



UvA-DARE (Digital Academic Repository)

Social preferences and emotions in repeated interactions

Hoyer, M.O.

Publication date

2018

Document Version

Final published version

License

Other

[Link to publication](#)

Citation for published version (APA):

Hoyer, M. O. (2018). *Social preferences and emotions in repeated interactions*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Maximilian Hoyer

Social Preferences and Emotions in Repeated Interactions

The way people make decisions in social situations is in many ways a product of how they feel about those affected by their decisions. Both positive and negative emotions are important factors in decision-making. This thesis presents results from a series of laboratory experiments designed to contribute to our understanding of emotions in social economic games, using both purely behavioral experiments and one fMRI experiment. A special focus lies on different aspects of the interaction between two different people and the difference between positive and negative ties that result from cooperative or destructive behavior by others. We do not find that destructive behavior causes stronger reactions than cooperative behavior. The final chapter, which focuses on the distinction between different kinds of positive and negative experiences instead of merely direction, finds clear differences between intentional and circumstantial shared experiences.

Maximilian Hoyer holds a BSc degree from Maastricht University and a MPhil/MSc degree from the University of Amsterdam and the Tinbergen Institute. He joined the Center for Research in Economics and Political Decision Making (CREED) as a PHD student in 2012. In 2016, he was part of the founding team of candidate select GmbH, which provides empirical analysis to support recruitment decisions.

Social Preferences and Emotions in Repeated Interactions
Maximilian Hoyer



Social Preferences and Emotions in Repeated Interactions

ISBN 978 90 3610 490 6

This book is no. **709** of the Tinbergen Institute Research Series, established through cooperation between Rozenberg Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Social Preferences and Emotions in Repeated Interactions

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
Prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op donderdag 29 maart 2018, te 12:00 uur

door

Maximilian Olaf Hoyer

geboren te Wuppertal, Duitsland

Promotiecommissie:

| | | |
|-----------------------|-------------------------------|----------------------------------|
| Promotor: | Prof. dr. F.A.A.M. van Winden | Universiteit van Amsterdam |
| Copromoter: | Dr. N.G.J. Bault | Università degli Studi di Trento |
| Overige leden: | Prof. dr. J.H. Sonnemans | Universiteit van Amsterdam |
| | Prof. dr. T.J.S. Offerman | Universiteit van Amsterdam |
| | Prof. dr. K.R. Ridderinkhof | Universiteit van Amsterdam |
| | Prof. dr. C.K.W. de Dreu | Universiteit Leiden |
| | Prof. dr. U. Weitzel | Universiteit Utrecht |

Acknowledgements

This thesis is the result of not only my own efforts, but a big group of people, to whom I am indebted in many different ways.

First and foremost, I would like to thank my supervisors, Frans van Winden and Nadge Bault. Over the years we worked together in a very closely on a number of projects. They were always willing to invest a great amount of time and dedication, listening to every question I might have had along the way. At the same time, they never treated me as anything less than their peer, always accepting my ideas and contributions in common projects as equally valuable and important as theirs.

Secondly, I would like to mention my friend and coauthor Ben, who was also involved in the series of projects that forms the basis of chapters two through four of this thesis. As such, we were close colleagues and friends for the better part of four years, an experience that was both enjoyable and interesting.

I would also like to thank the members of my defense committee, Carsten de Dreu, Theo Offerman, Richard Ridderinkhof, Joep Sonnemans, and Urs Weitzel for participating in my defense and generously investing their time and effort reading and evaluating my thesis.

During my time at CREED I got to work with a number of fantastic colleagues. Thanks to a number of building changes, I had the fortune to share an office with a great many PhD students. Starting in a room with Aaron and Ben, sports statistics were probably the most important office topic. Later, in a larger office together with Matze and Boris, we shifted more towards politics, as economists are wont to do at times. Finally, once I moved on to sharing offices with Simin and Margarita, the topics finally became more

diverse and a little bit less stereotypically male. All of them have been great friends inside and outside the office for many years.

