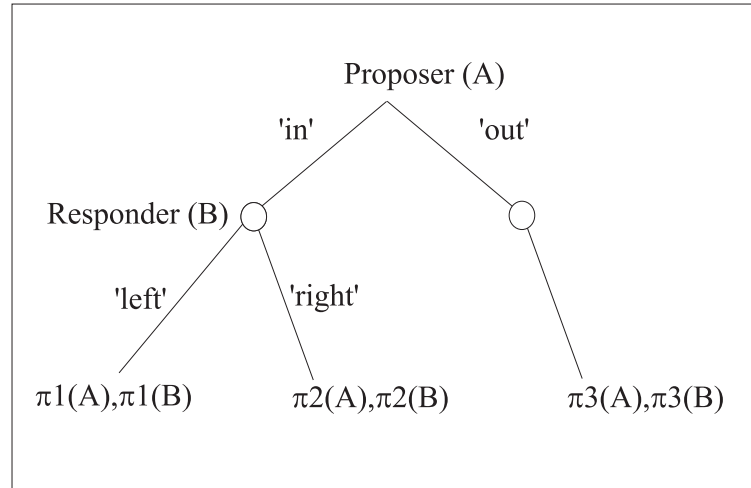


Figure 2.3: Game tree response game. Source: Own

by B , depending on B 's payoff given 'out' and B 's possible payoff given 'in' (see Table 2.4, 2.5).

Depending on the menu, B had the chance to reward nice behavior or retaliate in case of malign behavior. One has to keep in mind that this possible reciprocal behavior might be, depending on the menu, at the expense of B 's own payoff. A could have anticipated this reasoning and adapted his allocation behavior accordingly.

The response game was analyzed with respect to the previous dictator games by Charness and Rabin. This could be done since the allocations which could be chosen by B in the response games were similar to the ones in the dictator games (given that A chose 'in'). Given 'in' and A 's choice helps B (payoff B 'in' > payoff B 'out') and if helping was not costly for B (allocations where the payoff is the same for B in both choices but higher for A in one of the allocations), Charness and Rabin find that reciprocity seemed to dominate difference aversion in case there was a conflict between motivations! Remember that a share of one-third of the allocators exhibited pareto-damaging difference aversion in the dictator games (see Table 2.3, Berk29). This share was reduced to 0.06 in order to reward A for his choice (see Table 2.4, Barc7). Here, there was positive reciprocity. Interestingly, if A chose 'in' and B is worse off by this choice (not much, but worse off), then there seemed to be no negative reciprocal behavior, since B 's allocation behavior remained roughly the same as in the dictator setup. (see Table 2.4, Barc5, and Table 2.3, Berk29).

Table 2.4: Two-person response games - B 's payoffs identical

Treatment (# of obs)	Two-person response games B 's payoffs identical (π_A, π_B)	Out	Enter	Left	Right
Barc7 (36)	A chooses (750,0) or lets B choose (400,400) vs. (750,400)	0.47	0.53	0.06	0.94
Barc5 (36)	A chooses (550,550) or lets B choose (400,400) vs. (750,400)	0.39	0.61	0.33	0.67

Source: Charness and Rabin (2002), slightly altered compared to original.

Taking a look at cases where B had to sacrifice in order to help A , B did not much change his allocation behavior (with respect to the dictator games) as long as A acted neutral by choosing 'in' (see Table 2.5, Barc3, and Table 2.3, Barc2). As soon as A seemed to act selfishly (A hurt B by choosing 'in'), B 's willingness to help A was significantly reduced compared to the dictator game (see Table 2.5, Barc1, and Table 2.3, Barc2). Hence, there was evidence for negative reciprocity. However, if A sacrifices payoff by choosing 'in', B does exhibit no positive reciprocity compared to the dictator game (see Table 2.5, Barc6, and Table 2.3, Barc8).

Table 2.5: Two-person response games - B 's sacrifice helps A

Treatment (# of obs)	Two-person response games - B 's sacrifice helps A (π_A, π_B)	Out	Enter	Left	Right
Barc3 (42)	A chooses (725,0) or lets B choose (400,400) vs. (750,375)	0.74	0.26	0.62	0.38
Barc1 (44)	A chooses (550,550) or lets B choose (400,400) vs. (750,375)	0.96	0.04	0.93	0.07
Barc6 (36)	A chooses (750,100) or lets B choose (300,600) vs. (700,500)	0.92	0.08	0.75	0.25

Source: Charness and Rabin (2002), slightly altered compared to original.

Charness and Rabin tried to reconcile the data with distributional preferences of the players. They compared four different preference types. The 'narrow self-interest' type reflected the classical economics agent who was only interested in his own payoffs. The 'competitive' type mainly would like to maximize the relative distance between her and the other's payoff while still caring about the other's pay-off directly. The difference aversion type was

inspired by Fehr and Schmidt (1999) and Bolton and Ockenfels (2000). His preferences were such that his utility increases with income. But inequity reduces this types utility, hence she would sacrifice money to minimize inequity. The social welfare type (developed by Charness and Rabin) would like to maximize payoffs both for himself and for others. This type was more interested in getting a payoff for himself, if he was behind. But he was willing to accept payoff differences (even at his own expense) if overall payoffs were significantly increased. In these models, intentions and resulting reciprocity were not yet included. Charness and Rabin found that the social welfare preferences fits best with their data when checking for consistency of subjects' decisions.

Charness and Rabin also estimated different utility functions using the different preference types. The model explaining subjects' behavior best implies that allocators only care about the payoff of the ones being worse off than the allocator. They also find that reciprocity is an important component of peoples' preferences. Charness and Rabin conclude that there is strong evidence that subjects' preferences are best matched by the social welfare type.

2.2.14 Experiments that have some similarities with our setup

The following experiments are mentioned because they have some features in common with our setup. Their contribution to the evidence or theory of dictator games is rather modest.

Ben-Ner *et al.* (2004) ran two consecutive dictator games after one another where they changed the role of being allocator to see if reciprocity is important. Departing from the traditional setup, Ben-Ner *et al.* had two stages. In the first stage, there was a traditional dictator game. In the second stage, the roles of being allocator and recipient were exchanged. The allocators in the first stage were not informed that there was a second stage, which is not usual in experimental economics, where procedure is normally common knowledge to subjects. There were two treatments. The first stage was similar for both treatments. In the second stage, roughly half of the recipients were matched with the same allocator as in the round before (treatment 'same match') while the other half was matched randomly with a new dictator (treatment 'random match'). Ben-Ner *et al.* tested for reciprocity between the two stages by using an OLS regression with the share sent from the second stage as dependent variable and the share sent in the first stage plus a constant as independent variables. They ran a regression for both treatments. The results revealed that for 'same match' in both stages, for every \$ sent in the first stage, an

additional 0.83\$ was sent back. This parameter was significant and R^2 for the regression was 0.61. In the treatment with random matching, there was only 0.1\$ sent back for every \$ sent in the first stage and the parameter was insignificant while R^2 is only 0.013. Hence, reciprocity plays a role.

Diekmann (2004) used a design, where he also changed the role in two sequent dictator games and used the expression ‘sequential dictator game’ for his setup. Since the existence of second stage was common knowledge, we have a pure reciprocity case which is not worth investigating since I survey the results for reciprocity setups in section 2.1, the ultimatum game section.

Cason and Mui (1998) conduct two sequential dictator games after one another. Their objective was to see if subjects adapt their choice in the second dictator game if they receive information about the allocation choices of other subjects from the first dictator game. There were no clear results, I mention this experiment because the authors used the name ‘sequential dictator game’ which I would have used if it had not been taken by Cason and Mui.

Karni *et al.* (2001) were the first to introduce two recipients in a dictator game. Instead of assigning money, probabilities of winning 15\$ were assigned. The chances had to sum up to one. Subjects were randomly assigned to being A , B and C . A was the allocator while B and C were recipients. The authors tried to investigate if subjects preference structure can be modelled by minimizing the difference in probabilities between A , B and C or by only minimizing the difference in probabilities between the two recipients B and C . The authors find little evidence for former but some for the latter preference structure.

2.2.15 Hiding behind a random process

Andreoni and Bernheim (2007) investigate allocators’ behavior given that it is unclear for the recipient if the allocator or a random process decided on the allocation. Participants were gathered in a room. Allocator and recipient who were sitting opposite to each other were matched as a pair and received the same group number. Then, both allocator and recipient were one by one asked to stand up and state: ‘Hello, I am in group number X, I am your partner’. There were two treatments. In both treatments, nature could allocate instead of the subject with a certain probability which ranged from 0 (hence, it was certain that subjects allocated) to 0.75 (hence, in three out of four cases nature decided). Note, that the experimental setup was designed in such way that allocators knew if they decided or nature while the recipients were only aware of the probability that nature decided. Nature’s allocation was decided by

tossing a dice in front of the audience. In treatment 0, in case nature decided either the allocator or the recipient received the whole pie. Probability of both events was one half. Treatment 1 was similar to treatment 0, except that the allocation was 19\$ and 1\$. The important feature of the experiment was that if allocators chose the same allocation as in the case where nature chose, recipients could not infer from the allocation whether nature or the allocator made the choice.

In treatment 0, in case the probability that nature decided was zero, a proportion of 0.57 of the dictators divided equally and a share of 0.3 took the whole pie. If nature was introduced with probability of 0.25, the proportion of fair allocators was reduced to around 0.32 while proportion of allocators passing nothing increased to around 0.53. For the probability of nature coming into play with 0.5 (0.75), the proportion of fair offers decreased to 0.27 (0.27), respectively, while the proportion subjects taking all increased to 0.72 (0.70).

In treatment 1, given that nature was not involved, a proportion of 0.69 divided equally, a proportion of 0.17 kept the whole pie and a proportion of 0.03 kept 19\$. Increasing the likelihood that nature decided reduced the proportion of fair divisions to proportions of 0.45 to 0.35, respectively, while the proportion of subjects taking all was reduced to 0.1. Note, that the proportion of the subjects who chose to pass one (hence making it impossible for the recipient to determine if nature moved or not) increased from 0.03 to proportion of >0.35 if nature was introduced with probabilities ≥ 0.25 .

Andreoni and Bernheim used two random-effect probit models for their analysis. In the first model, the dependent binary variable was ‘choosing the allocation that cannot be distinguished from nature’ or ‘not choosing this allocation’. Independent variables were dummy variables for the different probabilities that nature comes into play and treatment 1. All dummies for the different probabilities >0 of nature coming into play were significant and positive. This makes sense since the more likely it was that nature came into play the more ‘comfortable’ it was to ‘hide’ one’s choice behind nature. Also the dummy for treatment 1 was significant. That might reflect the subjects who were unconcerned with social image and still chose the maximum amount for themselves instead of taking 19\$ and thereby hiding behind nature.

In the second regression, the dependent binary variable was ‘choosing the fair split’ or ‘not choosing the fair split’ while the regressors remained the same. Dummies for nature were significant again, but now with negative sign, which makes sense, since it became harder to attribute unfair behavior to allocators. The dummy for treatment 1 was not significant. Andreoni and Bernheim conclude that the degree of fairness among people varies. There is

strong support that people would like to be seen as fair, hence audience effects play an important role in subjects' decision behavior. Thus, many subjects will choose an unfair split if they are anonymous while being generous if their allocation is public information.

2.2.16 Altruism in dictator games - an experimental artefact?

Bardsley *et al.* (2008) were the first to actually allow allocators to take money from the recipient in a dictator game (the experiments were run before 2005). His study was designed to check if the results obtained from dictator games so far were a scientific artefact triggered by subjects desire to conform to the expectations of the experimenter. In one experiment, allocators received an endowment of £10 while the recipients received an endowment of £5. In treatment 1, allocator could give £0-3, while in treatment 2 they could take £0-3, but were not allowed to give. Increments were 0.25 in both treatments. A share of 0.55 gave money in treatment 1, while a share of 0.83 took money in treatment 2. The proportion of subjects giving in treatment 1 should not have been larger than the ones not taking in treatment 2, since this would mean that that a share of subjects was giving in one treatment and while taking in another treatment. *Ceteris paribus* this would not be rational.

Bardsley claims that opportunities trigger certain behavior. This effect could be explained in such way that subjects infer from the experimental situation to a particular behavior that seems to be adequate and try to conform with this. Hence, in the traditional dictator game, the mere opportunity to give conveys the impression to the subjects that they are expected to give something by the experimenter. In the dictator game where taking is allowed, subjects infer from the opportunity to take that they are expected by the experimenter to take something and hence conform to this expectation, taking higher shares than compatible with the results from the traditional dictator game.

Bardsley concludes that one cannot derive fixed preferences from experimental results in the way e.g. Andreoni and Miller (2002, see section 2.2.13) and Charness and Rabin (2002, see section 2.2.13) do.

List (2007) replicates the basic idea by Bardsley (2008). Experimental procedure was similar to FHSS unless some modifications with respect to endowment. There were four treatments. In treatment 1, both players received 5\$, allocators received an additional 5\$ that they could allocate in the spirit of the dictator game. Treatment 2 was similar to treatment 1, but subjects were

also allowed to take 1\$. Treatment 3 was identical to treatment 2, but allocators could take up to 5\$. Treatment 4 is identical to treatment 3 except that subjects had to spend some time earning the money during the experiment (by sorting and handling mail) instead of simply receiving their endowments

In treatment 1, subjects passed a mean share of 0.27, a proportion of 0.71 made positive offers. In treatment 2, mean share passed was 0.07, a proportion of 0.35 made a positive offer, while a proportion of 0.21 took 1\$. In treatment 3, mean share passed was -0.5, only a proportion of 0.1 made positive offers, while a proportion of 0.42 took 5\$. In treatment 4, mean share passed was -0.2, a proportion of 0.06 made a positive offers, while a proportion of 0.66 neither took nor gave something. The decrease of positive offers of treatment 2 compare to treatment 1 was significant. List claims that adding an additional option to the available option set has a significant impact. The proportion of positive offers of treatment 3 was significantly lower than the one in treatment 2. Hence, introducing symmetry in the option set by allowing a pass range from -5 to +5 instead of -1 to +5 induced subjects to significantly change their behavior. In treatment 4, a large share of the subjects neither took nor gave. The proportion of subjects taking was significantly lower than in treatment 3. Thus, List states that the origin of the endowment is also crucial for the experimental outcome. Taking away earned money increases ‘moral’ costs for subjects.

List comes to the same conclusion as Bardsley (2008). He states that experimental results from dictator games should not be interpreted by trying to construct preferences from the choices but should be seen as altered expectations on the supposed adequate behavior. List challenges the usefulness of dictator games at all. He claims that the experimenter is an authoritative figure who has credibility in the eyes of the subject. If this authority asks whether the subject would like to share money with someone who did not receive as much money as the allocator, subjects conform to what they perceive as socially adequate, and pass on money. He stresses the importance of being careful to export insights from lab experiments to the real world. Nonetheless, he acknowledges that lab experiments provide insight how institutions could affect individuals.

2.3 The impact of gender in dictator games

The evidence on gender effects in dictator games is mixed. Some authors find that gender does not matter, some authors observe significant effects of gender.

The first to investigate this issue were Bolton and Katok (1995). They used the 1Game-6Card and two other treatments by Bolton *et al.* (1998) (see section 2.2.5) to test for gender differences within each treatment but do not find any significant differences.

Eckel and Grossman (1998) conducted an experimental study with the objective to investigate the impact of gender. They applied the ‘double blind 1’ procedure by HMSS. In half of the experimental sessions all the allocators were men, in the other half women. It is not mentioned in the article, if the recipients were also unisex. Women roughly sacrificed twice as much as men with a mean share of passed money of 0.16, while men only donated 0.08. The difference between the distributions of passed money by men and by women was significant. Also comparing distributions conditional on passing a positive amount yielded a significant difference. Eckel and Grossman conclude that women are more generous than men.

Andreoni and Vesterlund (2001) used the data from Andreoni and Miller (2002, see section 2.2.13) to check for gender differences in behavior. Unlike Eckel and Grossman, there were no unisex groups, selection was not controlled for gender. Men passed an average of 2.56\$ while women passed 2.60\$ which is not significantly different. Taking the ratio of the hold and pass value (see section 2.2.9 for an explanation of the concept of hold and pass values) and thereby constructing the relative price of giving, a new picture emerges: Depending on this price, there were gender-related differences. The amount of passed money (in pass values) was increased for both male and women if the relative price of giving was reduced. But male subjects increased the proportion given with respect to their endowment while women did the opposite. This difference between proportions was significant for a price larger than one, while for a price lower than one the difference was marginally significant. Andreoni and Vesterlund conclude that there is a gender difference in the price elasticity of demand of ‘generosity’.

Andreoni and Vesterlund also analyzed whether there are differences between men and women with regard to preference types. They use the classification by Andreoni and Miller (2002, see section 2.2.13). They found that a share of 0.47 of the males were of the selfish type while this share of women was only 0.36. Men were more likely to see others payoff as perfect substitutes for their own. With regard to the Leontief preferences, the share of males was 0.25 while the share of women was 0.54. Testing, if the utility function was independent from gender was rejected significantly. Men were more likely to be of the traditional type of the perfectly selfish agent or to maximize joint payoffs while women were more likely to be in favor of an equitable outcome. Compar-

ing their results with the other results, Andreoni and Vesterlund reconcile the mixed results by Bolton *et al.* and Eckel and Grossman. They mention that the variance of giving was maximum at a relative price of giving equal to one. This might lead to different results due to this high variance if the sample size was not too large (which was the case in Bolton *et al.* 1998). They conclude that contingent on the setup men or women can be more altruistic.

2.4 The impact of study subject on the propensity to be altruistic

Marwell *et al.* (1981) were the first to find that study subject might matter in people's allocation behavior. In a public good provision experiment, graduate economists were significantly less fair than non-economists. Graduate economists invested only a share of 0.2 of their endowment for the public good, unlike the rest who, depending on the treatment, invested between 0.4 and 0.6. Marwell *et al.* (1981) tried to explain this phenomenon with two reasons. First, there could be a selection effect: People with a certain personality might be attracted by studying economics and this particular personality might increase the inclination to be egoistic. Second, there may be an indoctrination effect since economists might be more prone to act 'rationally' because they internalized the theory they have been taught and act accordingly.

Carter and Irons (1991) were inspired by this experiment and used a standard ultimatum game setup (the one suggested by Kahnemann, Knetsch and Thaler, 1986) to check if the previous results could be sustained. Student subjects belonged to four groups: Freshmen non-economists, freshmen economists, senior non-economists and senior economists. Carter and Irons (1991) compared the results using a regression analysis. First, the minimum amount accepted (not distinguishing between freshmen and seniors) was used as dependent variable which was regressed on a constant and a dummy for economist. They found that the dummy is negative and significant. They conclude that there is a difference between economists and non-economists. They used the same equation while adding a senior and a senior economist dummy in order to control for a maturation effect. The economist dummy remained negative and significant, the senior economist dummy was positive and significant and the senior dummy was negative and marginally significant. Carter and Irons (1991) conclude that there is a selection effect, economists are more greedy from the begin on, while studying economics does not increase greed. In a second step they ran regressions with the proposal as dependent variable. Using

only a dummy for the economists, this dummy was positive and significant. Including again a senior and a senior economist dummy lead to non-significant results for all coefficients. Despite the non-significance of the regressors, Carter and Irons state that this confirms the selection effect. From my point of view, using non-parametric sample comparison methods would have been a more adequate methodology for this kind of analysis.

Gross (2005) investigated the effect of studying economics in a dictator game. Gross did not run any test on sample distribution difference. But around 0.75¹⁶ of the economists pass nothing and around 0.09 split fairly while around 0.39 of the non-economists pass nothing and around 0.27 split fairly. Gross concludes that economists may have less concern for fairness.

Meier and Frey (2004) investigate the donations of a very large sample of students from Zürich with 180225 observations. Students were asked every semester whether they would like to donate money for needy or foreign students. Using a Probit Regression with ‘donating or not’ as binary dependent variable, they found that the probability that an economist student contributes is lower by 0.04 compared to a non-economist which was significant. They also controlled for other socio-economic features. Meier and Frey claim that economic students’ reduced generosity stems from a selection effect while there is no indoctrination effect (see Marwell, 1981).

2.5 The role of intentions in sequential allocation behavior

It is intuitive that it matters for people whether an action was performed intentionally or whether it was unintentional (maybe as a byproduct of another action or just accidentally). There is a perceptual difference if someone had the impression that another person stepped on her foot because this person did not pay attention or because she did it on purpose. The problem with these cases is that often the intentions of someone else are not that obvious to other people. Nonetheless, people still could attribute intentions to actions.

2.5.1 Intentional action triggers reciprocity in a common pool resource game

Rutte *et al.* (1987) investigated the impact of intentional action in a common pool resource (CPR) setup. They analyze whether subjects extraction behavior

¹⁶The exact share is unclear, I inferred the share from a figure

from a CPR is affected by previous subjects' extraction behavior from this CPR. Subjects could take 0 to 10 guilders from a resource pool and they were told that the pool initially contained 5 to 55 guilders (randomly chosen by a computer). There were six subjects extracting sequentially, subjects were told at which position of the sequence they were extracting. The game was played once. In the first treatment, subjects were aware of both the initial pool size and the current pool size at the position they were extracting while in the second treatment they were only told about the current pool size at the position they were extracting but not about the initial pool size. Hence, in the first treatment, it was possible to attribute intentions to previous subjects' behavior while in the second treatment this was impossible.¹⁷ Rutte *et al.* found that subjects in the first treatment responded to perceived intentions from previous subjects' extractions (in comparison to the second treatment). If previous extractors took deliberately more than the equal share in the first treatment, subjects reciprocated by being less generous (in comparison to the second treatment). If previous extractors took deliberately less than the equal share in the first treatment, subjects reciprocated by being more generous (in comparison to the second treatment). The difference in behavior between the two treatments was significant.

2.5.2 Intentional action triggers reciprocity in an ultimatum game

Blount (1995) did a study of a modified ultimatum game investigating the role of intentions by replacing the proposer with either a third party or a random process. There were three treatments. The first treatment, called 'baseline', was a traditional ultimatum game. In the second treatment, called 'third party', another student was involved. This third party was drawn from a sample of students, who were not in the sample for the experiment but from the same graduate school. The third party neither knew the proposer nor the respondent. The third party decided on the proposed allocation while not receiving anything from this allocation. The (passive) proposer would still receive the payoff if the proposition (by the third party) was accepted by the recipient. The last treatment, called 'random', was similar to 'baseline' except that a random process (generated by a computer) proposed the split instead of a human proposer, possible outcomes (increments of 1\$) were uniformly

¹⁷Note, that in this experiment, there was a large discrepancy between what subjects have been told and how the experiment was actually conducted. Since it was an experiment by psychologists, this is not unusual, but the standard for economic experiments is that description and actual procedure are identical.

distributed. In each treatment subjects were asked about their minimum acceptable offer

Mean minimum share of acceptable offers was 0.29 in the ‘baseline’, 0.21 in ‘third party’ and 0.12 in ‘random’.¹⁸ The distribution of acceptable offers of ‘random’ was significantly different from ‘third party’ and from ‘baseline’. ‘Third party’ and ‘baseline’ were not significantly different. Blount concludes that in the random treatment subjects acted more in spirit of the selfish agent who maximizes payoff, while in the other two treatments subjects also responded to intentions by forfeiting offers which they perceived as too low. Blount concludes that attribution plays a crucial role in social decision contexts. Perceived intentions are important, triggering the desire to retribute in case of a perceived unfair action.

2.5.3 Mainly negative intentions trigger reciprocity

Offerman (2002) investigated whether there was a difference in subjects’ response to actions that were either intentional or unintentional. He also analyzed whether the nature (positive/negative) of the intended action played a role. The experiment was a sequential game. At the first stage, the first mover decided between a helpful choice giving himself 8 guilders and the second mover 4 guilders or a hurtful choice where the first mover received 11 guilders and the second mover -4 guilders. At the second stage, the second mover observed the first mover’s choice and then decided, whether the payoff from the first stage of both players was increased, reduced or left unaltered. The choice was between three moves: ‘Cool’, where the first mover’s payoff was unaltered and the second mover received additional 10 guilders, ‘reward’, where the first mover obtained 4 additional guilders and the second mover 9 additional guilders or ‘punish’, where 4 guilders were subtracted from the first mover’s payoff from the first stage while the second mover obtained additionally 9 guilders. There were two treatments, in the first treatment ‘nature’, a random process decided on the first movers choice between ‘helping’ and ‘hurting’ with a probability of 0.5 for each choice. This was done by the first mover tossing a dice. In the second treatment ‘hot flesh’, the choice was made by the first mover.

In case of the helpful choice, a share of 0.75 reciprocated positively in the ‘hot flesh’ treatment (choosing ‘reward’) while only a share of 0.5 did so in the ‘nature’ treatment, the difference between both treatments was not signif-

¹⁸Responders had to state their minimum acceptable proposal in advance of getting to know the true proposal

icant. The rest chose the ‘cool’ reaction, none punished for helping. In case the first mover chose hurting, a share of 0.83 chose to punish in the ‘hot flesh’ while only a share of 0.17 did so in the ‘nature’ treatment, which was significant. Obviously, negative intentionality was more important than positive intentionality. Offerman also asked second movers about their emotions after they made their choice. For helping, positive emotions were triggered but there was no significant difference between the ‘nature’ or ‘hot-flesh’ treatment. But for hurting, negative emotions were always weak in the state of the ‘nature’ treatment while for the majority of the ‘hot-flesh’ treatment, negative emotions were strong. This difference was significant. Offerman explains the difference in reactions on positive and negative intentions with a self-serving bias. Subjects tend to attribute good events to internal causes like competence but bad events to external causes as uncontrollable circumstances. He claims, that subjects perceive it as normal that someone helps ‘someone as nice as oneself’. Hence, reciprocating is not imperative since this action is assumed to be standard. While being intentionally unfair is perceived as an insult which demands retribution.

Knobe (2006) claims that moral considerations play an important role in reasoning. If an action is good or harmful this has an impact on peoples’ judgement on the intentions of the action. Knobe used the following case in an experiment in order to shed light on the influence of moral judgement on people’s impression with respect to intentions. In the first treatment, subjects were confronted with a simple case: The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.’ The chairman of the board answered, ‘I dont care at all about harming the environment. I just want to make as much profit as I can.’ They started the new program. Sure enough, the environment was harmed.