That is also true of many of the other inspiring and gifted PhD students that I got to share a research environment and many non-academic interests and experiences with in Amsterdam. At CREED there were Andro, Anita, David, Francisco, Jindi, Junze, Lenny, Pedro, Stephan, and Yang. From the greater community of Tinbergen Institute students I would like to give special mention to Anghel, Eszter, Heiner, Gosia, Karo, Lucy, Oli, Rei, Sait and Violeta, who all became great friends over the years. Particular consideration goes to my flatmate of more than four years, Lukas, whose positive attitude towards life and sheer determination were a constant example to me. I would also like to thank all the gracious people who hosted me on my regular visits to Amsterdam since I officially moved away two years ago, Camille & Mark, Charlotte, Delphine, Gosia, Gunes, Heiner, Kath, Oli, Rei, Vince, and Violeta. I would also like to thank my old friends and new colleagues in Cologne, Jan and Philipp, with whom I was able to make great new experiences outside of academia during that time.

Finally, I would like to thank my family for supporting me over the years and my friends from outside of the academic world for never – visibly – tiring of hearing me talk about the weird little universe that is academic research.

Maximilian Hoyer

Cologne, 2018

Contents

| | |
|---|-----------|
| Acknowledgments | ii |
| 1 Introduction | 1 |
| 2 Destructive Behavior in a Fragile Public Good Game | 7 |
| 2.1 Introduction | 8 |
| 2.2 Experiment | 9 |
| 2.2.1 Symmetric Treatment (SYM) | 9 |
| 2.2.2 Framing Treatment (FRAME) | 10 |
| 2.2.3 Asymmetric Treatment (ASYM) | 10 |
| 2.3 Results | 11 |
| 2.4 Conclusion | 16 |
| 2.5 Appendix: Instructions and Test Questions | 17 |
| 2.5.1 Instructions Ring Test | 17 |
| 2.5.2 Instructions and Test Questions Main Part, SYM | 18 |
| 2.5.3 Instructions and Test Questions Main Part, FRAME | 21 |
| 2.5.4 Instructions and Test Questions Main Experiment, ASYM | 24 |
| 2.5.5 Interstitial Instruction Screens | 27 |
| 3 Asymmetry in the Development of Cooperative and Antagonistic Relationships. A Model-Based Analysis of a Fragile Public Good Game | |
| Experiment | 29 |
| 3.1 Introduction | 30 |
| 3.2 Theory | 32 |
| 3.2.1 Fragile Public Good Game | 32 |
| 3.2.2 Model | 33 |
| 3.2.3 Model analysis and hypotheses | 35 |
| 3.3 Methods | 38 |
| 3.3.1 Experiment | 38 |
| 3.3.2 Estimation | 40 |
| 3.4 Results | 41 |
| 3.4.1 Descriptive statistics | 41 |
| 3.4.2 Group level results | 43 |
| 3.4.3 Individual level results | 46 |
| 3.4.4 Predictive performance out-of-sample and model comparison | 48 |
| 3.5 Applying the Ties Model to the Repeated Prisoner’s Dilemma | 52 |
| 3.6 Conclusion | 60 |
| 3.7 Appendix A: Proof of Propositions | 61 |

| | | |
|----------|---|------------|
| 3.7.1 | Proposition 1: Contributions outside of $-5 \leq C_{it} \leq 5$ can never be part of any equilibrium if $-1 \leq \alpha \leq 1$. | 61 |
| 3.7.2 | Proposition 2.2: All symmetric equilibria with $-5 \leq C_{it} \leq 5$ are possible. | 62 |
| 3.7.3 | Proposition 3: Asymmetric equilibria exist if $C_i C_j > 0$ and $ C_i \leq 5$ and $ C_j \leq 5$ | 65 |
| 3.7.4 | Proposition 4: For the socially optimal choices to be a stable equilibrium under the myopic model, both players satisfying $0.2 < \frac{\delta_{2i}}{1-\delta_{1i}} < 0.25$ is a necessary, but not a sufficient condition. | 66 |
| 3.8 | Appendix B: Reference Point | 66 |
| 3.9 | Appendix C: Tie Model Parameters for different Repeated PD Strategies | 69 |
| 3.9.1 | Introduction | 69 |
| 3.9.2 | Tit-for-tat | 70 |
| 3.9.3 | Tit-for-2-tats | 71 |
| 3.10 | Appendix D: Payoff Matrix and SVO Example | 73 |
| 4 | The Processing of Positive and Negative Social Impulses | 75 |
| 4.1 | Introduction | 76 |
| 4.2 | Materials and Methods | 79 |
| 4.2.1 | Social Value Orientation | 80 |
| 4.2.2 | Fragile Public Good Game | 80 |
| 4.2.3 | Experimental Procedure | 81 |
| 4.3 | Results | 87 |
| 4.3.1 | Behavior | 87 |
| 4.3.2 | Model estimates | 90 |
| 4.3.3 | fMRI results | 91 |
| 4.4 | Discussion | 94 |
| 4.5 | Conclusion | 100 |
| 4.6 | Supplementary Material | 102 |
| 4.6.1 | Instructions | 102 |
| 4.6.2 | Payoff Matrix | 103 |
| 4.6.3 | Contributions in individual Dyads | 104 |
| 4.6.4 | SVO and Contribution Correlations | 110 |
| 4.6.5 | Additional fMRI data | 111 |
| 4.6.6 | Questionnaires | 111 |
| 4.6.7 | Subsample of subjects with high θ value | 114 |
| 5 | Investors have feelings too | 117 |
| 5.1 | Introduction | 118 |
| 5.2 | Literature | 122 |
| 5.3 | Design and Hypotheses | 127 |
| 5.3.1 | Treatments | 127 |
| 5.3.2 | Projects | 131 |
| 5.3.3 | Presentation and Organization | 134 |
| 5.3.4 | Hypotheses | 135 |
| 5.4 | Results | 138 |
| 5.5 | Discussion | 148 |
| 5.6 | Conclusion | 151 |

| | | |
|-------|--|------------|
| 5.7 | Appendix A | 152 |
| 5.7.1 | Additional Figures and Tables | 152 |
| 5.8 | Appendix B: Supplementary Material | 159 |
| 5.8.1 | Instructions | 159 |
| 5.8.2 | Questionnaires | 159 |
| 5.9 | Appendix C: Alternative Theories | 163 |
| 5.9.1 | The Role of Other Social Preference Models in our Design | 163 |
| 5.9.2 | The Case of Flexible Project Manager Payment | 164 |
| | Summary | 171 |
| | Samenvatting | 174 |
| | Bibliography | 178 |

Chapter 1

Introduction

Human history is, more than anything else, a story of relationships. Our superior ability to communicate and to form bonds was central to establishing a lasting period of human dominance over other animals on the planet. At the same time, our tendency to hold grudges or, even worse, cultivate hatred towards other humans, can prevent cooperation, cause violence, and may well ultimately be the reason that this period of human dominance ends. Shifting our gaze to interactions of less grandiose a scale, the relevance of attachments and social ties is just as relevant. This is almost tautologically true on the interpersonal level, but also when it comes to one of the central topics that economics as a science is interested in: the ability of individuals or groups to cooperate with each other for the common good.

This dissertation is wholly focused on the question of how such ties develop. In particular, the type of relationship we are interested in is relatively short term and characterized by the possibility to behave in a cooperative manner, or to be aggressive towards the other person. In particular, we are striving to compare these two options on an equal footing and to gain an understanding of how, if at all, positive and negative relationships differ systematically. To this end, all the experiments presented here are designed with the idea in mind to allow for the comparison of positive and negative relationships. In many cases the theoretical difference between a positive and a negative relationship will only be a positive or negative sign before a variable. This gives us the ability to analyze such relationships comparatively and focus on the way in which they differ. Specifically, our approach is intended to specify the difference in which emotional reaction to interaction

with another individual differs for positive and negative experiences.

In this dissertation, the term relationship will typically be meant to represent a repeated interaction between two individuals in a laboratory setting, which we use to analyze and test our models. As such we are talking of relationships in the most abstract sense. All four substantial chapters are based on experiments that pit participants against each other in a perfectly anonymous fashion, facilitated through a series of computerized experiments. While this is in line with typical practice in the research field of experimental economics, it may be surprising to the casual reader to see us use terms such as bond, tie, or relationship to describe such rather cold and short interactions. One might even wonder what such a stylized type of experiment could possibly tell us about the interaction of people in the "real world". In fact, there are a number of studies that analyze the type of social interaction we are interested in using field experiments, which are studying naturally developed social ties, or use naturally occurring ties in a lab setting (Goette et al., 2012; Leider et al., 2009; Reuben and Van Winden, 2008). However, while these approaches have advantages in terms of external validity, they bring with them certain limitations that would make it difficult to achieve the degree of control over events that form these relationships that is necessary for the type of rigorous analysis that we are striving for in the research presented here.