In the second treatment, the word ‘harm’ was replaced by ‘help’, such that the vignette became: The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, and it will also help the environment.’ The chairman of the board answered, ‘I dont care at all about helping the environment. I just want to make as much profit as I can. Lets start the new program.’ They started the new program. Sure enough, the environment was helped.

A share of 0.82 of the subjects who received the story about environmental harm said that the chairman harmed the environment intentionally, whereas a share of only 0.23 of the subjects who received the story about environmental help said that the chairman helped the environment intentionally.

This simple connection has been proven in several different contexts in different experimental studies (Knobe 2003a, Knobe 2003b, Knobe and Burra forthcoming, Leslie and Knobe 2004, Nadelhoffer 2004). Hence, if a side-effect was good, people were less likely to infer that it was brought about intentionally than in the case where the side effect was bad.

2.6 Common pool resource experiments

Common pool resource (CPR) experiments are usually more complicated than dictator games. In most cases there is a common resource that can be extracted over periods and if extraction is not too large, the resource will grow. In an experimental one generation setup of a CPR with multiple players and a number of periods larger than one there is usually a conflict between socially and individually optimal behavior: Often, it is socially optimal to let the resource grow and to not extract too much (depending on the setup) while there is an individual incentive to preempt others on extracting as much as possible.

Now assume a setup with an intergenerational component. The problem with regard to intergenerational CPR approaches is that they confront subjects with complicated strategic considerations. Given that subjects do not only care about their own payoff but also about the future payoff of all other generations, they have to build expectations on the following generations behavior to determine their optimal decision. It may be the case that subjects assume that future generations will be selfish and exploit their generosity by appropriating the whole surplus passed on instead of passing most of it on to the following generations. Thus, altruistic subjects would prefer to behave selfishly themselves. It may also be the case that altruistic subjects assume that following generations will be altruistic as well, then they would sacrifice resources for the sake of these future generations. Hence, although in both cases subjects have a latent desire to be altruistic, this desire only manifests given that particular circumstances with regard to expectations are met. Hence, these experiments make it difficult to interpret subjects' decisions in terms of other-regarding behavior.

2.6.1 An intergenerational approach

Chermak and Krause (2002) investigated intergenerational behavior with regard to a CPR. In addition, they analyze how information affects CPR-related consumption decisions. Subjects were randomly assigned to groups consisting

of three generations. Group membership was changed each trial.¹⁹ Subjects of the first generation played from period 1 to 3, the second generation from 2 to 4 and the third generation from 3 to 5. There were two treatments. In the first treatment, subjects knew ex-ante to which generation they belonged. This treatment was called ‘informed’. The second treatment, called ‘uninformed’, was similar, except that subjects got to know to which generation they belonged ex post and . Each player had the possibility to consume a part of the resource in each period. The common pool resource had a specific growth rate depending on the extent of the extraction of the three generations participating. The resource stock could increase, decline, stay constant or be completely depleted (in this case the game ends). The consumption determined the subjects payment. After every period, subjects were informed about the current resource stock. If the resource was not completely depleted at the last trial, the rest was given to a charity, chosen by the third generation.

In ‘informed’, a share of 0.14 exhausted its resource, in the ‘uninformed’ 0.17 did so. 0.13 of the subjects behaved sustainably, allowing the following subjects to extract the same amount they took, 0.18 were trying to maximize payoffs, extracting most in their last period, while 0.27 were following the classic economic behavior using backward induction and extracting most in the first period. The strategies of the rest were unclear. The mean level of the resource at the end of the last period was rather high (more than nine times the initial level though, though one has to note that a growth factor of three was part of the formula governing the resource path). On average, subjects withdrew less than the sustainable level. Standard deviation of extraction was substantial. It did matter if subjects were in the ‘uninformed’ or the ‘informed’ treatment. In the ‘informed’ treatment, the pattern of strategies followed was much clearer. The results of the ‘uninformed’ and ‘informed’ treatment did only differ significantly the first two periods, from the third period on they were not significantly different. Chermak and Krause conclude that subjects adapt their behavior to the position in an intergenerational chain in case they are aware of this position.

2.6.2 Intergenerational generosity: Triggered by the desire to give or to achieve equity?

The objective of the study by Sadrieh (2003) was to find out whether intergenerational generosity is triggered by subjects’ desire to pass something on (without the aim to achieve a specific equity distribution) or whether subjects

¹⁹If the membership to a generation was changed, too, is unclear from the article.

explicitly have the goal of obtaining an equitable outcome. Sadrieh identified the ‘warm glow’ type who derives utility from the mere act of giving and the altruistic type who derives utility from others’ increased utility or equitable outcome. Anonymity was ensured according to ‘double blind’ by HMSS.

The CPR was constructed as follows: The part of the resource that was not extracted grew and was then at the disposal of the future generations. The payoff of the subject was also subject to a formula, depending on the extraction. There was an individually optimal amount to be extracted. Extracting too little or too much resulted in less payoff. I will focus on two treatments of the experiment: In the ‘very slow’ treatment, the individually optimal extraction rate was higher than the growth rate of the resource, in the ‘very fast’ treatments, this was reversed. There were 15 periods, but both the number of periods and which period the subjects were assigned to was unclear to the subjects, they only knew that the number of periods was limited. A subject did not play all generations but only played for one generation, hence we have a one-shot game.

Subjects maximizing their own payoff simply had to choose an individually optimal extraction level that was unique. If subjects wanted to achieve intergenerational equity they would have had to alter their individually optimal extraction level depending on the treatment. They would have to exhibit ‘growth compensating’ behavior: In the ‘very fast’ treatment they would have to increase their extractions compared to the individually optimal behavior to compensate for the high growth rate in order to obtain an intergenerationally equitable outcome. In the ‘very slow’ treatment, extractions must be much smaller compared to the individually optimal behavior (and hence compared to the ‘very fast’ treatment as well) in order to compensate for the reduced growth rate. If subjects only wanted to obtain a warm glow of giving, such considerations should not matter, they would simply pass on more than individually optimal regardless of the treatment.

The results showed that subjects on average depart from the individually optimal level by choosing lower levels of extraction, hence displaying a propensity for altruism. The slower the growth rate the more subjects reduce the extraction below the individually optimal level, the difference between ‘very slow’ and ‘very fast’ was significant. Subjects caring about intergenerational equity should overexploit the resource in the fast growing treatment. But the overwhelming share of the subjects did the opposite. Only a share of 0.08 of the subjects chose to take more than individually optimal. Thus, there was clear evidence that altruism was not motivated by equity concerns. In ‘very slow’, a share of 0.36 chose the individually optimal level, while a share of 0.46

did so in ‘very fast’. Thus, roughly half of the subjects exhibited a propensity for intergenerational altruism, while the other half was strictly selfishly motivated. Sadrieh concludes that in case intergenerational altruism was present, it was mainly motivated by the warm glow of giving.

2.6.3 Investigating intergenerational altruism by eliciting expectations

The experiment by Fischer *et al.* (2004) was a CPR game that was close to the one by Sadrieh (2003). Fischer *et al.* tried to shed light on the motivations of the subjects with regard to altruism by asking for expectations on the other subjects’ behavior. With the help of the expectations it was easier to evaluate the motives behind an extraction choice. As in Sadrieh (2003), subjects knew there was a limited number of periods but not how many periods there were overall and to which period they were assigned to. Subjects only played for one generation, hence we had a one-shot game, too. In the experiment, there were four sequential generations of three players extracting once at the same time from the same resource. The features of the CPR can be described as follows: There was a hump-shaped production function of the resource, having the sum of the exploitation of all players as argument. Each subjects production (and hence their earnings from the experiment) depended positively on subjects’ individual extraction level and positively on the stock of the resource. As soon as total exploitation was larger than the social optimum, a negative externality arose, induced by the hump-shaped production function. In the individual optimum, assuming rational agents who care only about their own payoff, players extraction would be well above the social optimum. There were three treatments. The baseline treatment, called ‘restart’, had no intergenerational features, the stock of the resource was held constant every period. The two other treatments were intergenerational. In these treatments, the stock depended both on its growth rate and on extraction levels of the former periods. In the second treatment, called ‘slow’, the growth rate was rather low which meant if subjects wanted to obtain an intergenerationally equitable outcome, they would need to extract less than individually and socially optimal. In the last treatment, called ‘fast’, the resource growth rate was rather high, implying that subjects who would like to achieve intergenerational equity should extract more than individually and socially optimal.

Extraction levels were significantly smaller in ‘slow’ and in ‘restart’ than in ‘fast’, while there was no significant difference between ‘slow’ and ‘restart’. In all treatments, the means of extraction were significantly lower than in

the individual but larger than in the symmetric social optimum (each subject exploits the same amount and the social optimum is achieved). Hence, there was a willingness to forfeit personal consumption possibilities for the sake of intra- and intergenerational fairness. Compared to the growth compensating path, there were significantly smaller extractions in the ‘fast’ treatment and significantly larger extractions in the ‘slow’ treatment. Thus, there is definitely no evidence for growth compensating behavior confirming the evidence by Sadrieh (2003).

Subjects were also asked about their expectations about others’ decisions (the better the guess, the more money was paid). Comparing subjects expectations about others’ extraction levels and their own choices revealed that subjects deliberately chose a level below the individual optimum, subjects intended to be altruistic. Interestingly, there was more intended free riding in ‘slow’ compared to the other treatments. In general, subjects were too optimistic about cooperation of the others, especially in ‘slow’. Fischer *et al.* conclude, that the presence of an intergenerational element does positively affect subjects’ expectations about the others’ behavior, but since subjects did behave the same way in ‘slow’ and ‘restart’, the consciousness of the intergenerational problem does not result in behavioral alteration for the subjects themselves.

Chapter 3

Statistical methodology

The usual method for the analysis of dictator games is to compare different treatments of one setup. In most cases, the distributions of the allocations of different treatments are compared (see section 11.3.2). In experimental contexts different from dictator or ultimatum games, treatment comparison is usually done by the mean difference test, the z-test. But this test assumes a normal distribution of the sample which is usually not given in the non-normally distributed dictator game results. Hence, non-parametric tests are commonly used. I will present these non-parametric tests in the next section. In the second section of this chapter I will justify the choice of tests for the analysis of the results of the intergenerational dictator game.

3.1 Mann-Whitney test

The Mann-Whitney test (MW) (see e.g. Conover 1971) is almost similar to the Wilcoxon-Rank-Sum test (WI). For the WI the observations of the two samples to be compared are pooled and ordered. Then, ranks are assigned according to their position in the order (from lowest to largest). In case of ties the average of the ranks of the ties is taken instead. Then the sum of the ranks of each sample is taken, the smaller resulting value is the Wilcoxon-Rank-Sum-statistic whose formula is given by

$$WI = \sum_{i=1}^n R(X_i)$$

The smaller WI, the larger the difference between the two distributions of the samples. The intuition behind this is that for very different distributions, the lower ranks should concentrate on one distribution and hence WI is small.

The non-standardized Mann-Whitney test statistic is simply a linear trans-

formation of the Wilcoxon-Rank-Sum-statistic.

$$MW1 = WI - \frac{n(n+1)}{2}$$

where n denotes the sample size of the sample which yielded the smaller WI.

To calculate the MW statistic and the corresponding p-values I used the Matlab-routine `[p,h,stats]=ranksum(X1,X2)`, where X1 and X2 are the names of the two samples. For large samples, MW1 follows a normal distribution. The values given for MW in the following sections are calculated by standardizing MW1:

$$MW = \frac{MW1 - \mu_{MW}}{\sigma_{MW}}$$

$$\mu_{MW} = \frac{n_1 n_2}{2}$$

$$\sigma_{MW} = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

3.2 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test statistic, KS, (see e.g. Conover 1971) is based on the empirical distribution functions of the two samples which are compared. I denote the empirical distribution functions by $F(x)$ for this chapter. The empirical distribution function is a cumulative probability distribution function that concentrates probability $1/n$ at each of the n elements of the sample. The Kolmogorov-Smirnov test measures the largest vertical distance between the two cumulated distribution functions of the two samples. Thus, the larger this distance is (the number of observations is taken into account calculating p-values) the more likely it is that the two samples have different distributions. The test statistic is given by

$$KS = \sup_x |F_1(x) - F_2(x)|$$

3.3 Cramér-von Mises test

The Cramér-von Mises test, CM, is derived from the Kolmogorov-Smirnov test. Here, roughly speaking, we sum up the squared differences between the two empirical distribution functions of the two samples. Hence, all distances between the cumulated distribution functions of the two samples are taken into account. The larger this statistic the more likely we have two samples with

different distributions. The test statistic is given by

$$CM = \frac{n_1 n_2}{(n_1 + n_2)^2} \int_{-\infty}^{\infty} [F_1(x) - F_2(x)]^2 dF_{12}(x)$$

$F_{12}(x)$ is the joint empirical distribution function of both samples.

3.4 Anderson-Darling test

The Anderson-Darling test, AD, is very similar to the Cramér-von Mises test. The only difference is that the tails of the two cumulated distribution functions are weighted higher than in the Cramér-von Mises test which is accomplished by introducing the factor $w(x)$ which increases the closer $F_{12}(x)$ comes to 0 or 1.

$$AD = \frac{n_1 n_2}{(n_1 + n_2)^2} \int_{-\infty}^{\infty} w(x) [F_1(x) - F_2(x)]^2 dF_{12}(x)$$

where

$$w(x) = [F_{12}(x)(1 - F_{12}(x))]^{-1}$$

To the best of my knowledge, no program routines for the AD statistic is publicly available, hence I wrote a Matlab program for the AD (see appendix, section 11.2.1). Later, I obtained a GAUSS-program from Martin Sefton for the AD statistic (see appendix, section 11.2.2).

A problem with the AD statistic is that e.g. unlike the Epps-Singleton statistic, it does not follow a parametric distribution which can easily be obtained and hence is not included in standard statistical packages. Thus, the critical values have to be approximated. The AD code by Sefton includes a Monte Carlo simulation where this is done. The program generates two random samples and calculates the corresponding AD-statistic from comparing these two random samples. This is done e.g. 10000 times. Each time the AD-statistic of comparing the two random samples is larger than the AD-statistic from the actual sample comparison¹, we add 1/10000 to a starting value of zero. This yields the error-I statistic and hence the p-values.²

¹From the two samples we want to compare.

²Hence, by repeated random sampling we obtain the share of randomly generated AD-statistics that are larger than our statistic. For this share, we would have the case that we reject the null hypothesis of identical distributions of the samples although the distributions of the samples are actually the same (both are randomly generated). This is nothing else but the error-I statistic we need. Due to the law of large numbers this procedure is fairly accurate if done with enough repetitions.

Note also, that with repetitions of 10000, the p-values may vary marginally. This variation is usually limited to the third decimal place.

3.5 Epps-Singleton test

The Epps-Singleton test, ES, is based on the empirical characteristic functions (CF) of the two sample distributions. These CFs are Fourier transformations of the sample distribution functions. By X_{km} I denote observation m in sample k where $m = 1, 2, \dots, n_k$. Each observation X_{km} is transformed according to the following formula

$$g(X_{km}) = (\cos t_1 X_{km}, \sin t_1 X_{km}, \cos t_2 X_{km}, \sin t_2 X_{km})'$$

Then, for each sample we take the mean of the transformed samples, g_k , $k = 1, 2$

$$g_k = n_k^{-1} \sum_{m=1}^{n_k} g(X_{km}).$$

The transformation parameters t_1 and t_2 are given by

$$t_1 = \frac{0.4}{\hat{\sigma}}$$

and

$$t_2 = \frac{0.8}{\hat{\sigma}}$$

The parameters t_1 and t_2 have been chosen by Epps and Singleton after various simulations in order to maximize the power of the test. $\hat{\sigma}$ is used to standardize t_1 and t_2 .

$$\hat{\sigma} = 0.5 \left[\frac{(Y_U + Y_{U-1})}{2} - \frac{(Y_L + Y_{L+1})}{2} \right]$$

where Y is the combined sample of both samples in ascending order. L is the greatest integer in $\frac{n_1+n_2}{4}$ and $U = n_1 + n_2 - L$.³

The Epps-Singleton statistic is based on the squared distance between the means of the transformed samples:

$$W = (n_1 + n_2) (g_1 - g_2)' \hat{\Omega}^{-1} (g_1 - g_2).$$

This squared difference is weighted by $\hat{\Omega}$ which is an estimator for the co-

³Note, that unlike in the original article by Epps and Singleton (1986) and unlike in the article by Hoffman *et al.* (1994), where the formula for ES was copied from Epps and Singleton (1986), $\hat{\sigma}$ is not a sum but the difference of $0.5(Y_U + Y_{U-1})$ and $0.5(Y_L + Y_{L+1})$.

variance matrix of $\sqrt{(n_1 + n_2)}(g_1 - g_2)$. Hence, $\widehat{\Omega}$ normalizes W . $\widehat{\Omega}$ is given by

$$\widehat{\Omega} = 0.5(\widehat{S}_1 + \widehat{S}_2)(n_1 + n_2)(n_1^{-1} + n_2^{-1})$$

where \widehat{S}_1 and \widehat{S}_2 are the sample covariance matrices of the elements of g_1 and g_2 given by

$$\widehat{S}_k = n_k^{-1} \sum_{m=1}^{n_k} g(X_{km})g(X_{km})' - g_k g_k'$$

Finally, a small sample correction has to be applied if one of the samples has less than 50 observations (which is always the case in my experimental study). Thus, the Epps-Singleton statistic, ES , is given by

$$ES = W(1 + (n_1 + n_2)^{-0.45} + 10.1(n_1^{-1.7} + n_2^{-1.7}))^{-1}$$

P-values for ES are given by $\chi^2(4)$.

To the best of my knowledge, no program routines for the ES statistic is publicly available, hence I wrote a Matlab program for the ES statistic (see appendix, section 11.2.3)

3.6 Justification for the choice of tests for the data analysis

FHSS compare different non-parametric tests (KS, CM, AD, MW, and ES) using Monte Carlo simulations. They find that the AD and the ES test have the highest power. I have summarized which statistics have been used in the overview on experimental studies of dictator games (see appendix, section 11.3.2). Interestingly, FHSS have been the only ones who used the AD. The CM was not used since it has the same underlying concept as the AD but a lower power. If the KS is used, this is done mostly in addition to other tests. The ES statistic has been used more often. The test which has been most commonly used, especially in the later publications, is the MW. Since this test relies on a concept different from ES and AD, I will include it as well.

A shortcoming none of the authors so far addresses has to be mentioned. Often, there are replications conducted to test if the experimental study is in line with the results from standard studies. To do so, the tests mentioned above are applied and as soon as there is no significant difference⁴ it is assumed that the samples are drawn from the same population. However, the only thing that can be inferred from this is that these samples are not significantly different,

⁴Even though in some studies e.g. by HMSS the test statistic is close to being significant.

the reverse is of course not true, since there is no test controlling for type II error. Still, I apply the same methodology due to the lack of alternatives.

In later studies, it has become popular to not only compare distributions but also the proportions of subjects passing on a positive amount. Unfortunately, significant differences are only found here if the impact of a specific treatment is strong due to the rather small sample sizes which are used in dictator games.

Regressions (in order to control for socio-economic features) are rarely used in dictator games. Moreover, it is well known that in regressions with dictator game data, the error terms are mostly non-normally distributed. This is also the case for the first generation's allocation choice of this study, the Jarque-Bera statistic rejects the normality hypothesis with a p-value of 0.005. This implies that the t-statistic, and hence the p-value in general, cannot be interpreted in the same way as with normally distributed error terms. However, the regression is still unbiased and this effect vanishes when sample size increases. According to Ratcliffe (1968), a number of 80 observations is sufficient to ensure proper interpretation of the t-statistic. This is the case in my study with its 166 observations. For the second generations choice I find the error terms to be normally distributed, the Jarque-Bera statistic cannot reject the normality hypothesis, the p-value is 0.992.

Note, however, that in dictator games the subjects' decision process may be divided into two steps. In a first step, the candidates have to make a discrete choice: Do I pass money on or not? Once they have decided to pass on some money, subjects decide in a second step how *much* to pass on. It may therefore be the case that the factors determining the first decision are different from those determining the second one. OLS fails to differentiate between the impacts of the two sub-decisions and may therefore generate a biased estimate. Accordingly, I also employ the Tobit regression method which accounts for this problem.

Chapter 4

Experimental design

The experiments were conducted from October 2000 to January 2001 in Heidelberg with 117 participants in 9 sessions, and from May 2002 to August 2007 in Kiel with 572 participants in 54 sessions. The sessions in Heidelberg were carried out by Till Requate, the ones in Kiel by me. Each subject was allowed to participate only once. The basic design has already been briefly sketched in the introduction. Depending on the treatment, the experiments lasted between 35 to 90 minutes, and average payoff ranged between 13€ and 15.5 €, including a show-up fee of 3€. The show-up fee is in line with comparable experiments, e.g. ‘random exchange’ by HMSS also had a show-up fee of 3\$, though the standard experiment by FHSS had a show-up fee of 5\$ (see section 11.3.2 in the appendix). The instructions for the experiments are attached in the appendix.¹ The subjects for the experiment were recruited by posting flyers in student cafeterias. Accordingly, there was a wide range of study subjects among the participants. In each session, 9, 12, or 15 participants were gathered in a seminar room and asked to read the instructions. After 5 to 10 minutes (depending on the treatment), the experimenter answered questions concerning the experimental procedure. Then subjects were randomly assigned to groups of three, consisting of members *A*, *B*, and *C*, except for the replication treatment of the traditional dictator game. The position of being *A*, *B*, or *C* was also assigned randomly. For treatment 3, after random division into groups, subjects were asked within each group whether they knew each other. If this was the case, they were re-matched in order to preclude any prior relationship influencing the subjects’ behavior.

The basic game common to each of the four treatments was what I call the *intergenerational three-person dictator game*. The experimental set-up is

¹The instructions were in German, the attached version is a translation. Note that the German instructions were gender-neutral.

outlined in Figure 4.1.

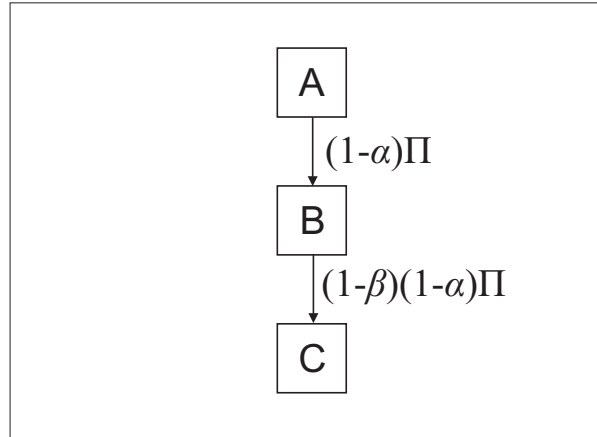


Figure 4.1: Basic experimental setup

Subject A received an amount Π of 30€ (I did not use experimental currency units). This pie size is rather large compared to other experiments where the pie-size was usually 5\$ or 10\$ (see section 11.3.2 in the appendix). In this setup there is a "third generation" hence pie-size should be increased by 50%. However, this would mean a pie-size of max. 15€. The use of 30€ instead was necessary to create a sufficient incentive for subjects to participate since some of the treatments lasted up to 90 minutes while the experiment by e.g. FHSS lasted 20 minutes. Since former experiments showed that pie-size alterations from 5\$ to 10\$ had no influence on subjects allocation behavior (see FHSS, section 2.2.2), a variation of pie the size to the extent introduced here should not matter, however this cannot be excluded completely.

From the pie of 30€ A had to choose a share α that she kept for herself, passing the rest $(1 - \alpha)\Pi$ on to B . Then B divided the rest $(1 - \alpha)\Pi$ between B and C , keeping a share β for herself and passing $(1 - \beta)(1 - \alpha)\Pi$ to C . All subjects knew the rules of the game. In particular, A knew that B had to divide the amount $(1 - \alpha)\Pi$ between herself and C .

This basic set-up was now divided into four different treatments. Treatments 1 and 4 ensured anonymity between A , B , and C , while in treatment 2 and 3 I introduced some social interaction between subjects A and B and between B and C , respectively.

In treatment 1 (anonymity), subjects A were taken into separate rooms each and asked to divide the initial pie Π between A and B . This division was done in writing on a form (Form 1, see appendix²), the money being paid out later. During their decision the subjects were left alone in the room. In addition,

²The forms were also in German, the attached version is a translation.

personal data (gender, study subject, income, and age) were requested and entered on the form. Then, Form 1 was collected, and subjects A received a second form (containing parts of Form 2, see appendix) in which they were asked about their motivation concerning the division of the pie and about their knowledge of the concept of sustainability. Afterwards, subjects A received their payoff, i.e. the amount $\alpha\Pi$ plus the show-up fee, and left without meeting subjects B or C , who were waiting in separate rooms, one for subjects B , one for those of type C . Then subjects B were taken into separate rooms for each, where the experimenter informed them how much of the initial pie Π had been passed on to them by subject A . Subjects B were then asked to divide the remaining amount $(1 - \alpha)\Pi$ between themselves and subjects C by filling out Form 1 that was adapted to B . When making their decisions, subjects B were left alone in their rooms. After Form 1 had been collected, subjects B were asked to fill out Form 2, where they had to answer the same questions as subjects A before. In addition, subjects B were asked which amount of money they had expected to be passed on by subjects A . Then subjects B were paid off, i.e. they received the amount $\beta(1 - \alpha)\Pi$ plus the show-up fee, and left. Finally, each subjects C were informed about the amount of money $(1 - \beta)(1 - \alpha)\Pi$ passed on to them and were then asked to fill out Form 2 (adapted to C). Then subjects C were paid off and left. Anonymity was ensured by the fact that A did not meet B and C directly, and B did not meet C after the treatment had started. The subjects met only briefly in one room while reading the instructions, being divided thereafter into groups and assigned to type A to C (additionally, for treatment 3 I ensured that subjects within one group did not know each other). I decided to choose a lower degree of anonymity than in the treatments by FHSS and HMSS (who put allocators and recipients into different rooms from the outset), because Frohlich *et al.* (2001) have shown that subjects may doubt the existence of recipients if there are no clues about them actually participating in the experiment.