On a theoretical level the experiments presented here are part of a series of projects built around the affective social tie model of van Dijk and van Winden (1997). This model formalizes the idea that the emotional disposition towards a particular other can be modeled using a parameter that tracks the degree of attachment towards the other. This parameter, the Social Tie parameter, is affected by the actions of the other and continuously updated. The individual's future actions are then affected by the parameter insofar as a more positive parameter leads to more positive actions towards the other and vice versa. We are using the theory, which has been already been applied in a series of earlier experiments (van Dijk et al., 2002; Sonnemans et al., 2006; Bault et al., 2017), and extend its empirical application primarily by veering further into the domain of destructive relationships. One of the leading themes of this dissertation is therefore whether the way

in which an individual's Social Tie parameter is updated differs for positive and negative actions of the other or not.

In particular, we were inspired by the observation that, if given the opportunity, individuals can be quite aggressive in giving up own income in order to hurt another individual that they interacted with previously. Specifically, earlier experiments have demonstrated that subjects react stronger to hurtful than to helpful behavior (Keysar et al., 2008; Offerman, 2002). This observation has multiple implications. One aspect is that, in repeated interactions, the dynamics of destructive relationships are different from the dynamics of constructive relationships. A much more micro-level aspect is that the way in which humans process destructive and constructive behaviors by others are distinct. Therefore, we designed a series of experiments that allow us to a) study the development of positive and negative ties, as opposed to purely positive ties, and b) study situations in which individuals have the possibility to either cooperate or destruct with similar costs as one single decision, as opposed to studying cooperation and destruction in isolation or artificially separating the two types of action.

The different chapters of this dissertation present separate essays, but are connected insofar as they are all based on experiments that make participants interact with each other repeatedly, allowing for the development of positive or negative short-term relationships in the sense of the Social Ties model. This is a feature that this document shares with the author's MPhil thesis (Hoyer, 2012), which focused on the effect that positive or negative experiences in one interaction have on interactions with other individuals. Chapters 2 to 4 are related even more closely: Chapters 2 and 3 share the same experiment as their source of data, but focus on different analytical aspects of that project. Chapter 4 uses a very similar experimental design, but this time the experiment is performed in a functional magnetic resonance imaging (fMRI) scanner. Hence, the analysis is enriched by an additional dimension of data and we are not only able to analyze the participants' behavior, but also their cognitive processing of the situations they encounter. Chapter 5 moves back to a purely behavioral experiment, but looks at a more specific situation by adding randomized results in an investor/manager relationship to the dynamic rela-

tionship between two participants. All chapters were co-authored with Frans van Winden and chapters 2 to 4 were also co-authored with Ben Loerakker and Nadège Bault.

Chapter 2 focuses on the development of negative, or destructive, relationships in an experiment where participants interact with the same one other participant for 35 rounds. They have the option to either contribute to a common, public good like project, or to costly take from the public good, hurting the other participant. This choice is not binary, but spread out over a range of 15 different potential choices. We analyze three treatment conditions, which differ in the design and framing of the payoff matrix that translates the two participants' decisions into outcomes in the form of payoffs. In this chapter, we particularly look at behavior in the context of a literature regarding the development of conflict in repeated games (Abbink and Herrmann, 2009; Fehl et al., 2012; Nikiforakis and Engelmann, 2011). The main results are the following. First, participants did engage in destructive behavior, despite it not being rational from a simple profit-maximizing perspective to do so within each individual round. Second, feuds, repeated periods of mutual destructive behavior, are most common in our so called asymmetric treatment, which was designed to make them particularly likely. While some of them stretch on for several rounds, very long feuds are rarely observed.

In chapter 3, the same data as in the previous chapter is used, but this time with a focus on applying and improving the estimation of the underlying Social Ties model. Building on techniques used in Bault et al. (2017), different estimation methods are developed and implemented. Out of sample prediction is used to improve on the estimation quality and comparisons to other relevant models are made. Finally, the estimated model parameters are used to define different automatic rules, so-called automata, for behavior in repeated prisoner's dilemma games (Dal Bó and Fréchette, 2011; Fudenberg et al., 2012). Using the model-based estimations from the Fragile Public Good game experimental data, it is possible to generate a distribution of different automata that explain the prisoner's dilemma game experimental data, as opposed merely being observed.

Chapter 4 takes the design from one of the treatments in chapters 2 and 3 and adapts it for use in a fMRI study, combining aspects of the experimental design found in the previous

chapters and in Bault et al. (2014). Again, we are interested in the difference between positive and negative relationships, but this time the focus lies on the neural processing of such experiences and decisions. We ran the experiment in a setup wherein one of the participants in each interaction is placed in a fMRI scanner. In difference to many experiments in the field of cognitive science, our participants are actually participating in a live interaction with other participants, who are not placed in a scanner, but interact with the scanned participants via a computer network, eliminating the need for any type of deception. While the added complexity of the more intricate technical and organizational setup presented some serious challenges, we do find several noteworthy results. Most importantly, our results are consistent with the theory that the cognitive processing of positive and negative social ties is linearly, as opposed to a quadratically, encoded in a brain area found previously to be implicated in social interaction.

Finally, in chapter 5 we look at another configuration of a dyadic interaction, namely one in which one of the two actors has the option to unilaterally end the interaction and choose another partner to interact with. In this chapter, entitled "Investors have feelings too", an investor chooses a manager, who in turn chooses a project to implement in the name of the investor. Other than in more traditional principal-agent type situations the manager's choice, while meaningful for the result, is free from any type conflict of interest. Neither are there different types of effort that the manager could exert. As a result, the outcome of the project is not objectively tied to the manager at all, other than through the blind choice she made early on. The goal of the experiment is to show whether the success or failure of the project is having an effect on the feeling of attachment that the investors may have towards the manager, as expressed through the investor's choice to either keep the manager to implement a new project or to "fire" and replace her with another manager. With this design it attempts to bring together the literature on "relationship banking" (Boot, 2000) and experiments on effects such as unfounded blame (Gurdal et al., 2013) and gift giving as a form of engendering favoritism (Malmendier and Schmidt, 2012). Between the last two concepts we find much stronger effects for the latter than the former. The effect of the sheer association of a manager with a success

or failure fails to be substantially more pronounced than a non-social control treatment, even though secondary measures such as decision time show that participants appear to consider the effects that their decisions have on other participants in the experiment. Gift giving, however, has a significant effect.

Chapter 2

Destructive Behavior in a Fragile Public Good Game¹

¹This chapter is based on work with Nadège Bault, Ben Loerakker and Frans van Winden, published in volume 123 of *Economics Letters*, 2014. Financial support by the Research Priority Area Behavioral Economics of the University of Amsterdam, the French National Research Agency (ANR-11-EMCO-01101) and the LABEX CORTEX (ANR-11- LABX-0042) is gratefully acknowledged.

2.1 Introduction

Many experimental studies have investigated the development of cooperation in a social dilemma or public good environment, and the effect of punishment mechanisms in this context (for a recent survey, see Chaudhuri, 2011). In the real world, however, people can often cooperate with or hurt one another. Relationships may even turn sour and induce persistent destructive behavior. Repeated and severe conflict is a real part of human interaction. Examples are neighborhood conflicts, family feuds, or the destruction of public property during riots. In some studies, a substantial proportion of individuals engaged in the destruction of others' earnings, even when rank egalitarianism and reciprocity motives were not present and when the destruction was costly (Zizzo, 2003; Abbink and Sadrieh, 2009). To study whether destructive behavior can be observed and modulated in a public good environment we designed a 'fragile public good' (FPG) game. A key feature of the FPG game is that it gives as much room for destructive behavior (taking) as for constructive behavior (contributing). More formally, it does so by shifting both the (standard) Nash equilibrium and the status quo – i.e., the initial allocation of tokens to the common account – to the middle of the action space, with perfect symmetry in the marginal cost of taking and contributing. Contrary to the relatively few public good experiments that allow for an interior Nash equilibrium (see surveys by Laury and Holt, 2008, and Saijo, 2008), we focus on destructive actions in a repeated context where subjects can identify the individual decisions of others.