Treatment 2 resembled treatment 1, except that A and B socially interacted after A had divided the pie. Again, subjects A were first taken to separate rooms and were asked to divide the pie Π between themselves, with a share to be passed on to subject B . Alone in the room, they filled out Form 1. After subjects A had completed the form, B was taken into the same room and was informed about the amount of money being passed on to B by A . Then A and B were asked to play a simple form of scrabble, where they cooperatively tried to gain as many points as possible by forming German words from the scrabble letters in a period of 15 minutes, following the usual rules of the scrabble game, i.e. assigning points to the letters they had managed to use. In this game they

could earn up to 3€, depending on their success. The resulting payoff from the scrabble game was split equally between A and B . The same procedure was repeated for B and C , who also played the scrabble game after B had made their decisions and filled out their forms.

Treatment 3 was similar to treatment 2, except that the choice on the allocation of the pie was made *after* playing the scrabble game. After the scrabble game was finished, subject B (or C) left the room, and A (or B) were left alone in the room while making their decisions and filling out their forms.

Treatment 4 was similar to treatment 1, except that B and C did not know whether the division of pie Π had been made by subject A or by a random mechanism. By tossing a coin (in the presence of A), I decided whether the division would be done by A or by the random mechanism. If A was to make the choice, the experiment continued as in treatment 1. If the random mechanism was selected for division, a 20-sided dice was used (again, in the presence of A) to determine the amount of money for A . Subject A could receive integer values between 10€ and 30€. Since there are 21 integer numbers from 10 to 30, I omitted the number 11 as a possible outcome so as to obtain 20 possible divisions that could be assigned to A . The intention was to keep the division (10,20) frequently chosen by A in the set of possible allocations. I excluded the values 0 through 9 since it was unlikely that A would take less than one third of the pie. Since the random process yielded only integer allocations, A (if A was allocating) was only allowed to choose integer allocations too, i.e. A was not allowed to pass on, say, 12.5 euros.

Finally, I also ran a replication of the traditional dictator game (in the following referred to as ‘TDG’) to ensure that my results were comparable to other studies. Here, the set-up was similar to treatment 1, except that the pie had only to be divided between subjects A and B .

Table 4.1 summarizes the the main features of the different treatments.

Table 4.1: Treatment overview

Treatment	Social interaction	Subject A
1	None	Human
2	After decision	Human
3	Before decision	Human
4	None	50% human, 50% random
TDG	None	Human

Chapter 5

Analysis of the first generation's allocation choice

In this chapter, I analyze the decisions of subjects A , the first generation, and compare the impact of the different treatments. I start with descriptive statistics, then I test for differences between treatments. Unlike in most of the studies, I do not only test for treatment differences of the entire sample but also for treatment differences, where the subjects taking the whole pie are excluded. Finally, I investigate factors determining A 's decision by running regressions using the asked personal data.

5.1 Treatment comparison entire sample

5.1.1 Descriptive analysis

Table 5.1 lists the number of observations for the different treatments and the means of the shares passed on by A .

Table 5.1: Descriptive results allocation choice A

Treatment	Number obs.	Mean share of A s passed on by A
1A	54	0.34
2A	40	0.48
3A	39	0.49
4A	33	0.33
TDG	40	0.26
FHSS	69	0.23
HMSS	24	0.27

The number of observations is in line with comparable experiments, e.g.

FHSS had 24 to 46 observations, HMSS had 24 to 41 observations per treatment (see section 11.3.2 in the appendix).¹

In treatments 1 and 4, subject *A* passed on an average of one third of the initial pie. Social interaction increased the average of shares passed on by *A* to about one half. In ‘TDG’, subjects *A* passed on a share of 0.26 on average, which is in line with the results by FHSS and HMSS, who observed mean shares of 0.23 and 0.27, respectively. Thus, in two-person dictator game treatments a lower mean share was passed on than in treatment 1 with its mean share of 0.34.

The results from the four treatments are displayed as histograms in Figure 5.1.

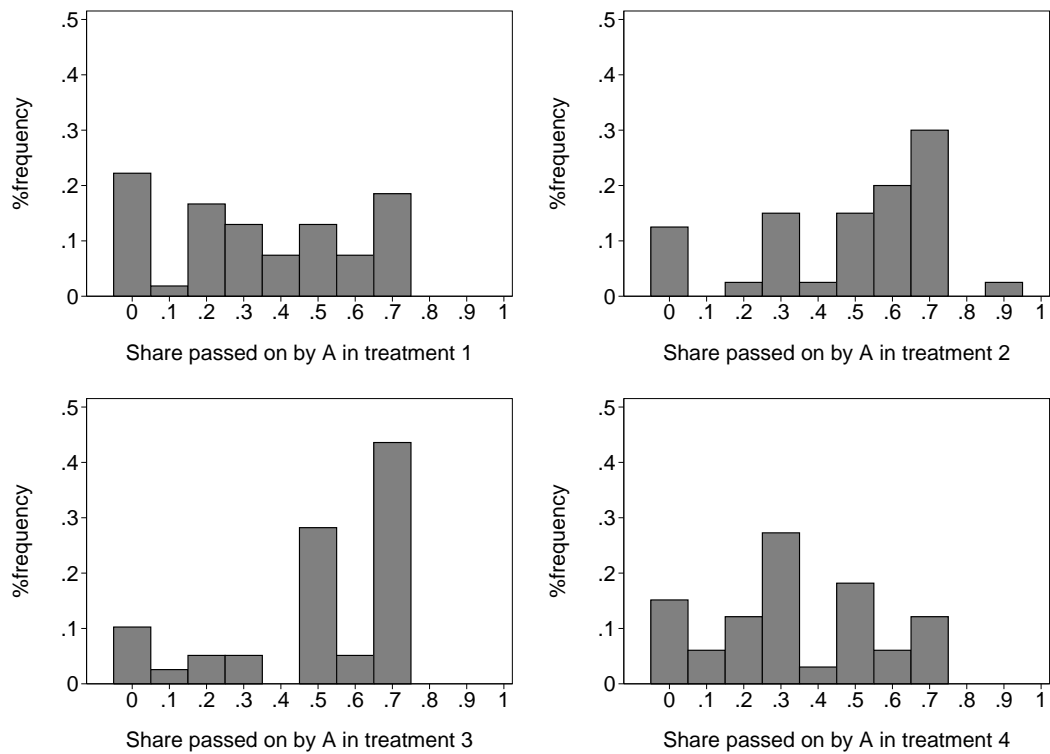


Figure 5.1: Frequency distribution of shares passed on by *A*, treatments 1-4

Here I see that in the treatments *with* social interaction the frequency of shares passed on by subject *A* peak at 0.7, while *without* social interaction the frequency of those shares is rather uniformly distributed between zero and 0.7. Note that the 0.7 column represents the subjects passing on two thirds of the initial endowment, which is the equal share.

Interestingly, in treatment 2 there is one case of *A* passing on 26€, which

¹The number of 69 observations in Table 5.1 stems from pooling treatments with 5\$ and 10\$ pie-size.

was substantially more than two thirds of the pie. This is surprising, since subject A 's choice increases social inequity in favor of B and C . Since subjects had to explain their division (in Form 2a), it turned out that this subject had understood the rules but deliberately chose this allocation. She claimed that her monthly income (which was around 350 to 500€ after having paid housing rent) was sufficient for living, and therefore she had no desire to earn additional money but had just participated for the fun of it. Since in her perception the other subjects' participation was mainly motivated by monetary considerations, she argued that it would be welfare-maximizing to pass on more money than the equal split. This case favors the social welfare-preference model by both Andreoni and Miller (2002) and Charness and Rabin (2002), in contrast to the difference-aversion models by both Fehr and Schmidt (1999) and Bolton and Ockenfels (2000). However, in my experiment, this kind of behavior was exceptional.

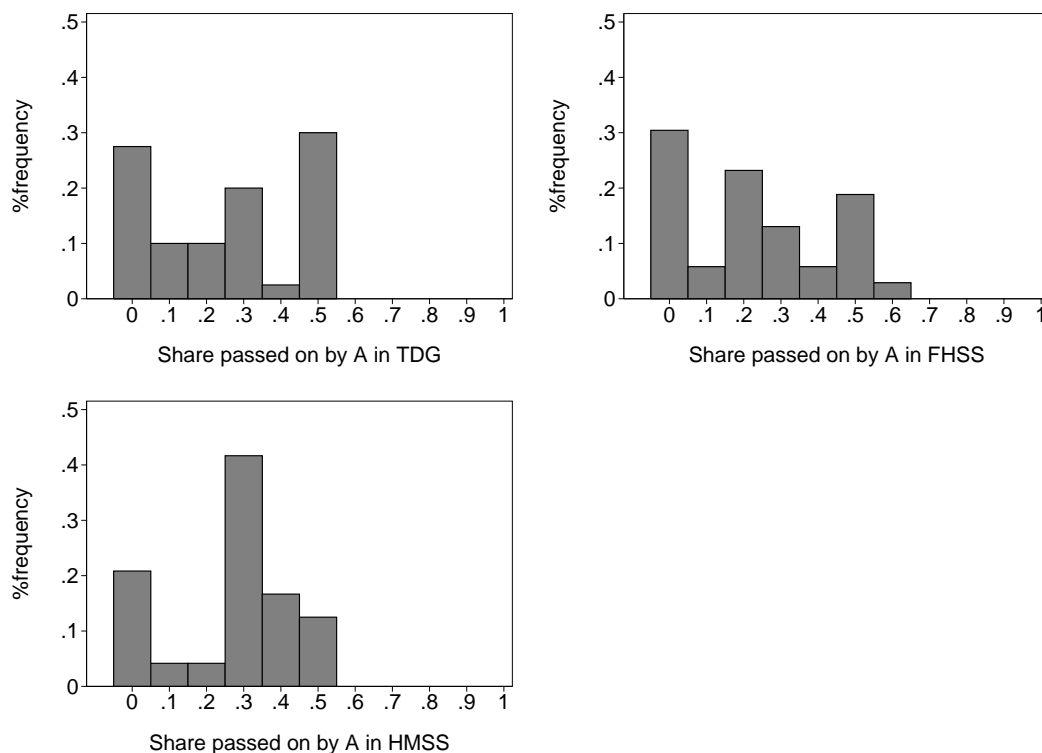


Figure 5.2: Frequency distribution of shares passed on by A , TDG (replication) vs. FHSS and HMSS

I also compared my results from ‘TDG’ with the traditional dictator game by FHSS² and HMSS.³ The results displayed in Figure 5.2 were roughly in line

²The sample for FHSS consists of a pooled sample combining the ‘dictator pay’ treatments for both the pie sizes with 5\$ and the one with 10\$.

³The sample for HMSS consists of the sample from the ‘Dictator random entitlement’ treatment, whose set-up was largely similar to the one by FHSS.

with those from FHSS and HMSS, however, the share of subjects allocating equally is higher in my set-up (though not significantly). In the investigations by FHSS and HMSS, subjects never passed on a share exceeding 0.5 (except for two out of 69 subjects in FHSS), while in treatment 1 of the three-person dictator game almost 30% of subjects A passed on shares larger than 0.5, but not exceeding 0.66. This is plausible, since a reference point is the fair division of one third to everybody. Although it is unlikely, I cannot rule out that some differences in distribution may have stemmed from the different pie-size used in my experiment compared to FHSS and HMSS.

5.1.2 Comparison of distributions

Testing for differences in the treatments, I first compare the outcomes of treatments 1-4. My main focus is on the impact of social interaction. Then I test for differences between the three-person dictator games and the traditional dictator games. I also statistically compare the results from ‘TDG’ to the outcomes of FHSS and HMSS. Finally, I investigate the influence of gender and other socio-economic factors.

Testing for differences between treatments

I now compare the outcomes from the treatments 1 through 4. In Table 5.2, I report the results from using the Mann-Whitney test (MW), the Anderson-Darling test (AD), and the Epps-Singleton test (ES) of testing the null hypothesis stating that the outcomes of the different treatments are equal. The table lists the corresponding test statistics and the respective p-values in parenthesis beneath (the same holds for the remaining tables reporting treatment comparisons).

My main findings from the comparisons are as follows: If I compare the outcome of a treatment without social interaction (i.e. 1 or 4) to the outcome of a treatment with social interaction (i.e. 2 or 3), in other words 1A vs. 2A, 1A vs. 3A, 2A vs. 4A, and 3A vs. 4A, I find that social interaction (either before or after the allocation choice) significantly increases generosity.

Comparing the two treatments without social interaction, the choices of subject A are not significantly different in treatments 1 and 4. Thus, introducing the possibility of a random element in the choice of A apparently does not influence the decision of subject A in any significant way. Comparing the two treatments with social interaction (treatments 2 and 3) I find that the outcomes do not differ significantly. Thus, given that social interaction takes place, it does not seem to play a major role for A ’s choice whether social in-

teraction takes place before or after A has to divide the pie between herself and subjects B and C .

Table 5.2: Treatment comparison A

	Treatments compared	MW	AD	ES
The effect of social interaction	1A vs 2A	-2.549 (0.011)	3.544 (0.014)	10.107 (0.039)
	1A vs 3A	-2.827 (0.005)	3.891 (0.010)	10.658 (0.031)
	2A vs 4A	2.934 (0.003)	4.184 (0.007)	11.795 (0.019)
	3A vs 4A	3.188 (0.001)	4.716 (0.004)	14.420 (0.006)
Possibility of a non-human allocator	1A vs 4A	0.295 (0.768)	0.343 (0.913)	1.599 (0.809)
Social interaction before or after choice	2A vs 3A	-0.384 (0.701)	0.785 (0.494)	1.559 (0.816)
Extension to three persons	1A vs TDG	1.850 (0.064)	2.921 (0.028)	12.331 (0.015)
	1A vs FHSS	-2.686 (0.007)	5.150 (0.002)	15.883 (0.003)
	1A vs HMSS	1.455 (0.146)	2.710 (0.038)	17.093 (0.002)
Comparison to standard setups	TDG vs FHSS	-0.751 (0.453)	0.985 (0.368)	5.220 (0.266)
	TDG vs HMSS	0.261 (0.794)	0.972 (0.375)	8.117 (0.087)

Nevertheless, it might be the case that the motivation to pass on a larger share compared to the treatments without social interaction is quite different in treatment 2 compared to treatment 3. In treatment 2, where A decides before social interaction, there are two possible explanations why A passes on more than in the anonymous dictator games in treatments 1 and 4. The first may be that A would like to give B an incentive to cooperate. If B feels unfairly treated by A , B might want to punish A by refusing to cooperate in the scrabble game and hence reciprocate according to the model suggested by Rabin (1993). Even if this behavior was at B 's own expense (since B earns half of the payoff from the scrabble game), this explanation would be consistent with results from previous experiments. In ultimatum games, e.g. in Güth *et al.* (1982, see section 2.1.1), subjects revealed a strong desire to reciprocate despite incurring costs by punishing someone they had perceived as behaving unfairly. In my set-up, the impact of this reasoning should not be too large, since the possible loss for A (through punishment by B) is rather low. Playing

scrabble alone would not reduce a subject's earnings in a dramatic way. If I additionally compare the rather low maximum possible payoff of 3 Euros in the scrabble game to the pie of 30€, the influence of the willingness to provide incentives to cooperate should not be too strong. Still, 10% of the subjects in treatment 2 explained their choice by the incentive argument, while 7.5% justified the low amounts they had passed on to *B* (and *C*) by indicating the negligible impact of cooperation due to the small amount of money at stake.

An alternative explanation for *A*'s generosity in treatment 2 might be as follows: If subject *A* treats subject *B* non-nicely by being greedy, *A* may have to undergo an awkward situation when being forced to spend 15 minutes with *B* in the same room working on a joint task. *A* might want to avoid this by being (relatively) generous. However, with the exception of one subject, no participant in treatment 2 explained her behavior with this argument. However, when looking at the subjects' reasoning in treatments 1 and 4, I observe that 20% and 12% of subjects *A* in treatments 1 and 4, respectively, justified their behavior by referring to anonymity. This enables us to reverse the argument: If subjects name anonymity as justifying greedy behavior, observing generosity in treatment 2 implies that social interaction is important.

Considering treatment 3, there may also be two effects that increase generosity. Firstly, subjects *A* may feel obliged to reward *B* for successful cooperation. However, only 5% of the subjects participating in treatment 3 used this argument (from an incentive point of view there would be no reason for this behavior). Secondly, subjects got to know each other during the scrabble game that induced increased affection between *A* and *B*. In fact, only one subject mentions this motive. There is nevertheless indirect evidence for the existence of the second effect. It is the same argument as for treatment 2, i.e. the high share of subjects mentioning anonymity in treatments 1 and 4.

There may be a third effect that plays a role in both treatments 2 and 3 which, in contrast to the effects described previously, reduces the amount that *A* passes on to *B*. Since *B* plays the scrabble game twice, while *A* and *C* play only once, *A* might account for this and pass on less than one third to *B* for reasons of equity. This argument would be supported by the results of both Fehr and Schmidt (1999) and Bolton and Ockenfels (2000). Actually, 5% of the subjects *A* taking part in treatment 2 and 18% of those taking part in treatment 3 put forward this argument to explain their decision.

I summarize these findings as follows:

Result 1: In the three-person dictator game as described above,
i) social interaction (either before or after the allocation choice) significantly increases generosity (amounts passed on by *A* are significantly higher in treat-

- ments 2 and 3 than in treatments 1 and 4);
- ii) introduction of a random element does not influence subject A 's choice in a significant way (the outcomes of treatments 1 vs. 4 are not significantly different with respect to the choice by subject A);
 - iii) the order of social interaction and division choice does not induce significantly different outcomes with respect to the choice by subject A (the outcomes of treatments 2 vs. 3 are not significantly different with respect to the choice of subject A).

Three-person intergenerational dictator game vs. traditional dictator game

Comparing the results from the intergenerational (or sequential) dictator game with the results from the traditional standard two-person dictator games, I find that there is a significant difference between my results and the results obtained by FHSS for all test statistics. Comparing my results to HMSS, I obtain significant differences for both the AD and the ES test statistic, while the MW test statistic is insignificant, which possibly has to do with the high number of ties in the sample by HMSS. So when people decide how much to keep for themselves, it obviously does matter whether they are sharing a pie with one or with several people.

I summarize my findings as follows:

Result 2: In the three-person sequential dictator game, the first dictator passes on more than the dictator in the traditional dictator games.

Note that these results contrast with the findings by Güth and van Damme (1998), who conduct an ultimatum game where allocations are made by a proposer to a respondent and a passive third person. The passive person, however, in contrast to the responder, receives only a very small share of the pie. An explanation might be the claim by Bardsley (2008) and List (2007) that subjects respond to options and infer from these to the behavior they believe is expected by them and tend to conform to these expectations. Being confronted with a setup where one party is powerless might induce subjects to infer that they are supposed to neglect this party.

Comparing the results from the replication TDG with the results from the traditional dictator game obtained by FHSS and HMSS yields no significant difference.⁴ Obviously, the slight differences in instructions, procedure and pie-size do not matter.

⁴Note that with a p-value of 0.087 the Epps-Singleton statistic for the comparison of my replication and HMSS is close to being significant.

The Impact of gender

Since experimental studies, such as those by Eckel and Grossman (1998) and Andreoni and Vesterlund (2001) often find significant differences in behavior with respect to gender, I compared male and female decisions in the samples. The descriptive results are summarized in Table 5.3.

Table 5.3: Descriptive results allocation choice A with regard to gender

Treatment	Mean share passed on by males	Obs males	Mean share passed on by females	Obs females
1A	0.31	37	0.40	17
2A	0.43	29	0.61	11
3A	0.46	28	0.56	11
4A	0.30	17	0.36	16
TDG	0.21	24	0.33	16

At first glance, males seem to pass on less than women in each treatment, especially in treatment 2. When testing for gender differences *within* the treatments, it can be seen, however, that this difference is only significant in treatment 2 (see Table 5.4). A possible explanation why females are more generous

Table 5.4: Testing for the impact of gender on A 's choice within treatments

Treatment	MW	AD	ES
1A	-1.145 (0.252)	1.039 (0.340)	2.492 (0.646)
2A	-2.266 (0.023)	3.227 (0.019)	5.3451 (0.254)
3A	-1.119 (0.263)	1.034 (0.342)	2.444 (0.655)
4A	-0.871 (0.384)	0.815 (0.476)	2.831 (0.587)
TDG	-1.514 (0.130)	1.536 (0.167)	5.817 (0.213)

in treatment 2 is that females may feel more uncomfortable than men about being in a room and cooperating with someone they had treated "unfairly." In fact, however, only one female subject justified her generosity by putting forward this argument.

I also checked for treatment differences within genders (see Table 5.5). Some differences significant for the overall sample are no longer significant for some 'uni-gender' subsamples. Notably, the Epps-Singleton hardly detects any

difference between treatments, which may be a result of the smaller sample size. Employing the Mann-Whitney and the Anderson-Darling tests, I do find differences between genders. First, for males the results from treatments 1

Table 5.5: Treatment comparison *A* for ‘uni-gender’ subsamples

	Treatments compared	MW	AD	ES
The effect of social interaction	Males 1A vs 2A	-1.704 (0.088)	1.734 (0.126)	5.271 (0.261)
	Females 1A vs 2A	-2.347 (0.019)	3.340 (0.015)	5.994 (0.200)
	Males 1A vs 3A	-2.209 (0.027)	3.027 (0.024)	6.821 (0.146)
	Females 1A vs 3A	-1.914 (0.056)	2.786 (0.031)	3.987 (0.408)
	Males 2A vs 4A	1.808 (0.071)	1.879 (0.106)	5.030 (0.284)
	Females 2A vs 4A	-3.118 (0.002)	5.810 (0.001)	10.830 (0.029)
	Males 3A vs 4A	2.175 (0.030)	2.627 (0.041)	7.416 (0.116)
	Females 3A vs 4A	-2.693 (0.007)	4.031 (0.006)	7.849 (0.097)
Possibility of a non-human allocator	Males 1A vs 4A	0.216 (0.829)	0.218 (0.991)	0.417 (0.981)
	Females 1A vs 4A	0.617 (0.537)	0.798 (0.491)	1.334 (0.856)
Social interaction before or after choice	Males 2A vs 3A	-0.686 (0.493)	0.654 (0.606)	0.839 (0.933)
	Females 2A vs 3A	-0.493 (0.622)	0.388 (0.892)	1.851 (0.763)
Extension to three persons	Males 1A vs TDG	1.615 (0.106)	2.133 (0.076)	7.462 (0.113)
	Females 1A vs TDG	1.147 (0.251)	1.329 (0.220)	8.035 (0.090)

vs. 2 and those from treatments 2 vs. 4 are no longer significantly different, whereas the differences between both treatments 1 vs. 3 and treatments 3 vs. 4 prevail. Obviously, males care about social interaction if it takes places after the allocation decision has been made, while social interaction seems to be less important if it happens before dividing the pie. There are two possible explanations why males behave less generously in treatment 2 than in treatment 3. Either they simply do not anticipate the uncomfortable situation of being together with someone they have not treated nicely, or they simply do not care. Females, by contrast, are sensitive to social interaction regardless of

when it takes place (though, possibly due to the small sample size, the comparison of treatments 1 vs. 3 is slightly above the 0.05 p-value from the MW test). Secondly, for both subsamples I again observe no significant differences between the different treatments *without* social interaction (i.e. 1 vs. 4) and those *with* (i.e. treatments 2 vs. 3). Thirdly, the significant difference between the intergenerational (or sequential) dictator game and the traditional standard two-person dictator games vanishes in uni-gender subsamples.

Besides sex, I controlled for age⁵, income, and the subjects studied the by participants. Using Spearman’s rank correlation coefficient to test for the impact of age, I find no significant correlation between the share taken by subjects *A* and their age in treatments 1, 3, and 4. In treatment 2, by contrast, I find a significantly positive relationship between the two variables. Obviously younger subjects are more fearful than older subjects of the prospect of spending 15 minutes in a room with someone they have not treated fairly and therefore pass on a higher share.

Controlling for income, I find a significantly negative relation between income and the share passed on by *A* only in treatment 1. Interestingly, in treatment 4, which has the same incentive structure as treatment 1, I do not find this relation. The same holds for treatment 2 and 3.

I do find any significant correlation between study subjects by the participants and share passed on by *A* for all treatments.

5.1.3 Comparison of proportions passing zero

Taking a look at the proportions of subjects passing on zero we find that the share of subjects passing nothing is lower in the treatments with social interaction (see Table 5.6). However, the difference is not substantial.

Table 5.6: Proportion of subjects *A* passing zero

Treatment	Proportion passing zero
1A	0.19
2A	0.13
3A	0.10
4A	0.15
TDG	0.25
FHSS	0.30
HMSS	0.21

Comparing proportions of the different treatments using Fisher’s exact test

⁵Carrying out an experimental study on contributions to a charity, List (2004) found that age did matter.

does not yield any significant differences between treatments, see Table 5.7. Even pooling treatment 1 and 4, and 2 and 3 respectively (there are no signifi-

Table 5.7: Comparison of proportion of subjects *A* passing zero

Treatments compared	P-value from Fisher exact test
1A vs 2A	0.571
1A vs 3A	0.381
1A vs 4A	0.776
2A vs 3A	1
2A vs 4A	0.747
3A vs 4A	0.723
14A vs 23A	0.378
1A vs TDG	0.623
1A vs FHSS	0.537
1A vs HMSS	1
TDG vs FHSS	0.661
TDG vs HMSS	0.769

cant differences between these treatments) and then comparing the two pooled treatments yields a test statistic far from being significant. It seems as if the decision to pass on zero is independent from the treatments. An explanation may be that only subjects that are already willing to pass on a positive amount seem to respond stronger to the different treatments.

Comparing proportions from the intergenerational (or sequential) dictator game with the results from the traditional standard two-person dictator games I do not find any significant differences. Again, subjects only seem to respond to the added third person given they pass on positive amounts.

Also comparing the proportions from the replication TDG with the results from the traditional dictator game obtained by FHSS and HMSS yields no significant difference.

When comparing proportions of males and females passing on zero, I find that in all treatments a higher share of men keeps the entire pie (see Table 5.8).

However, there are no significant differences between males and females for all treatments (see Table 5.9). Note, that comparing pooled samples for treatment 1 and 4 (14A) and for treatment 2 and 3 (23A) between males and females yields marginally or close to marginally significant results. Hence, sample size may matter, here.

Table 5.8: Proportion of male and female *As* passing zero

Treatment	Male: Proportion passing zero	Female: Proportion passing zero
1A	0.24	0.06
2A	0.17	0
3A	0.14	0
4A	0.24	0.06
TDG	0.29	0.19

Table 5.9: Comparison of proportion of male and female *As* passing zero within treatments

Treatment	P-value Fisher exact test
1A	0.259
2A	0.313
3A	0.558
4A	0.355
TDG	0.722
14A	0.080
23A	0.105

5.2 Treatment comparison, including only *As* passing on a positive amount

In this section, I exclude the observations where subjects took the whole pie. Of course the mean average passed on in all treatments is lower, now. The basic descriptive results of comparing means remain the same, social interaction increases generosity substantially (see Table 5.10).

Table 5.10: Descriptive results allocation choice *A* excluding observations where *A* passed zero

Treatment	Number observations	Mean share <i>A</i>
1	44	0.42
2	35	0.55
3	35	0.54
4	28	0.39
TDG	30	0.34
FHSS	48	0.32
HMSS	19	0.34

When comparing the results for the distribution of the treatments excluding the subjects taking all, I find no substantial differences compared to using the entire sample (see Table 5.11 compared to Table 5.2 in section 5.1.2). There are

Table 5.11: Treatment comparison A excluding observations where A passed zero

	Treatments compared	MW	AD	ES
The effect of social interaction	1A vs 2A	-2.650 (0.008)	4.118 (0.007)	10.086 (0.039)
	1A vs 3A	-2.732 (0.006)	4.981 (0.002)	9.377 (0.052)
	2A vs 4A	3.429 (0.001)	5.548 (0.001)	12.377 (0.015)
	3A vs 4A	3.485 (0.001)	5.419 (0.001)	14.080 (0.007)
Possibility of a non-human allocator	1A vs 4A	0.739 (0.460)	0.525 (0.730)	1.464 (0.833)
Social interaction before or after choice	2A vs 3A	0.251 (0.802)	0.602 (0.656)	2.471 (0.650)
Extension to three persons	1A vs TDG	1.887 (0.059)	2.828 (0.032)	12.521 (0.014)
	1A vs FHSS	-2.357 (0.018)	4.610 (0.004)	17.278 (0.002)
	1A vs HMSS	1.785 (0.074)	3.523 (0.013)	18.074 (0.001)
Comparison to standard setups	TDG vs FHSS	-0.441 (0.659)	1.284 (0.238)	6.787 (0.148)
	TDG vs HMSS	0.808 (0.419)	1.494 (0.177)	8.099 (0.088)

minor changes in statistics and p-values but all the basic results are sustained. It is interesting that some results are even more pronounced while some are less. Here, there are two effects at play: The sample size for the comparisons excluding the subjects passing a zero amount is reduced hence reducing the statistics and increasing the p-values. But if the share of subjects passing on nothing is rather constant over treatments, the results should be more pronounced, working in the opposite direction.

Excluding the observations from subjects who pass on zero does not change the results from comparing the intergenerational dictator game to the results from the traditional standard two-person dictator games by FHSS and HMSS.

The same holds for comparing my replication TDG to the results by FHSS and HMSS.

I now take a look the results for the impact of gender within the different treatments (see Table 5.12). Comparing these results to the results where the entire sample is considered (see Table 5.4 in section 5.1.2) reveals that the difference between men and women for treatment 2 is not significant any longer.

This may be interpreted in the way that the strongest difference between men and women is the share of subjects taking all, but the reduced sample size could also be the reason for this. The rest of the results are confirmed, but the change in p-values is relatively strong in treatment 1 increasing from 0.252 to 0.941 and also increasing in the other treatments.⁶

Table 5.12: Testing for the impact of gender on A 's allocation choice excluding observations where A passed zero

Treatment	MW	AD	ES
1A	-0.074 (0.941)	0.214 (0.993)	0.249 (0.993)
2A	-1.676 (0.094)	2.060 (0.082)	2.993 (0.559)
3A	-0.459 (0.646)	0.443 (0.826)	0.553 (0.968)
4A	0.023 (0.982)	0.504 (0.767)	1.056 (0.901)
TDG	-1.461 (0.144)	1.583 (0.153)	5.438 (0.245)

Then, I compare the results of the male and female subsamples, see Table 5.13 to the results where the entire sample is considered (see Table 5.5 in section 5.1.2). Again, we have some differences between the results from the entire sample and from the sample where observations passing zero are excluded. First, I investigate the male subsample. Comparing treatment 2 and 4 and comparing treatment 1a to TDG is now significant for MW and AD (p-value <0.05) while for the entire sample it was not. The rest of the results is sustained. For the women, there is no substantial change in results, having again some results being more, some being less pronounced.

⁶Note, that comparing results of the proportions of males and females taking all yields no significant difference, which would be evidence against this hypothesis, but the non-significant difference is possibly due to small sample size since the actual proportions differ substantially.

Table 5.13: Treatment comparison first generation

	Treatments compared	MW	AD	ES
The effect of social interaction	Male 1A vs 2A	-1.719 (0.086)	2.008 (0.089)	5.210 (0.267)
	Female 1A vs 2A	-2.191 (0.029)	2.982 (0.023)	5.074 (0.280)
	Male 1A vs 3A	-2.105 (0.035)	2.733 (0.036)	5.707 (0.222)
	Female 1A vs 3A	-1.734 (0.083)	2.437 (0.049)	3.187 (0.527)
	Male 2A vs 4A	-2.036 (0.042)	2.463 (0.049)	5.244 (0.263)
	Female 2A vs 4A	-2.995 (0.003)	5.424 (0.001)	9.807 (0.044)
	Male 3A vs 4A	2.267 (0.023)	2.748 (0.034)	6.550 (0.162)
	Female 3A vs 4A	-2.547 (0.011)	3.668 (0.009)	6.762 (0.149)
	Possibility of a non-human allocator	Male 1A vs 4A	0.395 (0.693)	0.249 (0.983)
Female 1A vs 4A		0.658 (0.511)	0.884 (0.431)	1.416 (0.841)
Social interaction before or after choice	Male 2A vs 3A	0.641 (0.522)	0.683 (0.582)	2.753 (0.600)
	Female 2A vs 3A	-0.493 (0.622)	0.388 (0.892)	1.851 (0.763)
Extension to three persons	Male 1A vs TDG	-2.044 (0.041)	2.858 (0.030)	7.566 (0.109)
	Female 1A vs TDG	0.623 (0.533)	1.168 (0.279)	7.636 (0.106)

5.3 Regression analysis of A 's choice

To shed more light on the factors determining the decisions of subjects A and also to elicit some quantitative effects, I now regress the share passed on by A , $(1 - \alpha)$, on gender, income, age, study subject, and knowledge about the concept of sustainability. Gender is coded by a dummy (0=female, 1=male). Income is divided into eight classes (since some students live with their parents, I asked about disposable income after having paid the rent) starting from less than 350€ and then increasing in steps of 150€ (350€ to 500€, 500€ to 650€, and so on). For the regression these classes are coded as 0,1,2,...,7. For some income data, corrections had to be made. These procedures are described in more detail in section 5.3.1. To capture the impact of the subjects studied, I split the participants into different groups. This procedure is described in more detail in subsection 5.3.2. Sustainability is captured by a dummy variable 'sust' taking the value of 1 if the subject gives an appropriate definition of sustainability and 0 otherwise. Treatments are also captured by a dummy variable for each of the treatments 2-4, leaving treatment 1 as baseline setting.

Note, that for very few observations, some categories were not given because e.g. a subject refused to state the income. Here, I used the mean of all other subjects.

5.3.1 Presentation of the different income adjustment methods

Initially (this is the case for 7 observations of treatment 2 and 32 observations of treatment 3) I asked the participants about their monthly income in general. Unfortunately, some of the students may live at their parents place (according to the 17th 'Sozialerhebung des Deutschen Studentenwerks' 2003, which is a survey about the socio-economic situation of students in Germany, this is the case for a share of 0.22). This implies that the income of students living at their parents' place is relatively higher than that of students who declared the same income but have to pay the rent in addition. Hence, I decided to ask for 'income after having paid the rent' for the rest of the study. Thus, I have to deal with the fact that there are two kinds of data. I take the observations where it was asked for 'income after having paid the rent' as benchmark since these answers should be more reliable and since I have roughly three times as many of these observation as of the other kind. I correct the data where I asked for 'income in general' using several methods. The first method to take the differences into account is adjustment method 1 which can be seen in

Table 5.14. Here, I simply replace all observations where I asked for income in general by the mean income of the other observations.

Table 5.14: Adjustment method 1

Income category on form 1	Income after having paid the rent	Corrected income
> 350	1	1.7234
350-500	2	1.7234
500-650	3	1.7234
650-800	4	1.7234
800-950	5	1.7234
950-1100	6	1.7234
>1100	7	1.7234

Adjustment method 2 is still rather simple. I just put every observation where I asked for income in general one income category lower compared to income after having paid rent. That means I use the adjustment scheme according to Table 5.15:

Table 5.15: Adjustment method 2

Income category on form 1	Income after having paid the rent	Corrected income
<350	1	0
350-500	2	1
500-650	3	2
650-800	4	3
800-950	5	4
950-1100	6	5
>1100	7	6

Adjustment method 3 follows a scheme where the same category on the questionnaire would be one category higher for the income in general than the one for the ‘income without having paid the rent’ but in a category that is lower than the next category for income in general. This scheme leads to Table 5.16.

Table 5.16: Adjustment method 3

Income category on form 1	Income after having paid the rent	Corrected income
<350	2	1
350-500	4	3
500-650	6	5
650-800	8	7
800-950	10	9
950-1100	12	11
>1100	14	13

Adjustment method 4 is more complicated than the other ones. I tried to extrapolate the relative share of the income that is devoted to rent and thereby calculate the mean values for the income categories. Therefore I used the 17th ‘Sozialerhebung des Deutschen Studentenwerks’ 2003. Actually, the relevant observations that have to be extrapolated are from the time October 2000 to January 2001, hence the 16th survey would be more appropriate since it is from 2000. However, this survey contained no data on the income-rent ratio, so I took the survey from 2003. This report states the average rent payments and the rent-income ratio for income quartiles. With the help of these two variables I calculate the average income of each quartile which can be seen in Table 5.17.

Table 5.17: Sozialerhebung des Deutschen Studentenwerks

Average income	Average rent	Average rent/income-ratio
515	197	0.382
665	229	0.344
798	258	0.323
1107	319	0.288

Then, I regress the average rent/income-ratio on income using OLS using the following regression equation⁷

$$\frac{\text{Average rent}}{\text{Income}} = \beta_0 + \beta_1 \text{Income} + \varepsilon.$$

The results of this regression can be seen in Table 5.18.

⁷I am aware that I have the independent variable included in the dependent variable.

Table 5.18: Extrapolating the average rent/income-ratio

Variable	Coefficient	P-value
β_0	0.452	0.002
β_1	-0.000153	0.021

Using β_0 and β_1 , I extrapolate the average rent/income-ratio for the different income categories for the people who filled out a form where I asked for the general income. Then, I inferred from the extrapolated average rent/income-ratio to the actual income without rent these participants should have. Note, that instead of values 1 to 8 for ‘income after having paid rent’ I used an assumed mean income with ad hoc estimations for the lowest and highest category (see Table 5.19).

Table 5.19: Extrapolating income without rent from the average rent/income-ratio

Income	Average rent/income-ratio	Income without rent
275	0.41	162.23
425	0.387	260.48
575	0.364	365.61
725	0.341	477.63
875	0.318	596.53
1025	0.295	722.31
1325	0.249	994.54

Thus, I have the following adjustment scheme for the corrected income (see Table 5.20).

Table 5.20: Adjustment method 4

Income category on form 1	Income after having paid the rent	Corrected income
<350	275	162.23
350-500	425	260.48
500-650	575	365.61
650-800	725	477.63
800-950	875	596.53
950-1100	1025	722.31
>1100	1325	994.54

5.3.2 Presentation of the categories for study subject

In the questionnaires participant were also asked about their study subject. Since there is a large range of study subjects, it is necessary to form categories. Table 5.21 indicates which subject was put into which category.

Table 5.21: Study subjects and categories

Subject	Category
Computer Science	Science
Physics	Science
Chemistry	Science
Biology	Science
Economics	Economics
Business	Economics
Languages	Humanities
Philosophy	Humanities
History	Humanities
Medical Science	Other
Psychology	Other
Sports	Other
Teacher	Other
Geography	Other
Agricultural Science	Other
Pharmacy	Other
Law	Other
Nutritional Science	Other
Other subjects	Other
Pupil	Other
Worker/Employee	Other

I chose "science" as category 1, "economics and business" as category 2, "humanities" as category 3 and "others" as category 4. The last group contained subjects that from my point of view did not really fit into the three other categories and the few participants who did not study.

5.3.3 Regression analysis - results

I have the following regression equation with the share passed on by subject A , $(1 - \alpha)$, as dependent variable. Note, that *subject* and *treatment* are multi-column vectors.

$$(1 - \alpha) = \beta_0 + \beta_1 \cdot gender + \beta_2 \cdot income + \beta_3 \cdot age \\ + \beta_4 \cdot subject + \beta_5 \cdot sust + \beta_6 \cdot treatment + \varepsilon$$

I compare the OLS regression results of all four income adjustment methods and it turns out that income adjustment method 1 seems to perform best, see Table 5.22.

Table 5.22: Comparison of regressions with different income adjustment methods - initial model

Income adjustment method	R^2	P-value income regressor
1	0.176	0.012
2	0.167	0.029
3	0.159	0.071
4	0.170	0.021

Hence, the results using income adjustment method 1 are used for the regression results in Table 5.23, p-values are in parenthesis beneath point estimates. The displayed values for a Tobit regression are discussed later in this section.

Table 5.23: Regression *A*, initial model. Dependent variable: Share passed on by *A*

Variable	OLS	Tobit
Constant	0.246 (0.041)	0.210 (0.115)
Gender	-0.129 (0.002)	-0.157 (0.001)
Income1	0.055 (0.011)	0.062 (0.009)
Age	0.003 (0.554)	0.003 (0.530)
Science	0.008 (0.873)	0.012 (0.829)
Economics	0.021 (0.645)	0.037 (0.471)
Humanities	-0.020 (0.774)	-0.021 (0.784)
Sustainability	0.008 (0.828)	0.009 (0.840)
Treatment 2	0.158 (0.002)	0.172 (0.002)
Treatment 3	0.162 (0.001)	0.177 (0.001)
Treatment 4	-0.014 (0.781)	-0.012 (0.832)

The treatment comparison from section 5.1 showed that there should be no

difference between treatment 2 and 3. A Wald-coefficient test confirms this evidence, treatment dummy 2 and treatment dummy 3 are not significantly different. I therefore replace both variables by a dummy for social interaction (SocInt), hence confirming the evidence from the former section. In order to select the appropriate model for my data I apply a procedure suggested by Herwartz (2007):

- Step 1: I regress a constant on my dependent variable.
- Step 2: The resulting residuals are now used as dependent variable.
- Step 3: I regress every variable alone and a constant on the residuals and choose the variable that has the highest LM statistic λ . $\lambda = TR^2$ where T denotes the number of observations and R^2 denotes the coefficient of determination.
- Step 4: If $\lambda > c_{1-\alpha}$, $c_{1-\alpha}$ being the $(1-\alpha)$ -quantile of a χ^2 -distribution ($c_{0.95} = 3.841146$), I restart at step 1, now regressing this variable and the constant on my initial dependent variable.
- Step 5: I repeat steps 1 to 5 until $\lambda < c_{1-\alpha}$

The results of an OLS-regression where I used this model selection procedure are reported in Table 5.24.

Table 5.24: Regression *A*, final model. Dependent variable: Share passed on by *A*

Variable	OLS	Tobit
Constant	0.305 (0.000)	0.285 (0.000)
Gender	-0.117 (0.002)	-0.141 (0.001)
Income1	0.057 (0.006)	0.064 (0.006)
SocInt	0.169 (0.000)	0.183 (0.000)

Again, I used income adjustment method 1 for the regression because comparing the different income modifications for the final OLS model shows that income adjustment method 1 performs best, see 5.25. My analysis reveals that gender has a significant effect. Females pass on around 11 percentage points more of the pie than males. Note further that female candidates were underrepresented in treatments 2 and 3. I had about 25% females in the pooled

Table 5.25: Comparison of regressions with different income adjustment methods - Final model

Method	R^2	P-value income
1	0.172	0.006
2	0.161	0.018
3	0.153	0.045
4	0.164	0.014

sample of treatments 2 and 3, compared to 31% and 48% females in treatments 1 and 4, respectively. These two facts explain the even stronger effect of social interaction in the regression analysis (significant at the 1%-level) increasing the share passed on by A by almost 17 percentage points while comparing simple treatment means yields a difference of only 14 percentage points. The reason for this is the non-weighted sample. Since women tend to be more generous, the social interaction effect is underestimated by comparing pure sample means if women are underrepresented in these samples.

Also, the regression analysis reveals that income plays a role. Its impact, however, is considerably smaller than that of both gender and social interaction. If income is increased by 150€ (the income difference between the different categories), subjects A pass on 4 percentage points more.

As is typical for regressions using dictator game data, the adjusted R-squared (0.16) is rather low. Thus personal data do matter, but their predictive power is not overwhelming. A White test rejects heteroscedasticity (p-value 0.094), hence in this respect OLS is a suitable tool.

I also employ the Tobit regression method that accounts for the two-stage nature of the decision process in dictator games (see section 3.6). However, the results summarized in Table 5.23 remain almost unchanged.⁸

⁸Note, that in a Tobit regression the regression coefficients can no longer be interpreted as elasticities.

Chapter 6

Analysis of the second generation's allocation choice

In this chapter, I analyze the decisions of subjects B , the second generation, with the same methodology as for the first generation. I start with descriptive statistics, then I test for differences between treatments. I also check for some treatment differences, where the subjects taking the whole pie are excluded. Finally, I investigate factors determining B 's decision by running regressions using the asked personal data.

6.1 Treatment comparison entire sample

6.1.1 Descriptive analysis

Table 6.1 reports the results for descriptive statistics of the different treatments. Treatments '4Bself' and '4Bdice' refer to the subsamples of treatment 4 where either candidates A or the random process via a 20-sided dice determines $(1 - \alpha)$, the amount to be passed on. Again, I find that in the

Table 6.1: Descriptive results allocation choice B

Treatment	Mean shares passed on by B	Obs
1B	0.29	44
2B	0.37	35
3B	0.39	35
4B	0.20	64
4Bself	0.19	28
4Bdice	0.20	36

treatments *with* social interaction subjects pass on more than in those without. Interestingly, in treatment 4, subjects B are substantially less generous

than in treatment 1, although the average share passed on by A was almost the same in both treatments and hence cannot explain the difference.

The results from the four treatments are displayed as histograms in Figure 6.1. Note that if subjects A pass on zero, there are no observations for subjects B . Hence we have fewer observations for subjects B compared to subjects A .

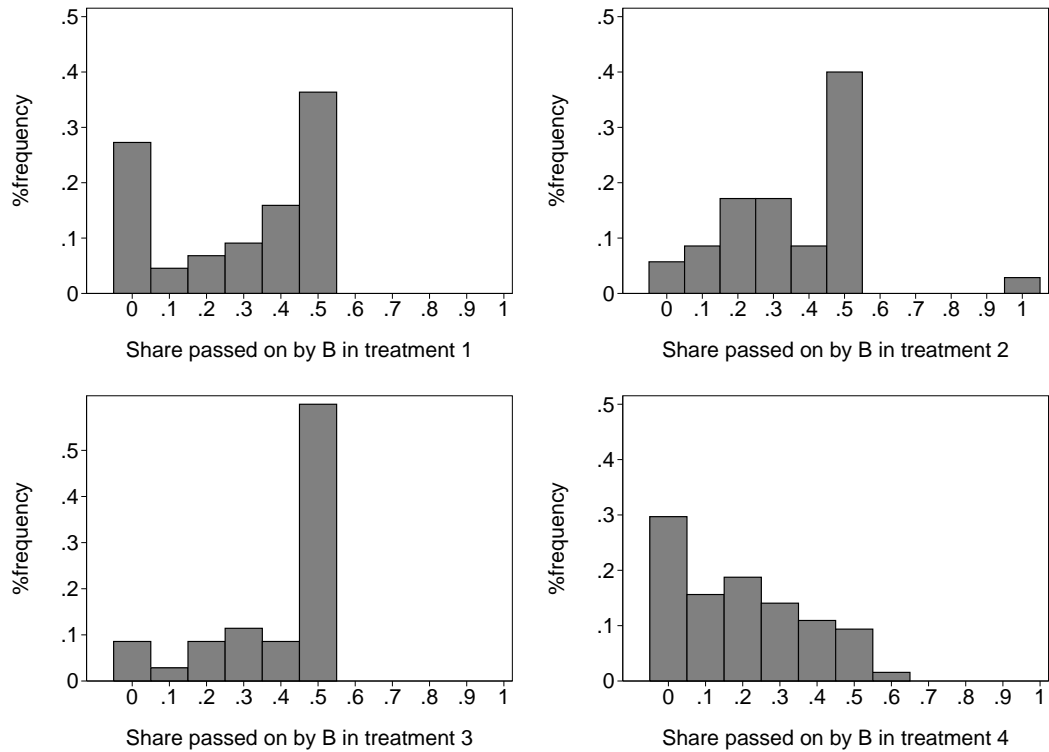


Figure 6.1: Frequency distribution of B 's shares, treatments 1-4

In treatments 1 through 3 the shares passed on from subjects B to C peak at shares of 0.5 of the remaining pie ($1 - \alpha$). In treatment 4, by contrast, the most frequent share passed on by subject B lies between 0 and 0.1. One subject from treatment 2 passed on the complete pie. Her explanation (that she trusted C) did not make sense to me. Another subject B from treatment 4 received 13€ from A and then passed on 8€ (=62%) to C . Her explanation was that 5 and 8 were her "lucky numbers." Interestingly, she chose to take the smaller amount for herself, but did not give any explanation for that.

6.1.2 Comparison of distributions

When analyzing the allocation behavior of subjects B there is an additional factor potentially influencing the choice made by B . Previous experimental evidence from, say, ultimatum games (see e.g. Güth et al., 1982) shows that subjects tend to reciprocate on A 's choice if they have the chance to do so.

Normally, reciprocity can be observed among the same subjects. Reciprocity then often follows a ‘tit-for-tat’ strategy: Once a person X has treated person Y nicely, Y reciprocates by treating person X nicely in return. In my case, however, subject B cannot apply this strategy to subject A but can only dictate the share allocated to a third party, here subject C . Even if C has not been involved in A ’s decision, B may want to reciprocate somehow and, therefore, reciprocates on C . Put differently, if A behaves generously (greedily) towards B , B may tend to be generous (greedy) to C .

I call this undirected reciprocity the *social history effect*. This effect could (and in fact does) have a strong impact on B ’s choice and, as a consequence, could make it impossible to distinguish between this effect and the treatment effect using treatment comparisons. Spearman’s rank correlation test strongly confirms the existence of the social history effect with a p-value of 0.000. Nonetheless, the results from treatment comparisons are presented in Table 6.2, since they show some interesting features.

Table 6.2: Treatment comparison B

	Treatments compared	MW	AD	ES
The effect of social interaction	1B vs 2B	-1.291 (0.197)	1.545 (0.167)	11.755 (0.019)
	1B vs 3B	-2.159 (0.031)	3.596 (0.013)	5.613 (0.230)
	2B vs 4B	-4.144 (0.000)	7.872 (0.000)	18.958 (0.001)
	3B vs 4B	-4.792 (0.000)	11.330 (0.000)	33.108 (0.000)
Possibility of a non-human allocator	1B vs 4B	-2.386 (0.017)	3.670 (0.012)	17.548 (0.002)
	4Bself vs 4Bdice	0.089 (0.929)	0.492 (0.766)	2.537 (0.638)
Social interaction before or after choice	2B vs 3B	1.032 (0.302)	1.129 (0.296)	4.914 (0.296)
B s choice vs. traditional dictator	1B vs TDG	0.700 (0.484)	1.100 (0.308)	4.036 (0.401)
	1B vs FHSS	-1.596 (0.111)	1.916 (0.100)	13.997 (0.007)
	1B vs HMSS	1.034 (0.301)	1.669 (0.139)	9.943 (0.041)

Similarly to the choice of A , I find that the outcomes from the treatments with social interaction, i.e. 2 and 3, differ from those without, i.e. treatments 1 and 4, the only exception is that treatment 1 and 2 do not differ any longer

for the MW- and AD-statistic. Note, that for comparing treatment 1 and 3, the ES statistic is not significant any longer, either. The differences between treatments are not necessarily triggered by the social interaction itself but could also be motivated by the social history effect. Again, it does not matter if social interaction takes place before or after the allocation choice.

When comparing the shares chosen by B to be passed on to C in treatment 1 (as the only recipient left) with the shares passed on by subject A in the replication treatment ‘TDG’, I do not find a significant difference. Hence, the second generation in a three-person dictator game does not behave significantly different from the allocator in a traditional dictator game. The same holds if I use the results by FHSS and HMSS instead of my replication TDG.¹

Unlike A ’s allocation behavior, I find that the outcomes of treatments 1 and 4 are now significantly different. This is rather surprising, since one would expect the decision set-up for type- B subjects to be not much different from that for type- A subjects. One might surmise that candidates B could somehow infer from the amount they observe whether the decision was taken by subject A or by the random process since subjects sometimes tend to choose prominent numbers and as a consequence will behave differently if the random process is operative. Looking at Table 6.3, I see that the pattern of shares chosen by A is not significantly different from the pattern generated by the dice. Note, however, that the MW and the AD statistics are close to being significant.

Table 6.3: Comparison of the subsamples: Subject vs. the dice in treatment 4a

Treatments compared	MW	AD	ES
4Aself vs 4Adice	-1.668 (0.095)	1.742 (0.127)	4.849 (0.303)

This evidence is further confirmed by comparing the outcomes of the subsamples for B ’s behavior in treatment 4, where either subject A (‘4Bself’) or the dice (‘4Bdice’) was responsible for the share passed on by the first generation to B . Here I do not find any significant difference (see Table ??). Hence in treatment 4, subjects B were not able to distinguish whether subject A or the random process was responsible for the division, or at least it had no bearing on subject B ’s decision.

So, there must be something else involved to explain the differences in B ’s behavior in treatments 1 and 4. The random process somehow seems to trigger

¹Employing the data from FHSS and HMMS (instead of my ‘TDG’) in this comparison, I find a significant difference for the ES statistic. I ignore this result, since comparing pooled data from FHSS to those of HMMS also yields a significant difference under the ES statistic.

different behavior. According to Malle and Knobe (1997), people distinguish between behavior they perceive to be intentional or unintentional. For this situation, this means that if subject B knew or presumed that subject A was responsible for the division of the initial pie, it would make sense to reciprocate on that action, whereas it would not make sense if B knew or presumed that the pie had been divided by a random process. Blount (1995) confirms the important role of attributing intentions. In a series of ultimatum games, the average minimum acceptable offer declined significantly if the proposer was replaced by a random process.