This chapter is related to a developing stream of literature on 'feuds' (Nikiforakis and Engelmann, 2011) and Nikiforakis et al. (2012) and 'vendettas' (Fehl et al., 2012; Abbink and Herrmann, 2009; Bolle et al., 2013). These experiments typically focus strongly on the punishment by explicitly separating contribution and punishment stages. For example, Nikiforakis and Engelmann (2011) use separate punishment rounds after a 4-player public good game, whereas Bolle et al. (2013) let subjects decrease others' probabilities of winning a prize.

Because framing can influence behavior in public good games (Brewer and Kramer, 1986;

Sonnemans et al., 1998; Willinger and Ziegelmeyer, 1999; for a survey, see Cookson, 2000), we study the sensitivity of our findings in two additional treatments, where we separate the status quo from the Nash outcome. In one case, we move the status quo to a corner so that subjects can only contribute, keeping everything else the same; this is a case of positive framing which may induce subjects to contribute more. In the other case, we move the Nash outcome away from the status quo towards taking by introducing a slight payoff asymmetry. Here the Nash choice may be read as aggression by subjects using the status quo as a reference point and induce destructive behavior.

Our main questions are: (1) does the FPG game generate destructive behavior and even cases where behavior equilibrates towards sour relationships?; (2) how does separating the Nash outcome from the status quo through framing or some minimal payoff asymmetry modulate taking and contributing? After the design we present our results, followed by a summary of our findings.

2.2 Experiment

Subjects played the FPG game in fixed dyads over 35 rounds in all three treatments.

2.2.1 Symmetric Treatment (SYM)

In each round both subjects of a dyad are endowed with a private account holding 7 tokens, earning 10 units each, and a common account holding 14 tokens, earning 10 each for *both* subjects. Subjects can contribute to or take up to 7 tokens from the common account, at increasing marginal costs: moving one token costs 2 units, while the marginal transfer cost of each additional token increases by 2^2 . Earnings are symmetric around the status quo which coincides with the selfish Nash outcome, while any combination of contributions of 4 or 5 is socially optimal.

²Formally, we use the following payoff function, where c_i can be positive or negative: $V_A(c_A, c_B) = 10(14 + c_A + c_B) + 10(7 - c_A) - (|c_A| + c_A^2)$. See figure 5 in appendix.

2.2.2 Framing Treatment (FRAME)

In FRAME subjects have exactly the same strategy space and equivalent earnings, but now they start each round with 14 tokens in their private accounts and the common account is empty. Thus, to reach an outcome equivalent to an outcome in SYM, subjects would have to contribute 7 more tokens than before³. Because only contributions can be made, this is a case of positive framing.

2.2.3 Asymmetric Treatment (ASYM)

ASYM differs from SYM in only two respects: tokens in the private account earn subjects 11 units instead of 10, and the first token transferred in either direction has zero costs. As in FRAME, the Nash equilibrium does not coincide with the status quo, but now it is the former that moves by prescribing to take one token out of the common account, while both subjects contributing 5 tokens is the social optimum⁴.

Subjects did not see the underlying formulas, but were supplied with graphs illustrating the marginal effects of every decision for themselves and the other, alongside with payoff tables⁵.

The public good game was preceded by a test of social value orientation (SVO; see Liebrand and McClintock, 1988a, taken from van Dijk et al., 2002). This test measures the preferences of subjects for distribution outcomes for themselves and a (generalized) other. Sessions were run in November and December 2012 and April 2013 at the CREED-lab in Amsterdam. SYM had 130 participants (50% female, 2% unreported gender, average age 22.2), FRAME 54 (41% female, average age 21.5), and ASYM 80 (43% female, average age 21.5). The experiment had an additional second part, which we do not cover in this chapter. The exchange rate of units into euros was 700 to one. Subjects earned on average 1.45 euro in the SVO-test and 10.82 euro in the public good game.

³Payoff function: $V_A(c_A, c_B) = 10(c_A + c_B) + 10(14 - c_A) - (|c_A - 7| + (c_A - 7)^2)$

⁴Payoff function: $V_A(c_A, c_B) = 10(14 + c_A + c_B) + 11(7 - c_A) + |c_A| - c_A^2$

⁵Instructions are available upon request.