This explanation, however, is not fully able to explain the difference between the two treatments 1 and 4, because on average the reciprocity effects should cancel out. Some of the subjects A have been generous, some have been greedy, but the average share passed on by A was approximately equal in both treatments 1 and 4. Hence something else must be operate. In a series of one-shot sequential experiments (where reciprocity is possible) Charness and Rabin (2002) find that *friendly* behavior has almost no influence on the response of the second mover and thus almost no explanatory power for respondents' behavior, whereas different levels of *unfriendly* behavior are reciprocated by unfriendly behavior. Offerman (2002) confirms this evidence so far. Additionally, in his study, it matters to subjects whether an action is chosen intentionally by another subject or by nature. If a negative action is chosen intentionally, subjects typically reciprocate strongly, while remaining largely neutral in the face of a positive intentional action. Conducting another psychological experiment, Knobe (2006) shows that some subjects are more inclined to perceive an action by another person as intentional rather than as a random by-product, if that action has a bad rather than a good effect. Applying this evidence to my experiment, I can interpret the results in the following way: If, in treatment 4, subjects B receive a high share, they tend to ignore this possible generosity by A since the random process may have been responsible as well. After receiving a small share, however, B candidates are more likely to attribute this result to the greed of A and not to the random process and will then indirectly reciprocate negatively in their behavior to C .

This reasoning is in line with my data from the questionnaires, where subjects explained their allocation behavior. In treatment 1, 9.1% of the subjects explained their behavior by some kind of positive reciprocity. In treatment 4, by contrast, no one named positive reciprocity as an explanation, while 4.7% of the subjects explained their behavior with reference to negative reciprocity. Note that in treatment 1, 15.9% of the subjects referred to negative reciprocity to explain their behavior. Thus my results are in line with previous evidence

showing that negative reciprocity is stronger than positive reciprocity.

Now, I analyze differences in allocation behavior with regard to gender for subjects B . Mean share passed is rather similar for both males and females except for treatment 1 (see Table 6.4).

Table 6.4: Descriptive results of B 's choice with regard to gender

Treatment	Mean share passed by males	Obs males	Mean share passed by females	Obs females
1B	0.25	28	0.36	16
2B	0.36	21	0.38	14
3B	0.37	23	0.43	12
4B	0.18	33	0.21	31
4BDice	0.17	20	0.23	16
4BSelf	0.20	13	0.19	15

Testing for gender effects, I do not find any significant differences between females and males in any of the treatments (see Tables 6.5). The gender effect in treatment 2 when A allocated vanishes, possibly being dominated by the social history effect.

Table 6.5: Testing for the impact of gender on B 's choice

Treatments compared	MW	AD	ES
1B	-1.246 (0.213)	1.720 (0.128)	4.722 (0.317)
2B	0.190 (0.849)	0.526 (0.738)	3.540 (0.472)
3B	0 (1)	1.118 (0.303)	8.230 (0.084)
4B	-0.422 (0.673)	0.728 (0.539)	2.248 (0.690)
4BSelf	-0.305 (0.760)	0.556 (0.712)	3.251 (0.517)
4BDice	-0.934 (0.350)	0.668 (0.600)	2.268 (0.687)

Other factors, such as income and subject of study, do not yield any significant results either when using Spearman's rank correlation coefficient.

Comparing A 's and B 's choice (within treatments) yields significant differences for all treatments, except in treatment 1, the MW statistic is not significant, see Table 6.6. Obviously, B passes on significantly less than A . I relegated the comparison of the treatments of the male and female subsamples to the appendix, see section 11.4, because this comparison does from my point

Table 6.6: Comparison of A's and B's allocation choice

Treatments	MW	AD	ES
1	1.192 (0.233)	3.029 (0.025)	20.140 (0.000)
2	3.220 (0.002)	6.650 (0.000)	27.190 (0.000)
3	3.245 (0.001)	5.774 (0.001)	33.011 (0.000)
4	-2.828 (0.005)	4.230 (0.006)	9.652 (0.047)

of view not lead to insightful results. The same holds for the comparison of *A*'s and *B*'s behavior for the male and female subsamples.

6.1.3 Comparison of proportions passing zero

Taking a look at proportions of subjects passing zero shows that there is a difference between the treatments without and with social interaction. Only a small share of subjects in the treatments with social interaction took the whole pie while in the treatments without social interaction this share was substantial (see Table 6.7).

Table 6.7: Proportion of subjects *B* passing zero

Treatment	Proportion passing zero
1B	0.27
2B	0.03
3B	0.09
4B	0.30
4Bself	0.32
4Bice	0.28
Repl	0.25
FHSS	0.30
HMSS	0.21

A comparison of proportions of subjects passing zero in the different treatments using Fisher's exact test confirms this impression, see Table 6.8. Treatment 2 is significantly different from treatment 1 and from 4, while between the other treatments there is no significant difference. Treatment 3 is marginally different from treatment 1 and 4, respectively.

The difference between proportions of males and females passing zero seems to be substantial, too, at least for Treatment 1-2 (see Table 6.9). But the comparison of proportions of males and females passing nothing shows no

Table 6.8: Comparison of proportion of subjects B passing zero

Treatments compared	P-value from Fisher exact test
1B vs 2B	0.013
1B vs 3B	0.092
1B vs 4B	1
2B vs 3B	0.615
2B vs 4B	0.007
3B vs 4B	0.073
1B vs Repl	0.829
1B vs FHSS	0.832
1B vs HMSS	1

Table 6.9: Proportion of male and female B s passing zero

Treatment	Male: Proportion passing zero	Female: Proportion passing zero
1B	0.27	0.13
2B	0.13	0
3B	0.05	0
4B	0.36	0.32
4BDice	0.23	0.40
4BSelf	0.30	0.25

significant difference (see Table 6.10), the effects seems to be too weak to show up with these sample sizes.

Table 6.10: Comparison of proportion of male and female B s passing zero

Treatment	P-value Fisher exact test
1B	0.300
2B	1
3B	0.54
4B	0.798

6.2 Treatment comparison, including only B s passing a positive amount

Comparing samples where the B s passing on zero are excluded does not reveal new insights. Just note, that some minor results that have been significant turn out to be not significant any longer. I skip mentioning these because a useful interpretation is very hard. For the sake of completeness I relegate these comparisons to the appendix.

6.3 Regression analysis of B 's choice

As mentioned earlier, by virtue of the social history effect, type- B subjects' decisions may be influenced by the varying amounts passed on by A . This effect must not be confused with a possible scale effect, which according to FHSS would be of minor importance. Hence I again use a regression analysis at this point to explain what factors determine B 's behavior. With the share passed on by B as the dependent variable, I use a similar regression equation as used for A , extended by some further regressors. In order to capture the social history effect, I add the share being passed on by A to B , $(1 - \alpha)$. Additionally, I include two dummies *more_expected*² and *less_expected*³, indicating whether B 's share passed on by A falls short of or exceeds B 's expectations, respectively.⁴ I obtain the following regression equation:

$$\begin{aligned} (1 - \beta) = & \beta_0 + \beta_1 \cdot (1 - \alpha) + \beta_2 \cdot \textit{less_expected} + \beta_3 \cdot \textit{more_expected} \\ & + \beta_4 \cdot \textit{gender} + \beta_5 \cdot \textit{income} + \beta_6 \cdot \textit{age} \\ & + \beta_7 \cdot \textit{subject} + \beta_8 \cdot \textit{sust} + \beta_9 \cdot \textit{treatment} + \varepsilon \end{aligned}$$

Using OLS yields the results from Table 6.11.

Applying once again the model selection procedure by Herwartz (2007), I obtain the final model implicit in Table 6.12.

Replacing *less_expected* by a regressor that captures the expected share of money passed by B reduces R^2 from 0.293 to 0.250. Hence, I abstain from doing so.

I observe that the social history effect represented by the variable $(1 - \alpha)$ and indicating how much A passes on to B is highly significant and much stronger than any other effect. If A increases the share to be passed to the B by 1 percentage point, B reciprocates by increasing the share passed on to C by almost 0.4 percentage points. This finding is confirmed by 18.6% of the subjects B explaining their behavior by reciprocity (note that more than one explanation was possible). Interestingly, 10.3% of the subjects A obviously anticipated this social history effect. They explained their choice by expressing the hope that they would act as a standard for B .

Gender may play a role because females are again slightly more generous than males (the result being significant close to the 5% significance level). However, in the regression the variable *gender* was discarded by the model selection

²Indicating that B expected more than B actually received.

³Indicating that B expected less than B actually received.

⁴Note that subjects were asked about their expectations after being notified about the share they received, see section 7 for a discussion on that matter.

Table 6.11: Regression B , initial model. Dependent variable: Share passed on by B

Variable	OLS	Tobit
Constant	0.099 (0.273)	0.007 (0.948)
$1 - \alpha$	0.392 (0.000)	0.521 (0.000)
Gender	-0.048 (0.084)	-0.056 (0.085)
Income	0.009 (0.506)	0.012 (0.463)
Less_expected	-0.148 (0.000)	-0.175 (0.000)
More_expected	0.042 (0.243)	0.058 (0.172)
Age	0.003 (0.333)	0.004 (0.343)
Science	-0.044 (0.253)	-0.051 (0.264)
Economics	-0.038 (0.227)	-0.052 (0.163)
Humanities	-0.060 (0.263)	-0.059 (0.344)
Sustainability	0.039 (0.154)	0.046 (0.155)
Exp2	0.052 (0.203)	0.071 (0.140)
Exp3	0.061 (0.141)	0.075 (0.126)
Exp4	-0.085 (0.014)	-0.086 (0.035)

process. Income is no longer significant. Age, study subject, and knowledge about the concept of sustainability are not important in this regression, either, nor is it important whether the candidate expected more money.

Interestingly, however, the expectations of subjects B about the generosity of A play an important role. I observe that subjects who had been positively surprised by A 's choice (*less_expected*) reciprocate negatively by passing on 16 percentage points less. At first glance, this result seems to be counterintuitive, as I found a strong propensity to reciprocate positively. There are two possible explanations for this paradox, though. First, subjects B may have an a-priori idea about how much C should obtain, say 33.3%. If A passes on more than expected, say 70% instead of 66.6%, B may feel inclined to keep the extra

Table 6.12: Regression B , final model. Dependent variable: Shares passed on by B

Variable	OLS	Tobit
Constant	0.207 (0.000)	0.145 (0.002)
$1 - \alpha$	0.386 (0.000)	0.502 (0.000)
Less_expected	-0.163 (0.000)	-0.193 (0.000)
Exp4	-0.107 (0.000)	-0.117 (0.001)

share of 3.3%. An alternative explanation is based on a self-selection effect. In an environment where subjects have no experience about how dictators customarily divide a pie, the adequate heuristic would be to infer from yourself. Being a selfish type, one would expect others to be selfish, too, and vice versa. Thus the value of the expectation dummy can be interpreted as a proxy for subject type. Under this hypothesis, there is a positive correlation between being greedy and being positively surprised. In other words, if the variable *less_expected* takes the value 1, it is likely to pick a selfish type that would pass on little anyway. This finding about subjects having different preferences with respect to altruism is consistent with similar results obtained by Andreoni and Miller (2002). Those authors distinguished between three types of preferences among subjects: The selfish type, whose utility depends only on her own payoff; the Rawlsian type, whose utility depends on her and the recipient's payoff and is represented by a Leontief utility function; and finally the utilitarian type, who considers the allocator's and the recipient's payoff as perfect substitutes.

Another interesting result I have obtained is that social interaction does not matter any longer, since the dummies for treatment 2 and 3 are not significant. Hence the results from the treatment comparison must be interpreted in the following way: There is a treatment effect that results not so much from the social interaction between B and C , but from the previous social interaction between A and B . This effect is transferred via the social history effect.

Finally, the outcomes from treatment 4 turn out to be significantly different from the outcomes in all other treatments (almost at the 1% significance level), confirming the results from the treatment comparison in the former section.

The adjusted R-squared is 0.29 which is considerably higher than 0.16, the corresponding value of the regression for subjects A . I attribute this to the strong influence of reciprocity. A White test rejects heteroscedasticity (p-value of 0.738), hence OLS is suitable. Running a Tobit regression as well (the

results of which are also reported in Table 6.12), I observe that significance levels improve slightly for the dummy variable representing B 's expectations.

One problem in my analysis may be possible multicollinearity between the variables $(1 - \alpha)$ and *less_expected* since both regressors are related. If α is low, B receives a large amount. As a consequence, it is likely that B expected less. Correlation between α and *less_expected* is -0.3766 indicating a medium negative relation. Farrar and Glauber (1967) cite a rule of thumb which says that multicollinearity is a severe problem starting from a correlation coefficient $> |0.8|$ (which is far from my value) but mention, that this rule has to be used with care. Farrar and Glauber (1967) also noted that the R^2 should be larger than the correlation between the two variables which is not the case in my analysis. Marquardt (1977) states that the variance inflation factor⁵, VIF, of a variable should be between 1.0 and 10. The VIF of *less_expected* is 1.15 and hence very far away from the values > 10 where multicollinearity problems start. Note, that if multicollinearity is present, the estimation remains best linear unbiased (BLUE) and not affected in its forecasting properties for the dependent variable. But: The interpretation of the regressors and the value of the regressors could change dramatically, since the effect of collinearity and relation to the dependent variable cannot be distinguished in the estimation process.

⁵Interpretation of the VIF according to Maddala (1988): The ratio of the actual variance of the regressor to what the variance of the regressor would have been if the variable was uncorrelated with the remaining variables. In the best case, this would be 1.

Chapter 7

Analysis of subjects' expectations

In form 2, subjects were also asked about their expectation on the amount of money they received. The results for B 's expectation are reported in Table 7.1. Subjects B were on average pretty reliable in there expectations on the

Table 7.1: Expectations B on A 's allocation choice

Treatment	Mean share received by B	Mean expectations B	Variance expectations B
1	0.36	0.39	0.23
2	0.48	0.44	0.18
3	0.49	0.51	0.24
4	0.31	0.35	0.18
4self	0.33	0.38	0.17
4dice	0.30	0.32	0.19

amount of money that was allocated, however, the variance of the expectations was comparably large. But this result has to be treated carefully. Subjects B had to report their expectation after they have been informed about A 's allocation. Hence, anchoring¹ could play a role such that there is a possible adjustment of subjects in the right direction. I decided to not asking subjects about their expectations ex-ante because this would have introduced a different factor possibly influencing the allocation behavior and the allocation choice was paramount for this study while expectations where of ancillary importance.

¹Anchoring means that subjects, if they have to estimate a value, take an initial value into account if they make their estimates and hence adjust these according to the staring value. Tversky and Kahneman (1974) have investigated this effect: In presence of subjects, a wheel of fortune was spun. Subjects were paid for accuracy. They asked e.g. a group of subjects to estimate the share of African countries in the UN. As an example for groups receiving a 10 by the wheel their median estimate was 25, while groups receiving a 65, median estimate was 45.

Also candidates C were asked about their expectations (see Table 7.2). Subjects C were also very reliable in their predictions but also the C can-

Table 7.2: Expectations C on B 's allocation choice

Treatment	Mean share received by C	Mean expectations C	Variance expectations C
1	0.11	0.09	0.10
2	0.18	0.12	0.11
3	0.20	0.18	0.12
4	0.07	0.12	0.12
4self	0.08	0.11	0.12
4dice	0.06	0.12	0.13

didates were informed about the allocation before they had to report their expectations. Note, that there is a rather high variance among the C candidate's expectation, too.

Regressing $(1 - \alpha)$ and $(1 - \beta)$ and personal data on the expected amount of money and applying the model selection process by Herwartz (2007, see 5.3), I obtain the following regression model (see Table 7.3). $R^2 = 0.201$ is

Table 7.3: Regression results with expectation B as dependent variable

Variable	OLS
Constant	0.227 (0.000)
$(1 - \beta)$	0.351 (0.000)
$(1 - \alpha)$	0.205 (0.007)

rather low. The strong impact of $(1 - \beta)$, the share subject B passed on, can again be explained by the idea that the share you take indicates that you are of a certain type, (see 6.3). Hence, if you passed on more (meaning that you are a generous type) you also expected more. The regressor $(1 - \alpha)$ is highly significant, thus anchoring is at play, here. Interestingly, the treatment dummies are not incorporated in the model selection process. Possibly their influence is absorbed by $(1 - \alpha)$.

I skip showing the regression results for the 'expectation of C on the share passed on by B because the regression conveys no new insights.

Chapter 8

Analysis of subjects' motivations for their allocation choice

After the allocation choice, subjects were asked to explain their behavior. The question was open, hence there was a vast array of answers. Of course it would have been possible to let subjects choose within a list of reasons. But maybe it would not have been possible to anticipate all the reasons why subjects chose their allocation and hence the open question format was used. I categorized subjects' explanations. Often people mentioned more than one reason. I included all explanations in the different categories hence the categories are not mutually exclusive. A closer explanation of the different categories is attached in the appendix, see 11.6.1, 11.6.2 and 11.6.3. The number in brackets behind the category refers to the explanations in the appendix.

8.1 The first generation's explanations

Table 8.1 gives an overview on the used explanations by the first generation.

First, I take a look at the explanations aggregated over all treatments. The explanation used mostly by the subjects *A* is 'fairness'. A share of 0.35 of the subjects explained their behavior with this reason. Fairness could either be attributed to the inequity aversion preferences by Fehr and Schmidt (1999) and Bolton and Ockenfels (1999) or to the 'social welfare' preferences by Andreoni and Miller (2002) and Charness and Rabin (2002), see section 2.2.13. A share of 0.24 of the subjects used egoism as explanation. Often, subjects used more than two reasons to explain their behavior and especially frequent is the combination of fairness and selfishness, hence I have a rather low support for

Table 8.1: Explanation *A*

Explanation <i>A</i>	T1	T2	T3	T4	all
Fairness (1)	0.185	0.450	0.538	0.273	0.349
Selfishness (2)	0.278	0.200	0.179	0.303	0.241
<i>B</i> and/or <i>C</i> should receive >0 (3)	0.204	0.175	0.077	0.273	0.181
Luck (4)	0.259	0.125	0.154	0.061	0.163
Personal neediness (5)	0.167	0.075	0.051	0.242	0.133
Role model for <i>B</i> (6)	0.056	0.050	0.179	0.121	0.096
Anonymity (7)	0.204	0.025	0	0.121	0.096
Others the same (8)	0.074	0.050	0.077	0.121	0.078
Power (9)	0.093	0.050	0.026	0.121	0.072
Unclear answer (10)	0.111	0.025	0	0.061	0.054
<i>B</i> plays twice (11)	0	0.050	0.179	0	0.054
Norm (intrinsic) (12)	0.019	0.125	0	0.061	0.048
Same effort (13)	0.074	0	0.077	0	0.042
Participation to earn (14)	0.019	0.025	0.077	0	0.030
Rationality (15)	0.074	0.025	0	0	0.030
Maintain incentive (16)	0	0.100	0	0	0.024
Norm (17)	0	0	0.051	0.061	0.024
Maintain incentive not important (18)	0	0.075	0	0	0.018
<i>C</i> should receive >0 (19)	0.037	0	0	0	0.012
<i>C</i> does not know (20)	0	0	0.026	0.030	0.012
Did not get to know <i>C</i> (21)	0	0	0.051	0	0.012
Others are not needy (22)	0.037	0	0	0	0.012
Rules of game accepted (23)	0	0	0.026	0.030	0.012
Reward for cooperation (24)	0	0	0.026	0	0.006
Have no play vis-à-vis with <i>B</i> (25)	0	0.025	0	0	0.006

the classical economic model of an agent with a utility function having his own payoff as single argument. A share of 0.18 used the explanation ‘*B* and/or *C* should receive >0 ’ which is rather similar to the ‘fairness’ category but these explanations have not explicitly mentioned fairness. Also, the connotation is little different: There would be an altruism term in *A*’s utility function but *A* does not necessarily care about an equitable outcome, it is only important that *B* and *C* will not end up with nothing. This preference type refers to the ‘warm glow of giving’ preferences by Sadrieh (2003), see section 2.6.2. Obviously, for some subjects (a share of 0.16) it does matter if they earned the right to be in position *A* or if it was pure chance if they were in this position. This is in line with the results by HMSS (1994) who found that it made a difference if the entitlement to be in the position of the allocator was earned or the result of a random process. A share of 0.13 claimed that personal neediness governed their behavior. This might also be seen as a confirmation of the models where others’ payoff is included in the utility function: Even though the subjects

where selfish, they said that if it would not be costly, they would like to help B and C or if they had more money than at the present stage, they would be more generous. The expressed ‘Hope to be role model for B ’ by a share of 0.1 of the A s confirms that A anticipates that B is prone to reciprocity. This stresses the importance of the model by Charness and Rabin who incorporate a reciprocity term in their preferences. Anonymity does matter. It is clear to see that we do care more about people we know personally, so being more generous with people we interacted socially would make sense. The study by Andreoni (2007), where the possible veil of hiding behind a random process lead to less generosity, confirms this evidence (see section 2.2.15). A share of 0.1 of the subjects referred to anonymity, explaining their choice. Note, that for treatment 2 and 3 this explanation does not make much sense¹, hence, the overall share of people who think that anonymity is important should be almost 0.2. ‘Others the same’ is a rather collectivistic explanation by a share of 0.08 of the subjects. It seems that some subjects would like to conform to other people’s behavior. This explanation is used both by subjects who take more than the fair share and by subjects who take less. ‘Power’ is used by a share of 0.07 of the subjects. Here, the mere possibility to behave unfair seems to be justification enough to do so.

When analyzing explanations with regard to treatments, I find some differences between treatments. ‘Fairness’ plays a much bigger role for the social interaction treatments while ‘Selfishness’ is much more important for the treatments without social interaction. Obviously, social interaction seems to trigger fairness while it oppresses selfishness. ‘ B and/or C should receive an amount >0 ’ was used when rather low amounts of money were passed to B and C which was more often the case for treatment 1 and 4 and hence it appears more often in this treatments. Neediness does also appear much more often in treatment 1 and 4. Since the share of people that are needy should be roughly equal in all treatments, subjects seem to feel compelled to use this explanation more often if they were greedy, to justify their behavior and maybe maintain a positive self-image.

The explanations for the replication of the TDG are roughly in line with treatment 1, but more subjects refer to fairness while less to selfishness, see Table 8.2.

¹The one subject in treatment 3, who mentioned anonymity referred to the fact that she did not get to know the personal data of B .

Table 8.2: Explanation *A* in the TDG

Explanation <i>A</i>	Replication
Fairness	0.300
Selfishness	0.175
<i>B</i> and/or <i>C</i> should receive	> 0.200
Luck	0.050
Personal neediness	0.175
Anonymity	0.15
S.th. unfair in real life (33)	0.05
Adequate Pay <i>B</i> (34)	0.05
'Positive Empathy' (35)	0.05
'Negative Empathy' (36)	0.025
Others the same	0.025
Power	0.025
Rationality	0.025
Gift (37)	0.025
Christian values (38)	0.025
Game	0.025

8.2 The second generation's explanations

Table 8.3 gives an overview on the used explanations by the second generation.

The analysis all treatments reveals that also most of the subjects *B* explain their behavior with 'fairness', though only a share of 0.31 overall used it compared to a share of 0.34 of the *A*'s. 'Selfishness' is second with an overall share of 0.21. A share of 0.19 justified their behavior with reciprocity. Given that reciprocity was named, a share of 0.59 used positive reciprocity and share of 0.38 used negative reciprocity as explanation. The rest imitated *A*'s behavior by taking the same share as *A*, hence not allowing a categorization of reciprocity in terms of positive or negative. 'Personal neediness' is still mentioned frequently (a share of 0.11), while the importance of 'luck' as explanation is greatly reduced from a share of 0.16 for the *As* to 0.03 for the *Bs*. The importance that a positive amount of money should be passed on is reduced to a share of 0.11. A share of 0.10 used the argument that there is not much left to pass on while 'power' is only used by a share of 0.03. Anonymity is not very important any longer, either, only a share of 0.06 of the subjects used it as explanation.

Comparing the different treatments with each other yields results that are partly similar to the analysis of *A*'s behavior. Fairness is much more important in the treatments with social interaction, again in treatment 3 most of the people use this explanation. Selfishness is most important in the treatments

Table 8.3: Explanation B

Explanation B	T1	T2	T3	T4	all
Fairness (1)	0.250	0.371	0.514	0.203	0.309
Selfishness (2)	0.205	0.171	0.143	0.370	0.208
Reciprocity (26)	0.295	0.229	0.343	0.063	0.208
Personal neediness (5)	0.136	0.086	0.029	0.156	0.112
Positive reciprocity (5)	0.091	0.200	0.257	0	0.112
C should receive >0 (3)	0.068	0.057	0.029	0.188	0.101
Not much left to pass (27)	0.136	0.086	0	0.125	0.096
Unclear (10)	0.045	0.143	0.057	0.094	0.084
Negative reciprocity (5)	0.159	0.029	0.086	0.047	0.079
Anonymity (7)	0.114	0	0	0.078	0.056
Wanted certain amount (28)	0.068	0	0	0.047	0.034
Power (9)	0.023	0.029	0.057	0.031	0.034
Luck (4)	0.045	0	0	0.047	0.028
Maintain incentive (16)	0	0.114	0	0	0.022
B plays twice (29)	0	0	0.114	0	0.022
Norm (intrinsic) (12)	0.045	0.029	0	0.016	0.022
Norm (17)	0.023	0.057	0.029	0.000	0.022
'Got to know each other' (30)	0	0	0.086	0	0.017
'Rationality' (15)	0.045	0	0	0.016	0.017
Compassion (31)	0	0.029	0	0.031	0.017
'Others the same' (8)	0	0	0	0.031	0.011
Participation to earn (14)	0	0.029	0	0.016	0.011
Same effort (13)	0.023	0	0	0	0.006
' C does not know' (20)	0	0.029	0	0	0.006

without social interaction, here, treatment 4 has the highest share. Reciprocity is mentioned roughly with the same frequency in treatment 1 to 3 while in treatment 4 it is rarely used, which makes sense since subjects should not be able to tell apart if a random process or a subject allocated.