2.3 Results

Table 2.1 gives an overview of average contributions, where we adjust for the Nash equilibrium (NE) in each game by subtracting 7 tokens from results in FRAME and adding 1 to results in ASYM.

Table 2.1: Average contributions

| Average contribution | SYM | FRAME | ASYM |
|----------------------|-------------|--------------|-------------|
| Overall | 2.28 (2.01) | 2.07 (1.69) | 1.83 (3.3) |
| First round | 1.26 (2.55) | -0.02 (2.55) | 0.92 (3.06) |
| Rounds 26-34 | 2.44 (2.35) | 2.49 (3.85) | 2.13 (3.85) |
| Last round (35) | 0.68 (2.43) | 0.98 (1.81) | 0.31 (3.81) |

Note: adjusted for the (standard) Nash equilibrium; standard deviation in parentheses, with dyad averages as separate observations.

Average contributions are approximately 2 tokens above the Nash-prediction in all treatments. The first round, however, reveals a different pattern as the average contribution in FRAME is significantly lower than in SYM ($p = 0.001$)⁶. Because SYM and ASYM are more similar to a taking game than FRAME (where only contributions are possible), this result contrasts with the general finding that there are typically lower contributions in taking framings, if there is any difference (Andreoni, 1995; Sonnemans et al., 1998; Goerg and Walkowitz, 2010). Khadjavi and Lange (2015) have a treatment with intermediate endowments similar to our SYM and ASYM treatments and find no differences between a contributing frame and this alternative.

Subjects appear to be reluctant to contribute early on in FRAME, but are able to compensate for this throughout the game, as the difference stays significant at 1% up until the fifth round of the game.

⁶We use the Mann-Whitney U-test with dyad averages as observations unless otherwise mentioned.

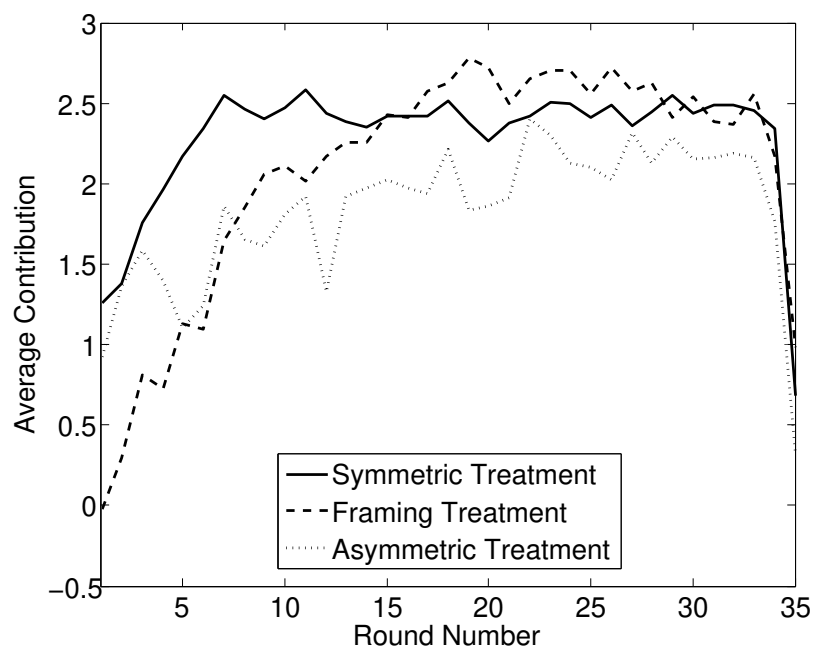


Figure 2.1: Average contributions, relative to Nash equilibrium

All treatments show an increase in contributions over time until the (usually observed) sharp decline at the end; see figure 2.1⁷. A simple regression shows significant positive time trends. Although the increase is at odds with the general observation of decreasing cooperation in public good experiments (Ledyard, 1995), it has been observed before in repeated two-player games using a comparable mechanism (van Dijk et al., 2002). Comparing SYM and ASYM, the hypothesis of equal contributions is rejected if they are calculated relative to the status quo ($p = 0.035$), but not relative to the Nash equilibrium outcome, which may suggest that the latter is a more important reference point.