Again, social interaction seems to trigger the desire to be fair while reducing selfishness. The higher share of subjects using positive reciprocity for the treatments with social interaction could have two reasons: Since subjects passed on more money on average in these treatments, it is likely that positive reciprocity is triggered more often. But also the social interaction effect could enhance the desire to reciprocate positively. However, social interaction may also evoke a stronger propensity to reciprocate negatively because you feel particularly upset if someone you interacted with is unfair to you.

8.3 A note on the validity of subjects' answers

Since subjects explanations were not linked to any incentive scheme that rewards answering truly, one might doubt the validity of the answers. Subjects may feel uncomfortably reporting the true motivations and hence make up explanations that they perceive as socially more adequate. However, an incentive scheme that rewards answering truly is unfeasible, hence, this is the only way to obtain insight into subjects' motivations.

It is even unclear if the stated explanations by subjects are the real causes since subjects may generally not be aware of the underlying mechanism that is responsible for their decision. This claim might sound a little presumptuous but maybe my point becomes clearer after an example. This example is not directly related to resource allocation but to moral reasoning, which should also be part of the allocation decision. Consider the following example first introduced by Foot (1967): There are two scenarios, in both scenarios there is a train out of control and will kill five workmen on the tracks if nothing is done. In the first scenario, a person is standing beside a switch, that will turn the train on a sideway, thereby killing one person. In the second scenario, a person is standing on a bridge together with a heavy man, who would, if thrown on the tracks, stop the train. Mikhail (2000) asked subjects, if it was morally permissible to throw the switch and if it was morally permissible to push the man down the tracks. An overwhelming majority of the subjects considered acting in the first scenario permissible while in the second scenario acting was not permissible, the difference was highly significant. The interesting aspect here is, that although there seemed to be a clear pattern, the subjects asked to justify their choice were not conscious of the underlying operational principle. They mostly used inconclusive answers like referring to intuition. Mikhail (2000) states that there is a moral grammar that is used by all subjects. This concept is the same as for using language in one's mother tongue, where the governing principles are, even if correctly applied, not necessarily consciously known. Mikhail (2000) claims that comparing the two scenarios, subjects unconsciously refer to the moral grammar: On the one hand, the negative consequence in the first scenario is only foreseen while in the second it is intended². In addition, in the first scenario the required action is neutral (flipping a switch, which is an impersonal act) while the second involves a negative, personal action (pushing a man down a bridge). Of course the moral ramifications seem to be more clear cut in allocation games, nevertheless it can

²In the first scenario, flipping the switch is the desired action, the person's death is a foreseen consequence. In the second scenario you stop the train by killing the heavy man.

be doubted if the numerous reasons used as explanations are really the true motivations or if there are not very few principles, governing subjects behavior in a partly unconscious manner.

Chapter 9

Suggestions for future experimental studies

Experiments in natural science are usually conducted in a controlled laboratory setup. Everything is held constant¹ and only one factor is varied to investigate its impact. Experimental economics try to ensure a similar procedure, although this is much more difficult since we work with human beings. Nonetheless, it is tried to achieve a procedure that is standardized to ensure that only the treatment variations and not any other, uncontrolled variations affect subjects behavior.

That starts with the subject recruitment. I always used the same flyer to recruit subjects. To facilitate the recruitment process, I used an online-sign-up procedure from the second session on. Here, after having noticed some problems, I introduced some additional features in the recruitment process. One important issue is absenteeism of subjects, especially if you need multiples of three. Usually, a share of 0.1 to 0.2 of the subjects does not show up. Hence, I started to recruit reserve candidates that received the show-up fee and participated only if needed. I also started sending a reminder one day before the experiment. Since some subjects tended to come late, I also included a notice in the reminder, that subjects who were late could not participate and did not receive any payoff.

With regard to the procedure, I had a protocol, but the protocol was not as explicit as it should have been. Every word and sentence that is used in the procedure should be standardized according to protocol. Still, I do not believe that this less detailed protocol had a measurable impact.

There is a shortcoming with respect to the asked income in the questionnaires. I should have asked for the income at disposal after paying rent from

¹The well-known ‘*ceteribus paribus*’ approach.

the begin on. Also, it would have been better to ask for the amount of money and not to use categories.

Another important issue is the payment of the subjects. The experimenter (not the other subjects) got to know the subjects' name during the experiment. It would have been better to preserve subjects' anonymity using simply *A*, *B*, and *C* and the corresponding numbers. I also paid subjects within groups *A*, *B*, and *C* together. It would have been better to pay all subjects privately which is standard for such experiments. This could be accomplished best by using envelopes that can be picked up a day later. Hence, subjects are not aware of the others payoff and the experimenter has time to assign the payoffs.

I believe it is unlikely that the criticized aspects influenced subjects' behavior. Nonetheless, it would have been better if the experiment had been run considering the named shortcomings from the begin on and thereby ensuring a higher level of standardization and/or conform more to standards in economic experiments.

Chapter 10

Conclusion: Intergenerational solidarity - it is feasible!

There is a large strand of literature concerning the experimental study of this thesis. The evidence from this literature is rich and it is not easy to find a common denominator for all the results. Rather, one has to regard the development of the experiments as an evolutionary evolution where the experiments one after another investigated possible explanations for the results so far and often improved the setup in response to former experiments. Especially in the first experiments, the development of the procedure could be seen as ‘learning by doing’ where a standard methodology still had to be developed.

The experiments by Güth *et al.* (1982) certainly were pioneering but suffered from methodological deficiencies and also some later experiments had serious shortcomings though often providing good ideas, see e.g. Frohlich *et al.* (2001). Güth *et al.* (1982) did not seem to ensure between-subject anonymity, instructions did not seem to have a real protocol and were done informally and the distance between experimenter and subjects was too close. But, as reaction to such shortcomings the experimental procedures were refined and standard setups were developed. FHSS are to be mentioned in this context, ensuring between-subject anonymity and having a clear protocol for instructions. HMSS set the other standard by implementing the double blind procedure that ensured also subject-experimenter anonymity. Both setups have been repeatedly used by other studies or were at least used as benchmark. The results from the ultimatum and especially the dictator games confronted the scientists with a puzzle: Subjects did not behave in accordance with the classic economic model of the purely selfish agent but displayed a propensity for other regarding behavior. There were a lot of attempts to explain this behavior but one could encompass most of the literature in two factions: One admitting that subjects

are altruistic and one rejecting this hypothesis while claiming that these results can also be traced back to the selfish agent. First, FHSS found evidence for altruism in a controlled standard setup. HMSS introduced the double blind procedure that reduced altruism to a negligible extent. HMSS claim that subjects' altruism is caused by subjects' expectation that being altruistic could lead to future pay-off via being considered for future experiments. If subjects cannot be identified and do not have the incentive to please the experimenter, they behave in accordance with the traditional model. Bohnet and Frey (1999) also ensured this experimenter-subject anonymity (with regard to allocators) but reduced social distance between subjects without creating an incentive to being nice due to repeated game effects. Their experimental results provide evidence for subjects' inclination to be altruistic. This evidence is confirmed by many following studies that reduce social distance. Frohlich *et al.* (2001) deliver an alternative explanation why the double blind procedure could cause subjects to be less generous, they claim that subjects could doubt the existence of a recipient. Unfortunately, they compromised their evidence by treatments with too few observations and by using the inadequate statistical instruments in their treatment comparison. Cherry *et al.* (2003) showed that earning the pie reduced altruism without using the double blind procedure.

Other approaches abstracting from anonymity but introducing preferences, governed by certain features, have been made to explain the subjects' behavior in these experiments. Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) explain the evidence by subjects' inequality aversion. Andreoni and Miller (2002) interpret their experimental results by using different preference types of players. Except for the selfish type, players have their own payoff and the other's payoff as arguments in their utility function. Charness and Rabin (2002) findings from an experimental study can be explained best by preferences that depend on social welfare. Charness' and Rabin's results from a response games (where two subjects interacted sequentially) also suggest that reciprocity is an important element in subjects' decision making process. In their experimental study, subjects' willingness to help others was greatly reduced if previously the other subjects themselves behaved selfishly. The results of the classic ultimatum games, where recipients frequently forfeit offers that they perceive as too low confirm this theory (see e.g. Güth *et al.*, 1982).

When explaining why altruism has emerged authors sometimes refer to our ancestral heritage, see e.g. Haley (2005) or Burnham (2003). They claim that altruism facilitates cooperation and was useful in early small-scale kinship societies. From their view it is a genetical atavism that manifests if triggered adequately. But the evidence by Henrich (2003) rather points in the direction

that cultural education plays an important role, too.

Bardsley (2005) and List (2007) gave the allocators the opportunity to take money from the recipient. They found that the share of subjects taking money was larger than subjects, who gave nothing in the traditional setup. Thus, they conclude 'heretically' that the evidence for other-regarding behavior so far could be an experimental artefact

Another approach explaining the experimental evidence of other-regarding behavior is given by neuro-scientists. According to Singer *et al.* (2005), people have the capacity to empathize. That means that people are able to put themselves in the position of others and try to build prospects of the others ideas. One could say they share the state of mind of others without a direct emotional stimulus of themselves. Hence, imagining or seeing someone else who feels certain sensations triggers automatically a representation of this state in the subjects brain.¹ According to Singer *et al.*, MRI² studies have shown that there are common neural responses to either seeing other people's sensations (both known or unknown, of course the extent varies) or experience these sensations yourself. They claim that mind reading and empathy have important impacts on economic behavior. People are able to predict other people's behavior that is necessary in order to act consistently if you would like to achieve your own aims. If your opponent is a reciprocator, it is not only important to be aware of his beliefs but also of his emotional state. In order to reach your goals you must be able to understand which action will render the opponent in an emotionally desirable state. But this capacity for empathy has a side-effect. People become more concerned about other people and their propensity towards altruistic behavior could be enhanced. Thus, a process that initially added to optimizing your ability to act in a self-interested manner induces other-regarding behavior due to its side-effects. However, the causality may as well be vice versa since cooperative behavior often allows to reach Pareto superior stages and this in turn could simply be a fit to evolution. It is important to note that the ability to empathize varies across people. This could also explain the differences in preferences between subject types found by Andreoni and Miller (2002).

Camerer *et al.* (2005) reject the approach of preferences that have the purpose to maximize humans utility. They point out that 'humans did not evolve to be happy, but to survive and to reproduce'. One attempt by the body to ensure this survival is called *homeostasis*. Here, as soon as detectors monitor

¹This process can work without a conscious control but its occurrence can be deliberately oppressed.

²Magnetic Resonance Imaging

a divergence from certain equilibrium states (e.g. a body temperature of 37 degree Celsius) unconscious and conscious actions result. The unconscious action is to start sweating. The conscious actions are normally attached to reward and punishment. Departing from the reference point feels unpleasant (e.g. like being to warm if it is hot outside), returning to this point feels good (taking a bath). Economists usually see preferences and the attached pleasure of fulfilling them as the starting point of human behavior. From the neuroscience perspective, pleasure must be seen as an information due to *homeostasis*.

Sanfey *et al.* (2006) also criticize the traditional economic models that have a unitary all purpose approach. They admit that these models are ‘formally explicit, analytically tractable, and can be used to make quantitatively precise predictions about decision-making in a wide variety of circumstances’. But they claim that there is a large strand of literature in psychology and especially neuroscience that indicates that human behavior is not the result of a single process but of the interaction of different subsystems. A very important distinction is the one between automatic and controlled processes. In addition there are a subsystems for emotions and for deliberation. This evidence is confirmed by the study by Sanfey *et al.* (2003) that stresses the trade of between emotions and deliberation. The evidence from neuroscience could by the way still be in line with the cultural learning hypothesis, since it is not clear if the neurological responses are manifestations that are genetically unalterable or part of a process that has been adopted via cultural learning.

For this study, I devised an experiment to test for people’s willingness to share resources with third parties they will not meet directly. One possible interpretation of this experiment is whether people are willing to bequeath resources to future generations they will never meet. The experimental evidence confirms the well-established fact that people are willing to sacrifice resources for the sake of others without being rewarded materially. I also find that the subjects are aware of the intergenerational context of my experimental set-up. They pass on more resources than we usually observe in the traditional dictator game. The subjects thus indicate that they take the upcoming generation into consideration. They also deliberately sacrifice resources to trigger generosity by the next generation via the reciprocity channel. Vice versa, I observe that if people realize that the previous generation behaved generously, they at least partially reciprocate by behaving generously to a third party they get involved with and that one may interpret as the ‘future generation’. Thus there is empirical evidence for the existence of a social history effect.

A further important finding is that social interaction enhances (intergenerational) generosity. The reasons for this generosity are not entirely obvious,

but there are some clues. People may feel uncomfortable if they behave greedily without the veil of complete anonymity. Thus social interaction can lead to increased affection triggering altruism. With respect to the first generation, I also find that gender matters in particular circumstances. Moreover, higher income (as socio-economic feature) increases the willingness to sacrifice resources at the initial stage.

The dominant factor determining the second generation's decision is the choice made by the first generation. On the one hand, I observe a strong social history effect. If the first generation increases the share passed on to the second generation, the second generation will indirectly reciprocate by increasing the share it passes on to the third generation. On the other hand, expectations about how much will be received, play an important role too. If the second generation is positively surprised by the size of share passed on to it, it is likely to keep this surplus for itself instead of passing it on to the third generation. By contrast, the impact of gender is considerably softened, and the impact of income vanishes completely.

Finally, I find that introducing the possibility that the first generation's allocation choice was made by a random process reduces the generosity of the second generation significantly compared to the case where the choice is made with certainty by a human allocator.

My model also contributes to experiments on common-pool resources where subjects prove to be altruistic. In those set-ups, the desire to be altruistic could, however, be mixed with strategic considerations, while in my set-up the wish to be altruistic is obvious. My results allow for an optimistic conclusion. People confronted with very simple kinds of intergenerational problems seem to be willing to sacrifice some wealth in order to increase intergenerational fairness. Of course I am aware that this is just a simple laboratory experiment. However, I consider it a start in eliciting social preferences about the well-being of other, anonymous individuals we may be associated with via a sequence of social interactions.

Chapter 11

Appendix

11.1 Instructions

INSTRUCTIONS FOR EXPERIMENT I

There are three candidates called A , B , and C who are in a group. Candidates are randomly assigned to a group. The candidates do not get to know each other. The task is to divide 30 euros.

Step 1: Candidate A chooses an amount between 0 and 30 euros that she keeps for herself. The rest R will be divided by candidate B between B and C .

Step 2: Candidate B is told of the amount R that A has passed on to B and C . Then B chooses an amount between 0 and R that she keeps for herself. Candidate C receives the rest of the money.

Step 3: The allocated amounts are paid to A , B , and C .

The experiment is over.

INSTRUCTIONS FOR EXPERIMENT II

There are three candidates called A , B , and C who are in a group. Candidates are randomly assigned to a group. The task is to divide 30 euros. A and B and B and C meet, A and C do not meet.

- Step 1: Candidate A chooses an amount between 0 and 30 euros that she keeps for herself. The rest R will be divided by candidate B between B and C .
- Step 2: Candidate A and B play a kind of Scrabble (a game with letters, for instructions see below), in which they can earn additional money. The better A and B cooperate, the higher the amount which they earn. Each of them receives half of the amount earned. Time limit for the game is 15 minutes.
- Step 3: Candidate B chooses an amount between 0 and R that she keeps for herself. C receives the rest of the money.
- Step 4: Candidate B is told of the amount R that A passed on to B and C .
- Step 5: Candidates B and C play the scrabble game. See Step 2.
- Step 6: The allocated amounts, including the profit from playing the scrabble game, are paid to A , B , and C .

The experiment is over.

Instructions for the Scrabble game

There are 101 letters, that can be used to form words. The numbers on the letters indicate the value of each letter. Words may be formed crosswise. For each letter used, you obtain points according to the value of the letter. Letters that appear in two words count twice.

One point is worth two cents.

The amount earned by player A and B (or B and C) is divided equally between them.

Note that you can earn a maximum of 6 euros (that means 3 euros per player).

INSTRUCTIONS FOR EXPERIMENT III

There are three candidates called A , B , and C who are in a group. Candidates are randomly assigned to a group. The task is to divide 30 euros. A and B and B and C meet, A and C do not meet.

- Step 1: Candidates A and B play a kind of Scrabble (a game with letters, for instructions see below), in which they can earn additional money. The better A and B cooperate, the higher the amount they earn. Each of them receives half of the amount earned. Time limit for the Scrabble game is 15 minutes. B leaves the room.
- Step 2: Candidate A chooses an amount between 0 and 30 euros that she keeps for herself. The rest R will be divided by candidate B between B and C .
- Step 3: Candidate B is told of the amount R that A passed on to B and C .
- Step 4: Candidates B and C play the Scrabble game. See Step 1.
- Step 5: Candidate B chooses an amount between 0 and R that she keeps for herself. Candidate C receives the rest of the money.
- Step 6: The allocated amounts, including the profit from playing the Scrabble game, are paid to A , B , and C .

The experiment is over.

Instructions for the Scrabble game

There are 101 letters, that can be used to form words. The numbers on the letters indicate the value of each letter. Words may be formed crosswise. For each letter used, you obtain points according to the value of the letter. Letters that appear in two words count twice.

One point is worth two cents.

The amount earned by player A and B (or B and C) is divided equally between them.

Note that you can earn a maximum of 6 euros (that means 3 euros per player).

INSTRUCTIONS FOR EXPERIMENT IV

There are three candidates called A , B , and C who are in a group. Candidates are randomly assigned to a group. The candidates do not get to know each other. The task is to divide 30 euros.

Step 1: It is decided by tossing a coin whether i) a 20-sided dice divides for candidate A or ii) candidate A does the dividing herself.

i) If the 20-sided dice divides for candidate A :

Step 2: The 20-sided dice determines an integer amount (in euros, no cents) between 10 and 30 euros that A receives. The rest R will be divided by candidate B between B and C .

Step 3: Candidate B is told of the amount R that A has passed on to B and C . Then B chooses an amount between 0 and R that she keeps for herself. Candidate C receives the rest of the money.

Step 4: The allocated amounts are paid to A , B , and C .

The experiment is finished.

i) In case candidate A divides herself:

Step 2: Candidate A chooses an amount between 0 and 30 Euros that she keeps for herself. The rest R will be divided by candidate B between B and C .

Step 3: Candidate B is told the amount R that A passed on to B and C . Then B chooses an amount between 0 and R that she keeps for herself. Candidate C receives the rest of the money.

Step 4: The allocated amounts are paid to A , B , and C .

The experiment is over.

INSTRUCTIONS FOR THE EXPERIMENT V

There are two candidates called A and B who are in a group. Candidates are randomly assigned to a group. The candidates do not get to know each other. The task is to divide 30 euros.

Step 1: Candidate A chooses an amount between 0 and 30 euros that she keeps for herself. The rest R is given to candidate B .

Step 2: The allocated amounts are paid to A and B .

The experiment is over.

FORM 1

Date: _____

Group:

Type:

The amount of 30 euros is divided into

EUR for me

EUR for B and C (B will then divide this amount between herself and C).

Statistical data

1.) Age: years

2.) Sex: female male

3.) Subject studied: _____

4.) After having paid my rent, I have an average monthly income of

below 350 euros 650 - 800 euros above 1100 euros

350 - 500 euros 800 - 950 euros

500 - 650 euros 950 - 1100 euros

FORM 2

Questionnaire player *B*

Fill in the your subject number (e.g. B1)

Fill in the date

1.) Explain why you divided the amount the way you did.

2.) Did the amount of money you received from Player A correspond to your expectations? Tick accordingly:

a.) Yes, roughly.

b.) I expected more, roughly euros.

c.) I expected less, roughly euros.

3.) Have you ever heard the term 'sustainability'?

Yes

No

If yes, what do you think it means?

11.2 Program codes

11.2.1 Anderson-Darling test - my Matlab code

Note, that the p-value approximation is taken from Sefton's code.

```
i=1;
j=1;
i1=1;
j1=1;
z=0;
k=0;
m=0;
X1=sample1;
X2=sample2;
X1=sort(X1);
X2=sort(X2);
n1=size(X1,1);
n2=size(X2,1);
n12=n1+n2;
for n=1:(n12-1)
MinX1=min(X1(i:end));
MinX2=min(X2(j:end));
if MinX2<MinX1 & j1<(n2)
j=j+1;
j1=j1+1;
k=k+1;
elseif MinX2<MinX1 & j1==(n2)
j1=j1+1;
k=k+1;
elseif MinX2<MinX1 & j1>(n2) & i1<(n1)
i=i+1;
i1=i1+1;
k=k+1;
elseif
MinX2<MinX1 & j1>(n2) & i1==(n1)
i1=i1+1;
k=k+1;
elseif MinX2>MinX1 & i1<(n1) i=i+1
```

```

i1=i1+1;
k=k+1;
elseif MinX2>MinX1 & i1==(n1)
i1=i1+1;
k=k+1;
elseif MinX2>MinX1 & i1>(n1) & j1<(n2)
j=j+1;
j1=j1+1;
k=k+1;
elseif MinX2>MinX1 & i1>(n1) & j1==(n2)
j1=j1+1;
k=k+1;
end
Z1=(i1-1)/n1;
Z2=(j1-1)/n2;
w=(k/n12*(1-k/n12))(-1);
a=Z1-Z2;
a2=a(2);
c=w*(a2);
z=z+c;
end
AD=z*n1*n2/(n12)(2)

pval=0;
for i=1:10000
h=rand(n1+n2,1);
h=[h[ones(n1,1);zeros(n2,1)]];
h=sortrows(h,1);
mi=cumsum(h(1:n1+n2-1,2));
imin=1:n1+n2-1;
imin=imin';
Fmn=(mi*(n1+n2)-n1*imin);
Fmn=Fmn.(2);
ADart=(1/(n1*n2))*sum(Fmn./((imin*(n1+n2)-imin).(2)));
if ADart>AD
pval=pval+1/10000;
else pval=pval;
end
end
end

```

pval

11.2.2 Anderson-Darling test - GAUSS code by Martin Sefton

```
/* AD.PRG computes Anderson-Darling test */
n1=rows(X1);
n2=rows(X2);
h=(X1—X2) (ones(n1,1)—zeros(n2,1));
h=sortc(h,1);
mi=cumsumc(h[1:n1+n2-1,2]);
imin=seqa(1,1,n1+n2-1);
Fmn=(mi*(n1+n2)-n1*imin)^(2);
AD=(1/(n1*n2))*sumc(Fmn./(imin*(n1+n2)-imin)^(2));
rejad=AD.>2.492;

/* approximating p-value */
reps=10000
rep=1;
pval=0;do until rep>reps;
h=rndu(n1+n2,1) (ones(n1,1)—zeros(n2,1));
h=sortc(h,1);
mi=cumsumc(h[1:n1+n2-1,2]);
Fmn=(mi*(n1+n2)-n1*imin)^(2);
ADart=(1/(n1*n2))*sumc(Fmn./(imin*(n1+n2)-imin)^(2));
pval=pval+(ADart.>AD)/reps;
rep=rep+1; endo;

/* screen output */ output file=test.out; output on; print "number of ob-
servations " n1 " " n2; ? ;
print "AD statistic " AD; ? ; print " approx p-value " pval;
print " output saved in TEST.OUT ";output off;
```

11.2.3 Epps-Singleton test - my Matlab code

```
X1=sample1;
X2=sample2;
X1=sort(X1);
X2=sort(X2);
```

```

X12=[X1;X2];
X12=sort(X12);
n1=size(X1,1);
n2=size(X2,1);
n12=n1+n2;
Lreal=n12/4;
L=fix(Lreal);
U=n12-L;
Yu=X12(n12-U,1);
Yu1=X12(n12-U+1,1);
Yl=X12(n12-L,1);
Yl1=X12(n12-L-1,1);
fi=0.5*((Yl+Yl1)/2-(Yu+Yu1)/2);
t1=0.4/fi t2=0.8/fi;
g1=[cos(t1*X1) sin(t1*X1) cos(t2*X1) sin(t2*X1)]';
g2=[cos(t1*X2) sin(t1*X2) cos(t2*X2) sin(t2*X2)]';
g_11=sum(g1');
g_22=sum(g2');
g_1=n1^(-1)*sum(g1');
g_2=n2^(-1)*sum(g2');
s1=g1*g1'-n1*(g_1'*g_1);
s2=g2*g2'-n2*(g_2'*g_2);
s11=1/n1*s1;
s22=1/n2*s2;
omega=(s11+s22)*n12*(1/n1+1/n2)/2;
C=(1+(n1+n2)^(-0.45)+10.1*(n1^(-1.7)+n2^(-1.7)))^(-1);
if n1>=50 & n2>=50
Epps=n12*(((g_1'-g_2')')*(omega)^(-1)*(g_1'-g_2'))
Prob=1-chiSquareProb(4,0,Epps);
else
Epps=n12*(((g_1'-g_2')')*(omega)^(-1)*(g_1'-g_2'))*C
Prob=1-chiSquareProb(4,0,Epps)
end

```

11.3 Overview experiments

Most of the table's content is self-explaining. Still, there are some remarks to be made: In case there was more than 1 treatment, there is one main row for the experiment (with reference to the section). The rows in italics below refer to the treatments. If there are columns with no entry, there are two possibilities: Either the entry is in the main row for the experiment and then holds for all treatments. Or there are differences between treatments and then the entry in the main row will be empty while there will be entries in the rows for the treatment. In some experiments, there were more than 1 observation per subject, in some experiments, only some observations were paid. To inform about this, I use the column with '#Obs'/'#Subj.'/'#Obs paid'. '#Obs' represents the overall number of observations, '#Subj.' refers to the number of subjects who allocated and '#Obs paid' stands for the number of observations where subjects received actual payments for their allocation choice. Hence, 100/10/10 means that there were 10 proposers who had to make 10 proposals each and 1 out of 10 proposals was paid. The column '#Sess' stands for number of sessions of the experiment since most experiments were run on more than one day. For the test-statistics, which were applied for the analysis, I used the following abbreviations: 'FE': Fisher exact test (tests for differences of proportions), 'Sp': Spearman's rank Correlation test (non-parametric correlation test), 'Pr': Probit regression (binary maximum likelihood regression), ' χ^2 ': χ^2 -test (contingency table test to check for sample differences), and Jonckheere (non-parametric order test). In case there is a '?' as entry, the information was not mentioned in the article. In case there is a '-', the experimental setup makes the use of this category impossible. For the pie-size, the amounts being combined with ' \approx ' refer to a pie-size that was in another currency. To ensure comparability, I converted these in US-\$.

11.3.1 Ultimatum games

Experiment <i>Treatment</i>	Pie-size	Show-up fee	#Obs/ #Obs paid	#Subj./	#Sess.	Test-stat	\bar{O} -share passed
2.1.1 Güth <i>et al.</i> (1982)	$\approx 2\$-4\$$?	21/21/3	1	1	-	0.35
2.1.1 Forsythe <i>et al.</i> (1994)	5\$	3\$	43/43/43	2	2	AD,ES	0.46
2.1.1 Bolton and Zwick (1995)	4\$	0	100/10/100	1	1	FE	-
2.1.2 Roth <i>et al.</i> (1991)	3\$	3\$?	?	?	MW	?
<i>USA</i>	5\$		27/27/27				0.45
<i>Israel</i>	5\$		30/30/30				0.36
<i>Japan</i>	5\$		29/29/29				0.45
<i>Yugoslavia</i>	0\$		30/30/30				?
2.1.2 Henrich (2000)	?	?	?	?	?	ES,MW	?
<i>Los Angeles</i>	160\$		15/15/15				0.48
<i>Machiguenga</i>	$\approx 160\$$		21/21/21				0.26
2.1.3 Sanfey <i>et al.</i> (2003)	10\$?	190/19/19	19	19	-	-
2.1.4 Wallace <i>et al.</i> (2007)	15\$	0\$?	?	?	Sp,Pr	-
<i>Monozygotic</i>			511/511/511				
<i>Dizygotic</i>			142/142/142				
2.1.5 Güth and van Damme (1998)	$\approx 6.80\$$?	?	?	?	-	-
<i>Full information</i>			180/24/180	6	6		
<i>Essential information</i>			180/24/180	6	6		
<i>Irrelevant information</i>			180/24/180	6	6		

11.3.2 Dictator games

Experiment <i>Treatment</i>	Pie- size	Show-up fee	#Obs/ #Obs paid	#Subj./ #Sess.	Test-stat	\bar{O} -share passed	Share passing 0
2.2.1 Kahneman <i>et al.</i> (1986)		?			-	-	
<i>Experiment 1</i>	20\$		161/161/8	2			
<i>Experiment 2</i>	-		32/32/32	1			
2.2.2 Forsythe <i>et al.</i> (1994)		3\$			AD,ES		
<i>Pay</i>	5\$		45/45/45	6		0.22	0.36
<i>No Pay</i>	5\$		46/46/46	6		0.42	0.13
<i>Pay</i>	10\$		24/24/24	3		0.23	0.21
<i>Pooled Pay</i>	5\$,10\$		69/69/69	9		0.23	0.30
2.2.3 Sefton (1992)	5\$	3\$			CM,KS		?
<i>Full pay</i>	5\$		24/24/24	3		0.22	
<i>Random pay</i>	5\$		24/24/8	3		0.37	
2.2.4 Hoffman <i>et al.</i> (1994)	10\$						
<i>Random exchange</i>		3\$	24/24/24	4	MW, ES	0.27	0.21
<i>Contest exchange</i>		3\$	24/24/24	4	MW, ES	0.13	0.41
<i>Double blind 1</i>		5\$	36/36/36	3	MW, ES	0.09	0.64
<i>Double blind 2</i>		5\$	41/41/41	4	MW, ES	0.11	0.59
<i>Single blind 1</i>		5\$	37/37/37	3	Jonchkeere	?	?
<i>Single blind 2</i>		5\$	43/43/43	4	Jonchkeere	?	?
<i>FHSS-V</i>		5\$	28/28/28	2	Jonchkeere	?	?
<i>FHSS-R</i>		5\$	28/28/28	2	Jonchkeere	?	?

Experiment	Pie-size	Show-up fee	#Obs/ #Obs paid	#Subj./ #Sess.	Test-stat	Ø-share passed	Share passing 0
<i>Treatment</i>							
2.2.5 Bolton and Zwick (1995)	4\$	0\$	100/10/100	10	FE	-	-
2.2.5 Bolton <i>et al.</i> (1998)		5\$?	MW,FE, χ^2	-	-
<i>1Game6Card</i>	10\$		27/27/27			-	0.56
<i>1Game2Card</i>	10\$		28/28/28			-	0.93
<i>Anonymity</i>	10\$		33/33/33			-	?
2.2.6 Bohnet and Frey (1999)	\approx 10\$	\approx 5\$?	MW,KS,FE	?	
<i>Double blind 2</i>			39/39/39				0.28
<i>1-way ident.</i>			18/18/18				0.11
<i>1-way ident. with info</i>			25/25/25				0
<i>2-way ident.</i>			24/24/24				0
2.2.7 Frohlich <i>et al.</i> (2001)	5\$	5\$			t-test,Pr		
<i>Two room</i>			41/41/41	3		0.22	0.41
<i>One room</i>			17/17/17	1		0.25	0.35
2.2.8 Johannesson and Persson (2000)	\approx 12\$	\approx 6\$			MW,KS, χ^2		
<i>Baseline - double blind 1</i>			21/21/21	1		0.13	0.67
<i>General population</i>			54/54/54	2		0.09	0.68
2.2.8 Charness and Gneezy (2003)	\approx 16\$?		?	MW,KS		
<i>Baseline</i>			30/30/30			0.18	0.43
<i>Name</i>			30/30/30			0.27	0.27

Experiment	Pie-size	Show-up fee	#Obs/ #Obs paid	#Subj./ #Sess.	Test-stat	Ø-share passed	Share passing 0
<i>Treatment</i>							
2.2.8 Burnham (2003)	5\$	5\$			MW		
<i>Double blind 1</i>			26/26/26	2		0.24	0.54
<i>Recipient photo</i>			24/24/24	2		0.39	0.58
<i>Dictator photo</i>			24/24/24	2		0.34	0.54
2.2.8 Brañas (2006)	5\$	2\$			MW, KS		
<i>Double blind 1</i>			27/27/27	1		0.25	0.19
<i>Friends</i>			22/22/22	1		0.37	0.14
2.2.9 Goeree <i>et al.</i> (2007)	6\$	2\$	800/80/80	8	Regression	-	-
2.2.9 Leider <i>et al.</i> (2007)	5\$	18\$		-	Tobit	-	-
<i>Nameless</i>			1158/193/193				
<i>Social distance</i>			905/181/181				
2.2.10 Haley <i>et al.</i> (2005)	10\$	5\$			MW		
<i>Baseline</i>			22/22/22	1		0.25	0.47
<i>Eyepots</i>			25/25/25	1		0.38	0.21
2.2.10 Brañas and Morales (2005)	-	-		1	MW, KS		
<i>Baseline</i>			20/20/20			0.12	0.55
<i>Helping</i>			20/20/20			0.31	0.1
2.2.10 Dufwenberg and Muren (2006)	110\$	0\$	352/352/2	?	MW, χ^2		
2.2.11 Eckel and Grossman (1998)	5\$	3\$			ES		
<i>Replication</i>			12/12/12	1		0.15	0.58
<i>Charity</i>			48/48/48	4		0.30	0.27

Experiment <i>Treatment</i>	Pie- size	Show-up fee	#Obs/ #Obs paid	#Subj./ #Sess.	Test-stat	Ø-share passed	Share passing 0
2.2.11 Ruffle (1998)		3\$?	ES		
<i>Coin</i>	4\$		25/25/25			0.31	0.36
<i>Coin</i>	10\$		27/27/27			0.34	0.30
<i>Skill</i>	4\$		28/28/28			0.23	0.46
<i>Skill</i>	10\$		28/28/28			0.45	0.07
2.2.11 Cherry <i>et al.</i> (2003)		5\$?	MW, FE	?	
<i>Baseline</i>	10\$		26/26/26				0.19
<i>Baseline</i>	40\$		26/26/26				0.15
<i>Earning</i>	10\$		28/28/28				0.79
<i>Earning</i>	40\$		33/33/33				0.70
<i>Double blind with earning</i>	10\$		25/25/25				0.95
<i>Double blind with earning</i>	40\$		36/36/36				0.97
2.2.11 List and Cherry (2004)		?		3	MW, FE		
<i>Baseline</i>	20\$		20/20/20			0.17	0.50
<i>Baseline</i>	100\$		20/20/20			0.13	0.55
<i>Symmetric - successful</i>	20\$		34/34/34			0.04	0.92
<i>Symmetric - unsuccessful</i>	100\$		43/43/43			0.04	0.88
<i>Asymmetric - successful</i>	20\$		37/37/37			0.19	0.71
<i>Asymmetric - unsuccessful</i>	100\$		41/41/41			0.09	0.58

Experiment	Pie-size	Show-up fee	#Obs/ #Obs paid	#Sess.	Test-stat	Ø-share passed	Share passing 0
2.2.12 Carpenter <i>et al.</i> (2005)	100\$	10\$?	MW, KS	?		
<i>Affluent students</i>			21/21/21			0.25	
<i>Older and less affluent students</i>			26/26/26			0.33	
<i>Workers</i>			37/37/37			0.45	
2.2.13 Andreoni and Miller (2002)	-	?	176/1936/176	5	-	-	-
2.2.13 Charness and Rabin (2002)	-		-	-	-	-	-
<i>Barcelona</i>		≈ 6\$					
<i>Berkeley</i>		11\$					
2.2.14 Ben-Ner <i>et al.</i> (2004)	10\$	15\$?	Regression			
<i>Same match</i>			≈50/≈ 50/≈50			0.28	0.29
<i>Random match</i>			≈50/≈ 50/≈50			0.25	0.30
2.2.14 Cason and Mui (1998)	40\$	5\$?	KS		?	
<i>1st Stage Relevant Info</i>			40/40/40				0.29
<i>2nd Stage Relevant Info</i>			40/40/40				0.29
<i>1st Irrelevant Info</i>			20/20/20				0.30
<i>2nd Irrelevant Info</i>			20/20/20				0.30
2.2.14 Karni <i>et al.</i> (2001)	15\$	10\$	45/280/45	?	MW	-	-

Experiment	Pie-size	Show-up fee	#Obs/ #Obs paid	#Subj./ #Sess.	Test-stat	Ø-share passed	Share passing 0
2.2.15 Andreoni and Bernheim (2007)	20\$	5\$	60/266/60		Probit	?	
<i>Treat. 0, prob. nature 0</i>			30/30/?	3			0.3
<i>Treat. 0, prob. nature 0.25</i>			≈45/≈45/?	3			0.53
<i>Treat. 0, prob. nature 0.50</i>			30/30/?	3			0.72
<i>Treat. 0, prob. nature 0.75</i>			30/30/?	3			0.70
<i>Treat. 1, prob. nature 0</i>			29/29/?	3			0.18
<i>Treat. 1, prob. nature 0.25</i>			≈44/≈44/?	3			0.14
<i>Treat. 1, prob. nature 0.50</i>			29/29/?	3			0.14
<i>Treat. 1, prob. nature 0.75</i>			29/29/?	3			0.10
2.2.16 Bardsley et al. (2008)					FE	-	
<i>Treatment 1</i>	-	-	31/31/31	?			0.55
<i>Treatment 2</i>	-	-	29/29/29	?			-
2.2.16 List (2007)	5\$	5\$			FE	-	
<i>Treatment 1</i>			24/24/24	3		0.27	0.29
<i>Treatment 2</i>			46/46/46	6		33	0.65
<i>Treatment 3</i>			50/50/50	6		-0.50	0.90
<i>Treatment 4</i>			47/47/47	6		-0.20	0.94

Experiment	Pie-size	Show-up fee	#Obs/ #Obs paid	#Subj./ #Sess.	Test-stat	Ø-share passed	Share passing 0
<i>Treatment</i>					ES,KS, χ^2		
2.3 Eckel and Grossman (1998)							
<i>Women</i>			60/60/60	5		0.16	0.47
<i>Men</i>			60/60/60	5		0.08	0.60
2.3 Andreoni and Vesterlund (2001)	-	-		4	-	-	-
<i>Women</i>			47/376/47				
<i>Men</i>			95/760/95				
2.4 Carter and Irons (1991)	10\$	2\$	46/46/46	?	OLS		-
<i>Freshmen Non-Economists</i>						0.57	
<i>Freshmen Economists</i>						0.63	
<i>Senior Non-Economists</i>						0.52	
<i>Senior Economists</i>						0.60	
2.4 Gross (2005)	10\$?		?	-	-	-
<i>Non-Economists</i>			24/24/24				0.75
<i>Economists</i>	-	-	61/61/61				0.39
2.4 Meier and Frey (2004)			180225/180225/-	-	Probit	-	-

Experiment <i>Treatment</i>	Pie- size	Show-up fee	#Obs/ #Obs paid	#Subj./ #Sess.	Test-stat	Ø-share passed	Share passing 0
2.5.1 Rutte <i>et al.</i> (1987)	-	-			-		?
<i>First</i>			18/18/6	3		0.56	
<i>Second</i>			18/18/6	3		0.49	
2.5.2 Blount (1995)	10\$?		?	MW	-	-
<i>Baseline</i>			17/17/17				
<i>Third party</i>			18/18/18				
2.5.3 Offerman (2002)	-	0\$		8	MW	-	-
<i>Nature</i>			28/28/28				
<i>Hot flesh</i>			28/28/28				
2.6.1 Chermak and Krause (2002)	-	3\$			-	-	-
<i>Informed</i>			35/30/20	2			
<i>Uninformed</i>			45/47/28	3			
2.6.2 Sadrieh (2003)	-	-		1	-	-	-
<i>Very Slow</i>			24/24/24				
<i>Very Fast</i>			25/25/25				
2.6.3 Fischer <i>et al.</i> (2004)	-	-		1	-	-	-
<i>Restart</i>			48/48/48				
<i>Slow</i>			48/48/48				
<i>Fast</i>			48/48/48				

11.4 Some treatment comparison results for B

Table 11.1: Treatment comparison B for ‘uni-gender’ subsamples

	Treatments compared	MW	AD	ES
The effect of social interaction	Males 1B vs 2B	-1.668 (0.095)	2.371 (0.056)	7.848 (0.097)
	Females 1B vs 2B	0.212 (0.832)	0.452 (0.819)	4.256 (0.373)
	Males 1B vs 3B	-1.985 (0.047)	2.089 (0.079)	9.348 (0.053)
	Females 1B vs 3B	-0.839 (0.402)	0.567 (0.701)	2.677 (0.613)
	Males 2B vs 4B	-3.453 (0.001)	6.634 (0.000)	16.323 (0.003)
	Females 2B vs 4B	-2.387 (0.017)	2.970 (0.025)	8.382 (0.079)
	Males 3B vs 4B	-3.500 (0.001)	7.455 (0.000)	24.258 (0.000)
	Females 3B vs 4B	-3.245 (0.001)	5.374 (0.001)	20.021 (0.000)
	Possibility of a non-human allocator	Males 1B vs 4B	-1.143 (0.253)	3.276 (0.018)
Females 1B vs 4B		-2.412 (0.016)	3.171 (0.020)	8.756 (0.068)
Males 4Bself vs 4Bdice		-0.596 (0.551)	0.885 (0.428)	5.444 (0.245)
Females 4Bself vs 4Bdice		0.684 (0.494)	0.305 (0.955)	1.295 (0.862)
Social interaction before or after choice	Males 2B vs 3B	0.635 (0.525)	0.894 (0.422)	2.235 (0.693)
	Females 2B vs 3B	1.2963 (0.207)	1.693 (0.133)	4.768 (0.312)
Extension to three persons	Males 1B vs TDG	-0.426 (0.670)	0.849 (0.453)	6.297 (0.178)
	Females 1B vs TDG	-0.369 (0.712)	0.355 (0.915)	0.298 (0.990)

Table 11.2: Comparison of A 's and B 's allocation choice in uni-gender subsamples within treatments

Treatments	MW	AD	ES
Males 1	1.071 (0.281)	1.686 (0.138)	11.497 (0.022)
Females 1	-0.709 (0.479)	1.205 (0.264)	8.290 (0.082)
Males 2	1.915 (0.056)	4.103 (0.007)	15.172 (0.004)
Females 2	-3.032 (0.002)	5.301 (0.001)	23.112 (0.000)
Males 3	2.183 (0.029)	3.162 (0.022)	16.915 (0.002)
Females 3	-2.2788 (0.005)	5.175 (0.001)	17.739 (0.001)
Males 4	-1.541 (0.123)	2.064 (0.082)	4.918 (0.296)
Females 4	-2.356 (0.019)	2.817 (0.031)	5.497 (0.240)

11.5 Results for the subsample of B passing on a positive amount

Table 11.3: Allocation choice B excluding observations where B passed zero

Treatment	Mean Share B
1B	0.40
2B	0.38
3B	0.42
4B	0.28
4Bself	0.28
4Bdice	0.28

Table 11.4: Treatment comparison B s excluding observations where B passed zero

	Treatments compared	MW	AD	ES
The effect of social interaction	1B vs 2B	-0.795 (0.427)	0.702 (0.562)	3.133 (0.536)
	1B vs 3B	0.947 (0.344)	1.263 (0.243)	1.547 (0.818)
	2B vs 4B	-2.606 (0.009)	3.145 (0.022)	7.252 (0.123)
	3B vs 4B	-4.306 (0.000)	8.605 (0.000)	25.550 (0.000)
Possibility of a non-human allocator	1B vs 4B	-3.570 (0.000)	6.050 (0.001)	18.522 (0.001)
	4Bself vs 4Bdice	-0.334 (0.738)	0.949 (0.390)	3.632 (0.458)
Social interaction before or after choice	2B vs 3B	-1.617 (0.106)	1.924 (0.098)	4.375 (0.358)
B s choice vs. traditional dictator	1B vs TDG	1.416 (0.157)	1.480 (0.181)	4.358 (0.360)
	1B vs FHSS	-2.192 (0.028)	3.044 (0.025)	16.664 (0.002)
	1B vs HMSS	2.273 (0.024)	3.061 (0.023)	10.062 (0.039)

Table 11.5: Treatment comparison of A 's and B 's allocation choice excluding observations where B passed zero

Treatments	MW	AD	ES
1	0.649 (0.516)	3.036 (0.025)	22.042 (0.000)
2	4.470 (0.000)	10.871 (0.000)	34.255 (0.000)
3	3.965 (0.000)	7.390 (0.000)	34.376 (0.000)
4	-2.442 (0.015)	3.155 (0.022)	7.752 (0.101)

Table 11.6: Treatment comparison of A 's and B 's allocation in uni-gender samples excluding observations where B passed zero

Treatments	MW	AD	ES
Males 1	0.467 (0.641)	1.483 (0.178)	13.684 (0.008)
Females 1	0.441 (0.659)	1.461 (0.183)	9.188 (0.057)
Males 2	3.224 (0.001)	6.422 (0.000)	14.526 (0.006)
Females 2	-3.032 (0.002)	5.301 (0.001)	23.112 (0.000)
Males 3	2.873 (0.004)	4.070 (0.006)	17.760 (0.001)
Females 3	-2.788 (0.005)	5.175 (0.001)	17.739 (0.001)
Males 4	-2.042 (0.041)	2.797 (0.032)	4.951 (0.292)
Females 4	-1.275 (0.202)	1.003 (0.357)	2.201 (0.699)

Table 11.7: Treatment comparison B s in uni-gender samples excluding observations where B passed zero

	Treatments compared	MW	AD	ES
The effect of social interaction	Males 1B vs 2B	-0.310 (0.757)	0.994 (0.361)	2.652 (0.618)
	Females 1B vs 2B	-0.894 (0.371)	0.873 (0.437)	3.003 (0.557)
	Males 1B vs 3B	0.862 (0.389)	0.745 (0.532)	7.991 (0.092)
	Females 1B vs 3B	-0.295 (0.768)	0.265 (0.981)	1.353 (0.852)
	Males 2B vs 4B	-2.736 (0.006)	4.476 (0.004)	13.828 (0.008)
	Females 2B vs 4B	-0.899 (0.369)	0.569 (0.692)	1.641 (0.801)
	Males 3B vs 4B	-3.689 (0.000)	7.762 (0.000)	28.854 (0.000)
	Females 3B vs 4B	-2.214 (0.027)	2.756 (0.032)	9.298 (0.054)
Possibility of a non-human allocator	Males 1B vs 4B	-2.993 (0.003)	5.543 (0.001)	13.620 (0.009)
	Females 1B vs 4B	-1.944 (0.052)	2.146 (0.072)	6.379 (0.173)
	Males 4Bself vs 4Bdice	-0.411 (0.681)	1.252 (0.247)	5.557 (0.235)
	Females 4Bself vs 4Bdice	0 (1)	0.202 (0.999)	0.902 (0.924)
Social interaction before or after choice	Males 2B vs 3B	1.354 (0.176)	1.445 (0.189)	1.423 (0.840)
	Females 2B vs 3B	-1.263 (0.207)	1.693 (0.133)	4.768 (0.312)
B s choice vs. traditional dictator	Males 1B vs TDG	-1.535 (0.125)	1.934 (0.096)	7.360 (0.118)
	Females 1B vs TDG	0.076 (0.939)	0.265 (0.980)	4.342 (0.362)

Table 11.8: Testing for the impact of gender on B 's allocation choice excluding observations where B passed zero

Treatment	MW	AD	ES
1B	0.060 (0.952)	0.789 (.498)	2.338 (0.674)
2B	0.451 (0.652)	0.589 (0.673)	2.831 (0.587)
3B	0.788 (0.431)	1.230 (0.257)	9.567 (0.048)
4B	-1.289 (0.197)	1.375 (0.208)	2.550 (0.636)
4BDice	-1.083 (0.279)	0.800 (0.491)	2.208 (0.698)
4BSelf	0.096 (0.923)	0.884 (0.432)	2.019 (0.732)

11.6 Presentation of the categories for the qualitative analysis to explain allocation behavior

11.6.1 *A*'s explanations

1. 'Selfishness' refers to explanations including claims like: 'Selfishness', 'I wanted most of the money for me', 'I think of my personal welfare first'. Note, that an explanation using 'selfishness' does not necessarily mean that the subject took the whole pie, there were also subjects who took little more than one third and still justified their behavior with selfishness. Possibly these subjects would like to stress that their own payoff is one argument of their utility function but that there are other arguments like e.g. altruism.

2. 'Fairness' refers to explanations including claims like: 'Fairness', 'I wanted to achieve an equal distribution of the pie'. The latter was only used by subjects taking one third of the pie while the former also includes subjects who took more than one third. As in the case for 'selfishness', these subjects possibly wanted to highlight that 'fairness' is one argument of their utility function.

3. '*B* and/or *C* should receive > 0 ' refers to explanations including claims like: 'I chose the amount in this way to gain a large amount for me without leaving the others with nothing.' These subjects hence indicated that they wanted to pass some money to *B* and/or *C* so that everybody earned a positive amount. Unlike in the category 'Fairness', it seems likely that subjects receive a 'warm glow of giving' according to the model by Sadrieh (2003).

4. 'Luck' refers to explanations including claims like: 'Because it was an experiment and none had any costs and it was chance, in which group you were drawn, I decided, that it is an adequate amount (...)' These subjects stressed that it was due to chance or luck that they became *A* or they referred to the randomness of the assignment process as explanation. Subjects used it either to explain taking more than one third, in this case they said that they would like to exploit their luck or they interpreted randomness in a way that they have not made any effort to be in the position and hence passed on a fair share.

5. 'Personal neediness' refers to explanations including claims like: 'I do not have got that much money (...)', '(...) my financial position played a role'. Subjects either used their personal income situation declaring themselves as poor or not having much money or they stated that they needed money.

Sometimes these explanations were followed by very specific details for what purpose the money was needed, e.g. to buy a used tricycle for her daughter, to pay a fine for speeding or to pay the repair of a bicycle.

6. ‘Hope to be role model for *B*’ refers to explanations including claims like: ‘I have divided the 30€ in 10€ for me and 20€ for the other candidates, in the hope that candidate *B* will give 10€ to candidate *C*, too (...)’ Subjects in this category tried to induce positive reciprocity. They specifically mentioned that they hoped that *B* would imitate their fair behavior.

7. ‘Anonymity’ refers to explanations including claims like: ‘Anonymity’, ‘I do not know the others’. Anonymity is used in two different senses: Almost all (for treatment 1 and 4) subjects use it in the sense that it is completely unclear who *B* and *C* will be. But there is also one subject in treatment 2 who used anonymity as explanation: In this case it must be understood in the way that the subject did not personally know the other subject though it met the subject during the interaction of the experiment.

8. ‘Others would have done the same’ refers to subjects who claimed that if someone else had been in their position she would have made the same decision.

9. ‘Power’ refers to explanations including claims like: ‘I divided the amount of money in 15€ for me and 15€ for *B* and *C*, because the one with the highest decision power (hence *A*) obtains the highest amount of money in any case (...)’ or ‘I wanted to exploit my privileged position(...)’ One of the subjects, a political science student, even mentioned Machiavelli.

10. ‘Unclear answer’ refers to explanations including claims like: ‘If I divided the amount into 10€ : 10€ : 10€ this result would not reflect the reality.’ ‘I wanted to have 15€ for me and 15€ for *B* and *C*. But I love prime numbers. Hence, I took the larger number (17€) for me and passed on the smaller number.’ ‘Why not? A profit distribution is constructed in a similar way 1 : 2/3.’ ‘I am not entirely sure. Theoretically, I do not know it at all, I just followed a feeling’ ‘22 is my jersey number. Besides, we have an even number. Additionally, I am 22.’ Hence, the explanations in this categories did not make sense or were idiosyncratic.

11. ‘*B* plays twice’ refers to explanations including claims like: ‘I gave myself 12€ because *B* has the chance to earn twice playing scrabble twice. I hope that *B* passes on 12€ to *C*, then everybody leaves the experiment with the most equitable amount of money.’ Most of the subjects use the explanation in line with the first example hence specifically mention their desire to obtain an equitable outcome and hence revealing a preference structure according to the equity model by Fehr and Schmit (1998) and Bolton and Ockenfels (1998).

But there were also few subjects taking e.g. amounts of larger than 15€ up to 30€ and mentioning that *B* plays twice (mostly, this was accompanied by other explanations from other categories).

12. ‘Norm (intrinsic)’ refers to explanations including claims like: ‘I divided the amount fairly in my opinion. (...) I would have a bad consciousness, otherwise.’ Hence, subjects considered their allocation (mostly they took one third) as a moral obligation. Interestingly, the explanation does not so much refer to the desire to be altruistic but to avoiding to have a bad consciousness.

13. ‘Same effort’ refers to explanations including claims like: ‘(...). Something different would be unfair, since I have not made a bigger effort than the others (...)’ These subjects claim that overall, *A*, *B* or *C* had to make the same effort during the experiment (which was not true, since *B* played the scrabble game twice and the amount of time that had to be spent was different) and hence do deserve a roughly equal amount of the pie. These subjects probably consider a concept of wage per hour worked. Mostly, *A* took one third in this case.

14. ‘Participation to earn’ refers to explanations including claims like: ‘(...) I was here for the money’. These explanations refer so subjects who specifically claimed that they had a financial motivation to participate in the experiment (though also the ones not mentioning this had this reason). This explanation was used to justify a choice were *A* took more than one third. Possibly, subjects wanted to indicate that they still have a desire to be fair but that under this setup fairness was not so important for them.

15. ‘Rationality’ refers to explanations including claims like: ‘I divided the amount accordingly because it would not be rational to pass on money to people I do not know, (...)’ These explanations specifically mention the concept of rationality in the economic sense. Interestingly, only one of the five subjects using this explanation studied economics.

16. ‘Maintain incentive’ refers to explanations including claims like: ‘(...) I wanted to ensure that my partner is motivated in order to perform better in constructing words in the second part’. These subjects saw a fair or close to fair allocation as incentive compatible behavior. The idea was that if *A* took a large share for himself that could be perceived as unfair by *B* and hence *B* could reciprocate by not cooperating (even though *B* would forfeit potential gain). This explanation was only used in treatment 2, since it only made sense to use this argument if you allocated before you played.

17. ‘Norm’ refers to explanations including claims like: ‘Taking more would be inadequate’, ‘I did not deserve more’, ‘The others should not receive zero and e.g. (passing) 1 Euro would have been pretty mean’. Explanations in

this category refer to subjects who claim that with their allocation choice they followed a moral obligation. In this category, unlike the category ‘internal norm’, it is unclear if the subjects refer to their own perception or to the perception of the society.

18. ‘Maintain incentive not important’ refers to explanations including claims like: ‘I kept 30€ for myself. I try to maximize my payoff. The 30 I have for sure. In case *B* was sour and refused to play with me, I would play alone. Even if I only manage to construct few words I will have the highest payoff among *A*, *B* and *C*.’, ‘There were at the maximum 5€ to win in the letter game, hence negligible. So I kept the whole 30€. I will not play again with the other people and hence cannot lose anything with respect to the future’. Thus, these subjects argue that even if *B* reciprocated on *A*’s unfair behavior by refusing to cooperate in the scrabble game, this would have almost no impact on *A*’s payoff due to the very low amount that could be earned in the scrabble game.

19. ‘*C* should receive >0’ refers to explanations including claims like: ‘(...). And then *C* hopefully will receive a small sum, too.’ Hence, these subjects explicitly considered *C* in their reasoning. For both subjects, who mentioned this idea, this was not the only explanation.

20. ‘*C* does not know’ refers to an explanations including claims like: ‘(...)
B will keep the entire amount left for himself in any case and tell *C* that *A* did not leave anything’

21. ‘Did not get to know *C*’ refers to explanations including claims like: ‘I did not get to know *C*. Hence, I give *B* the fair share of 15€ and then *B* can see if she would like to pass something on to *C*.’

22. ‘Others are not needy’ refers to explanations including claims like: ‘Students’ existence is normally not in danger’ This reasoning is based on the idea that subjects, who participate in these experiments, are usually students whose basic needs are covered.

23. ‘Rules of game accepted’ refers to explanations including claims like: ‘(...). I chose this, because all participants knew the experimental setup and agreed on it.’ These subjects claim that subjects *B* and *C* were aware of the experimental setup and that they knew they may receive an amount which is smaller than the fair share. Maybe the subjects referred also to the announcements in the student cafeteria, where I told the potential participants that they would earn between 3 to 35€ for participation. Maybe the subjects derived from that that it was be morally justifiable to pass only a small amount or nothing due to this.

24. ‘Reward for cooperation’ refers to an explanation by one subject: ‘(...)

I did not take the whole amount, because I did not wanted to fleece B of his share. The cooperation with him was very successful (...)' . Hence, this subject wanted to reward the other subject for the successful cooperation.

25. 'Have to play vis-à-vis with B ' refers to an explanation by one subject: '(...) Because B will personally sit in front of me during the game, maybe I was too shy to be more greedy.'

11.6.2 B 's explanations

A lot of explanations that A used are similar to the ones used by B . But there are some additional explanations by B which would not make sense if A used them.

26. 'Reciprocity' refers to explanations including claims like: ' A was so nice to pass on 20 Euro to me and C , hence I divided the overall money fairly. I was very happy about this and was of the opinion that C would be as glad as me if I behaved in the same way A did...' or 'I considered the amount of money A left to us ridiculous and mean, that is why I reacted mean and ripped off C ...'. There were also few subjects which imitated A 's behavior by taking the same share of the left money: '8€ for me and 5€ for C because A took also 17/30 of the overall money.' 'One adapts to one's environment: An allocation of 75% to 25% resembles roughly the one by A '. Hence, this category captures all claims by B who said that A 's behavior influenced their allocation decision. I also reported the relative shares for positive and negative reciprocity which are subsets of the set 'reciprocity'.

27. 'Not much left to pass' refers to explanations including claims like: 'There was not much left to pass, hence I took the whole amount'.

28. 'Wanted certain amount of money' refers to explanations including claims like: '(...) Average payoff of the experiment: 13€. I wanted to ensure to obtain this amount. I would have passed more money if possible.' These subjects B claim that they had an ex-ante expectation of how much money they would like to have for themselves. This argument was often used to justify an unfair division of $(1 - \alpha)$ if A passed less than two thirds of the pie.

29. ' B played twice' refers to explanations including claims like: '(...) I have played twice (...)' The argument that B played twice was used by B too, but mostly in another sense: Now B used it in her favor claiming that having to play twice is twice the effort and hence entitles to more payoff.

30. 'Got to know each other' refers to explanations including claims like: 'Because I got to know him I thought it would be adequate to pass on a certain amount.'

31. ‘Compassion’ refers to explanations including claims like: ‘Due to compassion for *C* I did not take all’ Subjects *A* did not use this explanation. The reason for this is possibly that *B* could directly see how much *C* would receive. They would feel sorry for *C* if she passed on nothing.

11.6.3 *A*’s explanations in TDG

In the replication treatment, the subjects used some new explanations. Actually, with regard to number of candidates (40), there were surprisingly many new explanations.

33. ‘S.th. unfair in real life’ refers to explanations including claims like: ‘Someone stole money from me and hence I am unfair, too, now.’ Hence, this explanations refer to subjects who claimed that something unfair happened to them, recently, and that is why they were unfair, too. That would again be a claim for the social history effect on an even more abstract level.

34. ‘Adequate pay *B*’ refers to explanations including claims like: ‘*B* should be rewarded adequately for his effort’. One added, that the amount passed on would have been satisfying for her if she were *B*.

35. ‘Positive empathy’ refers to explanations including claims like: ‘If I were *B* I would be in the need of *A*’s benevolence’ or ‘If I were *B* I would be happy to receive some money, too’. Hence, these subjects empathized positively with *B*.

36. ‘Negative empathy’ refers to an explanation by one subject: ‘If I were *B* I would have expected a zero payoff.’ This subject took 30 €.

37. ‘Gift’ refers to an explanation by one subject: ‘I see it (the experiment) as a game, being *A* means I won, *B* just gets a consolation gift.’

38. ‘Christian values’ refers to an explanation by one subject: ‘I am obliged to Christian values.’ This subject split fairly.

Bibliography

- AFRIAT, SYDNEY, (1967). The construction of a utility function from expenditure data. *International Economic Review* 8, 67-77.
- ANDREONI, JAMES, AND BERNHEIM, DOUGLAS B., (2007). Social image and the 50-50 norm. A theoretical and experimental analysis of audience effects. *Working paper*.
- ANDREONI, JAMES, AND MILLER, JOHN, (2002). Giving according to Garp: An experimental test of the consistency of preferences for altruism. *Econometrica* 70, 737-753.
- ANDREONI, JAMES, AND VESTERLUND, LISA, (2001). Which is the fair sex? Gender differences in altruism. *The Quarterly Journal of Economics* 116, 239-312.
- BARDSLEY, NICHOLAS, (2008). Dictator game giving: Altruism or artefact? *Experimental Economics* 11, 122-133.
- BEN-NER, AVNER, PUTTERMAN, LOUIS, KONG, FANMIN, AND MANGAN, DAN, (2004). Reciprocity in a two-part dictator game. *Journal of Economic Behavior and Organization* 53, 333-352.
- BLOUNT, SALLY, (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes* 63, 131-144.
- BOHNET, IRIS, AND FREY, BRUNO, S., (1999). Social distance and other-regarding behavior in dictator games: Comment. *The American Economic Review* 89, 335-339.
- BOLLE, FRIEDEL, (1990). High reward experiments without high expenditure for the experimenter? *Journal of Economic Psychology* 11, 157-167.
- BOLTON, GARY E., AND KATOK, ELENA, (1995). An experimental test for gender in beneficent behavior. *Economic Letters* 48, 67-77.

- BOLTON, GARY E., KATOK, ELENA, AND ZWICK, RAMI, (1998). Dictator game giving: Rules of fairness versus acts of kindness. *International Journal of Game Theory* 27, 269-299.
- BOLTON, GARY E., AND OCKENFELS, AXEL, (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review* 90, 166-193.
- BOLTON, GARY E., AND ZWICK, RAMI, (1995). Anonymity versus punishment in ultimatum bargaining. *Games and Economics Behavior* 10, 95-121.
- BRAÑAS-GARZA, PABLO, DURÀN, MIGUEL A., AND ESPINOSA, MARIA PAZ, (2005). Do experimental subjects favor their friends? *Central European Working paper*, Universidad de Granada.
- BRAÑAS-GARZA, PABLO, MORALES, ANTONIO J., (2005) Moral framing in dictator games by short sentences. *Working paper*, Universidad de Granada.
- BRANDSTÄTTER, HERMANN, AND GÜTH, WERNER, (2002). Personality in dictator and ultimatum games. *Journal of Operations Research* 10, 191-215.
- BURNHAM, TERENCE C., (2003). Engineering altruism: A theoretical and experimental investigation of anonymity and gift giving. *Journal of Economic Behavior & Organization* 50, 133-144.
- BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG, (2005). Die wirtschaftliche und soziale Lage der Studierenden in der Bundesrepublik Deutschland 2003. 17. Sozialerhebung des Deutschen Studentenwerks durchgeführt durch HIS Hochschul-Information-System .
- CAMERER, COLIN, LOEWENSTEIN, GEORGE, AND PRELEC, DRAZEN, (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature* 43, 9-64.
- CAMERER, COLIN, AND THALER, RICHARD H., (1995). Anomalies: Ultimatums, dictators, manners. *The Journal of Economic Perspectives* 9, 209-219.
- CARPENTER, JEFFREY P., BURKS, STEPHEN, AND VERHOOGEN, ERIC, (2005). Comparing students to workers: The effects of social framing on the behavior in distribution games. *Working paper*.
- CARTER, JOHN R., AND IRONS, MICHAEL D., (1991). Are economists different, and if so, why? *Journal of Economic Perspectives* 5, 171-177.

- CASON, TIMOTHY N., AND MUI, VAI-LAM, (1998). Social influence in the sequential dictator game. *Journal of Mathematical Psychology* 42, 248-265.
- CHARNESS, GARY, AND GNEEZY, URI, (2008). What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization* 68, 29-35.
- CHARNESS, GARY, AND RABIN, MATTHEW, (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics* 117, 817-869.
- CHERMAK, JANIE M., AND KRAUSE, KATE, (2002). Individual response, information and intergenerational common pool problems. *Journal of Environmental Economics and Management* 43, 47-70.
- CHERRY, TODD L., FRYKBLOM, PETER, AND SHOGREN, JASON F., (2002). Hardnose the dictator. *American Economic Review* 92, 1218-1221.
- CONOVER, WILLIAM J., (1971). Practical nonparametric statistics. New York: Wiley.
- COSMIDES, LEDA, AND TOOBY, JOHN, (1992). Cognitive adaptations for social exchange. In Jerome H. Barkow, Leda Cosmides, and John Tooby, eds., *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, New York: Oxford University Press.
- DIEKMANN, ANDREAS, (2004). The power of reciprocity. Fairness, reciprocity, and stakes in variants of the dictator game. *Journal of Conflict Resolution* 48, 487-505.
- DUFWENBERG, MARTIN, AND KIRCHSTEIGER, GEORG, (1998). A theory of sequential reciprocity. *Discussion paper: 9837*, Tilburg Center for Economic Research.
- DUFWENBERG, MARTIN, AND MUREN, ASTRI, (2006). Generosity, anonymity, gender. *Journal of Economic Behavior & Organization* 61, 42-49.
- ECKEL, CATHERIN C., AND GROSSMAN, PHILIP J., (1996). Altruism in anonymous dictator games. *Games and Economic Behavior* 16, 181-191.
- ECKEL, CATHERIN C., AND GROSSMAN, PHILIP J., (1998). Are women less selfish than men? Evidence from dictator experiments. *The Economic Journal* 108, 726-735.

- FALK, ARMIN, AND FISCHBACHER URS, (1998). A theory of reciprocity. *Working paper 6*, Institute for Empirical Research in Economics, University of Zürich.
- FARRAR, DONALD E., AND GLAUBER, ROBERT R., (1967). Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics* 49, 92-107.
- FEHR, ERNST AND SCHMIDT, KLAUS, (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* 114, 817-868.
- FOOT, PHILIPPA, (1967). The problem of abortion and the doctrine of double effect. *Oxford Review* 5, 5-15.
- FISCHER, MARIA-ELISABETH, IRLENBUSCH, BERND, AND SADRIEH, ABDOLKARIM, (2004). An intergenerational common pool resource experiment. *Journal of Environmental Economics and Management* 48, 811-836.
- FORSYTHE, ROBERT, HOROWITZ, JOEL L., SAVIN, N. E., AND SEFTON, MARTIN, (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior* 6, (347-369).
- FROHLICH, NORMAN, OPPENHEIMER, JOE, AND MOORE, J. BERNARD, (2001). Some doubts about measuring self-interest using dictator-experiments: The cost of anonymity. *Journal of Economic Behavior & Organization* 46, 271-290.
- GOEREE, JACOB K., MCCONNELL, MARGARET A., MITCHELL, TIFFANY, TROMP, TRACEY, AND YARIV, LEEAT, (2007). Linking and giving among teenage girls. *Working Paper*.
- GREENE, WILLIAM H., (2003). *Econometric Analysis*, fifth edition, New York: Prentice-Hall.
- GROSS, LAUREN, (2005). Altruism, fairness and social intelligence: Are economicists different? *Working paper*.
- GÜTH, WERNER, SCHMITTBERGER, ROLF, AND SCHWARZE, BERND, (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization* 3, 367-388.
- GÜTH, W., AND VAN DAMME, E., 1998. Information, strategic behavior and fairness in ultimatum bargaining - an experimental study. *Journal of Mathematical Psychology* 42, 227-247.

- HALEY, KEVIN J., AND FESSLER, DANIEL M.T., (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior* 26, 245-256.
- HENRICH, JOSEPH, (2000). Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *The American Economic Review* 90, 973-979.
- HENRICH, JOSEPH, BOYD, ROBERT, BOWLES, SAMUEL, CAMERER, COLIN, FEHR, ERNST, GINTIS, HERBERT, AND MCELREATH, RICHARD (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *The American Economic Review Papers and Proceedings* 90, 73-78.
- HERWARTZ, HELMUT,(2007). A note on model selection in (time series) regression models - general-to-specific or specific-to-general? *Working Paper*.
- HOFFMAN, ELIZABETH, MCCABE, KEVIN, SHACHAT, KEITH, AND SMITH, VERNON, (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economics Behavior* 7, 346-380.
- HOFFMAN, ELIZABETH, MCCABE, KEVIN, AND SMITH, VERNON, (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review* 86, 653-660.
- JOHANNESSON, MAGNUS, AND PERSSON, BJÖRN, (2000). Non-reciprocal altruism in dictator games. *Economic Letters* 69, 137-142.
- KAHNEMAN, DANIEL, KNETSCH, JACK L., AND THALER, RICHARD H., (1986). Fairness and the assumptions of economics. *The Journal of Business* 59, 285-300.
- KAHNEMAN, DANIEL, AND TVERSKY, AMOS, (1974). Judgment under uncertainty: Heuristics and biases. *Science* 185, 1124-1131.
- KARNI, EDI, SALMON, TIM, AND SOPHER, BARRY, (2001). Individual sense of fairness: An experimental study. *Working Paper*.
- KNOBE, JOSHUA, (2003). Intentional action and side effects in ordinary language. *Analysis* 63, 190-193.
- KNOBE, JOSHUA, (2003). Intention, intentional action in folk psychology. *Philosophical Psychology* 16, 309-324.
- KNOBE, JOSHUA, (2006). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies* 130, 203-231.

- KNOBE, JOSHUA, AND BURRA, ARUDRA, (2006). The folk concepts of intention and intentional action: A cross-cultural study *Journal of Cognition and Culture* 6, 113-132.
- KNOCH, DARIA, PASCUAL-LEONE, ALVARO, MEYER, KASPAR, TREYER, VALERIE, AND FEHR, ERNST, (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex *Science* 314, 829-832.
- LEIDER, STEPHEN, MÖBIUS, MARKUS M., ROSENBLAT, TANYA, AND DO, QUOC-ANH, (2007). How much is a friend worth? Directed altruism and enforced reciprocity in social networks. *Working paper*.
- LESLIE, ALAN M., AND KNOBE, JOSHUA, (2006). Acting intentionally and the side-effect: Theory of mind and moral judgement. *Psychological Science* 17, 421-427.
- LIST, JOHN A., (2004). Young, selfish and male: Field evidence of social preferences. *The Economic Journal* 114, 121-149.
- LIST, JOHN A., (2007). On the interpretation of giving in dictator games. *Journal of Political Economy* 115, 482-493.
- LIST, JOHN A., AND CHERRY, TODD L., (2004). Examining the role of fairness in high stakes allocation decisions. *The Economic Journal* 114, 121-149.
- LOEWENSTEIN, GEORGE, (2000). Emotions in economic theory and economic behavior. *The American Economic Review* 90, 426-432.
- MADDALA, GANGADHARRAO S., (1988). Introduction to econometrics, New York: Wiley.
- MALLE, BERTRAM F. AND KNOBE, JOSHUA, (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology* 33, 101-121.
- MARQUARDT, DONALD W., (1970). Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation. *Technometrics* 12, 591-612.
- MARWELL, GERALD, AND AMES, RUTH E., (1981). Economists free-ride, does anyone else? *Journal of Public Economics* 15, 295-310.
- MCCABE, KEVIN, (1989). Fiat money as a store of value in an experimental market. *Journal of Economic Behavior and Organization* 12, 215-231.

- MEIER, STEPHAN, AND FREY, BRUNO S., (2004). Do business students make good citizens? *International Journal of Economics of Business* 11, 141-163.
- MIKHAIL, JOHN M. , (2000). Rawls linguistic analogy: A study of the generative grammar model of moral theory described by John Rawls in a theory of justice. *Unpublished PhD*, Cornell University, Ithaca, New York.
- NADELHOFFER, THOMAS, (2004). The butler problem revisited. *Analysis* 64, 277-284.
- OFFERMAN, THEO, (2002). Hurting hurts more than helping helps. *European Economic Review* 46, 1423-1437.
- PARDUCCI, ALLEN, AND WEDELL, DOUGLAS H., (1986). The category effect with rating scales: Number of categories, number of stimuli and method of presentation. *Journal of Experimental Psychology: Human Perception and Performance* 12, 496-516.
- RABIN, MATTHEW, (1993). Incorporating fairness into game theory. *The American Economic Review* 83, 1281-1302.
- RATCLIFFE J. F., (1968). The effect on the t distribution of non-normality in the sampled population. *Applied Statistics*, 17 42-48.
- REUBEN, ERNESTO, AND VAN WINDEN, FRANS, (2004). Reciprocity and emotions - when reciprocators know each other. Tinbergen Discussion Paper.
- ROETHLISBERGER, FRITZ J., AND DICKSON, WILLIAM J., (1939). Management and the worker. New York: Wiley.
- ROTH, ALVIN E., PRASNIKAR, VESNA, OKUNO-FUJIWARA, MASAHIRO, AND ZAMIR, SHMUEL, (1991). Bargaining and market behavior in Jerusalem, Ljuljana, Pittsburgh, and Tokyo: An experimental study. *The American Economic Review* 81, 1068-1095.
- RUFFLE, BRADLEY J., (1998). More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavior* 23, 247-265.
- RUTTE, CHRISTEL G., WILKE, HENK A.M., AND MESSICK, DAVID M., (1987). Scarcity or abundance caused by people or the environment as determinants of behavior in the resource dilemma. *Journal of Experimental Social Psychology* 23, 208-216.

- SADRIEH, ABDOLKARIM, (2003). Equity versus warm glow in the inter-generational giving. *CentER Discussion Paper 2003-35*, Tilburg University.
- SANFEY, ALAN G., LOEWENSTEIN, GEORGE, MCCLURE, SAMUEL M., AND COHEN, JONATHAN D., (2006). Neuroeconomics: Cross-currents in research on decision-making. *TRENDS in Cognitive Science* 10, 108-116.
- SANFEY, ALAN G., RILLING JAMES K., ARONSON, JESSICA A., NYSTROM, LEIGH E., AND COHEN, JONATHAN D., (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755-1758.
- SEFTON, MARTIN, (1992). Incentives in simple bargaining games. *Journal of Economic Psychology* 13, 263-276.
- SINGER, TANIA, AND FEHR, ERNST, (2005). The neuroeconomics of mind reading and empathy. *The American Economic Review* 95, 340-345.
- STÅHL, INGOLF, (1972). Bargaining theory, Stockholm. Economic research institute.
- STANLEY, TOM D., AND TRAN, UME, (1998). Economics students need not be greedy: Fairness and the ultimatum game. *Journal of Socio-Economics* 27, 657-663.
- WALLACE, BJÖRN, CESARINI, DAVID, LICHTENSTEIN, PAUL, AND JOHANNESSON, MAGNUS, (2007). Heritability of ultimatum game responder behavior. *Proceedings of the National Academy of Sciences* 104, 15631-15634.
- WORLD COMMISSION ON ENVIRONMENT AND DEVELOPMENT, (1987). Our common future, report of the world commission on environment and development. *Annex to general assembly document A/42/427, development and international co-operation: Environment*.
- YEZER, ANTHONY M., GOLDFARB, ROBERT S., AND POPPEN, PAUL J., (1996). Does studying economics discourage cooperation? Watch what we do, not what we say or how we play. *Journal of Economic Perspectives* 10, 177-186.

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich meine Doktorarbeit ‘Sustainability - is it contagious?’ selbständig und ohne fremde Hilfe angefertigt habe und dass ich alle von anderen Autoren wörtlich übernommenen Stellen, wie auch die sich an die Gedanken anderer Autoren eng anlehenden Ausführungen meiner Arbeit, besonders gekennzeichnet und die Quellen nach den mir angegebenen Richtlinien zitiert habe.

Kiel, den 21.9.2009

Curriculum Vitae

Personal information

Name Gunter Bahr
Address Schauenburgerstraße 69
24118 Kiel
Telephone 0170/7507822
E-mail gunterbahr@freenet.de
Date of birth: 01.12.1977 in Kiel



Work experience

Since 5/2008

Junior Product Developer in the Energy business unit of the HSH Nordbank AG

- Management of the design, implementation and realization of a standardized research process, especially conducting market analysis
- Member of a team to set up a fund in the renewable energy sector. Analysis of cash-flows and the legal structure

10/2003 – 9/2007

Research assistant at the chair of Innovation, Competition Policy and New Institutional Economics at Kiel University

- Tutor for 'Introductory Economics' and 'Introductory Microeconomics'
- Supervision of students in seminars on Environmental Economics, Innovation, Competition Policy and New Institutional Economics

Study

Since 10/2003

Postgraduate study 'Quantitative economics' at Kiel University

- Presumably 9/2009: Degree: Dr. sc. pol.
- Dissertation: 'Fairness in an intergenerational three-person dictator game with social interaction', an experimental study on the preference structure of economics subjects
- 7/2007: Presentation on the 'Economic Science Association World congress' in Rome
- 2/2007: Presentation on the workshop 'Umwelt- und Ressourcenökonomik für Nachwuchswissenschaftler' at ETH Zürich

4/1999 - 5/2003

Diploma study of economics at Kiel University

- 5/2003: Degree: Diplom-Sozialökonom (1.8)
- 8/2002 - 5/2003: Erasmus-exchange Stockholm University

10/1998 - 3/1999

Magister study (Political science, economics and public law) at Freiburg University

Work experience during study

10/2001 – 5/2002

Student assistant at the chair of Innovation, Competition Policy und New Institutional Economics at Kiel University

- Literature research
- Assistance in economic experiments
- Statistical analysis of data
- Preparation of figures and presentations

9/2000 – 10/2000

Internship at the faction of the Landtag Bündnis 90/Die Grünen

- Conceptual design of a hearing with regard to ‘Evaluation at universities’
- Formulation of a proposition for the introduction of foreign languages at elementary school

Extracurricular activity

1/2000 - 7/2002

Work in the student body of Sociology and Political Science at Kiel University

- from 10/2001: Leading position
- Accounting of the budget
- Organization and realization: Information events for freshman students, excursions for freshman students, student body parties

Civilian service

7/1997 - 8/1998

Civilian service at Academy for Nature and Environment, Neumünster

- Maintenance of a trail on nature-orientated garden design
- Guiding pupils on the nature trail
- Assistance at the presentation of the academy at fairs
- Organizational tasks

School

6/1997

Abitur (1.8)

8/1988 - 6/1997

High school, Dahlmannschule in Bad Segeberg

8/1984 - 7/1988

Elementary school in Rickling

Language skills

English (fluent), Swedish (very good), French (good), Italian (satisfactory), Spanish (basic knowledge)

Technical skills

MS-Office, CorelDraw, LaTeX, EViews, STATA, Matlab

Hobbies

Windsurfing, football, piano, guitar