Table 2.2: Percentage of destructive decisions

| Percentage of destructive decisions | SYM | FRAME | ASYM |
|-------------------------------------|--------|--------|--------|
| Overall | 11.25% | 6.93 % | 21.14% |
| Last round | 13.9% | 5.36% | 25% |

Most relevant to this chapter is the occurrence and development of destructive behavior hurting both partners. Relative to all decisions, destructive decisions count 11% in SYM,

⁷Contributions in the last round are lower than in the ten rounds before ($p < 0.01$ in all treatments.)

7% in FRAME, and 21% in ASYM; see table 2.2. The percentages of subjects choosing below the Nash at least once are, respectively, 42%, 46%, and 56%. The higher number of destructive decisions in ASYM, despite similar average contribution levels relative to the Nash outcome (see table 2.1), suggests distributional differences. Indeed, the variance of subjects' decisions is larger in ASYM than in SYM and FRAME in 31 of the 35 rounds (Levene's test, $p < 0.01$; see also figure 2.2). Interestingly, this difference only becomes significant from the third round onwards, which indicates that it is partly driven by the dynamics in the game. Not only the variance across subjects, but also the variance within each subject's set of 35 decisions is greater in ASYM⁸. Summing the destructive decisions of each dyad we find a difference only between SYM and ASYM ($p = 0.094$)⁹. The higher level of conflict observed in ASYM is confirmed by the observation that the percentage of destructive decisions in the last round (when there are no strategic considerations present) is higher in ASYM than in the other treatments.

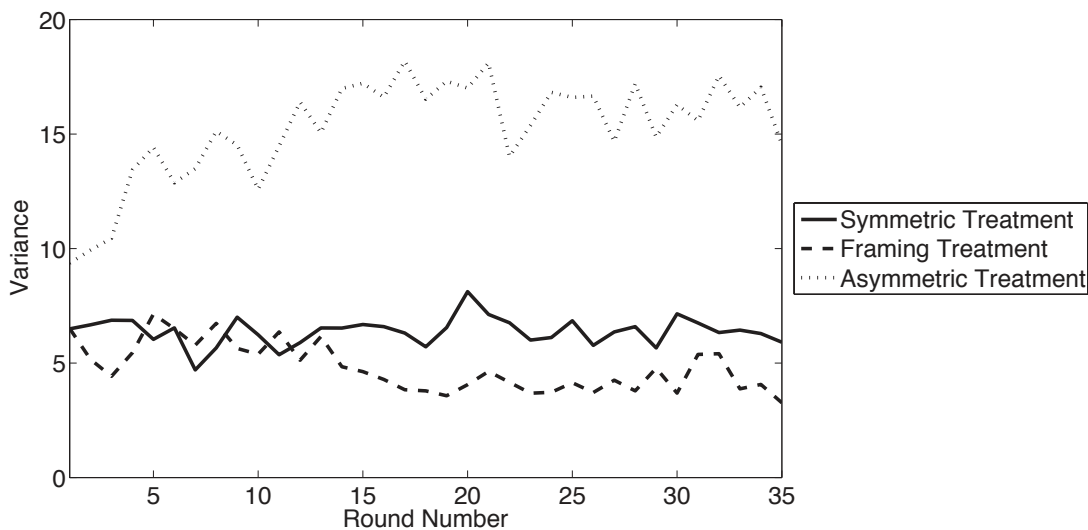


Figure 2.2: Between-subject variance across treatments/rounds

Background measures seem to play a role in explaining destructive behavior: In ASYM and SYM contributions are lower and destruction rates higher in female dyads than in

⁸Means of within-subject variances: 2.57 in SYM, 2.54 in FRAME, and 4.59 in ASYM. The differences between SYM and ASYM ($p < 0.001$) and FRAME vs ASYM ($p = 0.038$) are significant.

⁹In line with a result of Nikiforakis et al. (2012), who find higher rates of counter-punishment in a treatment with increased normative conflict.

male dyads at 5% significance levels and SVO correlates with individual decisions¹⁰.

It appears that sour relationships do indeed develop in the FPG game, as illustrated in two examples in figure 2.3.

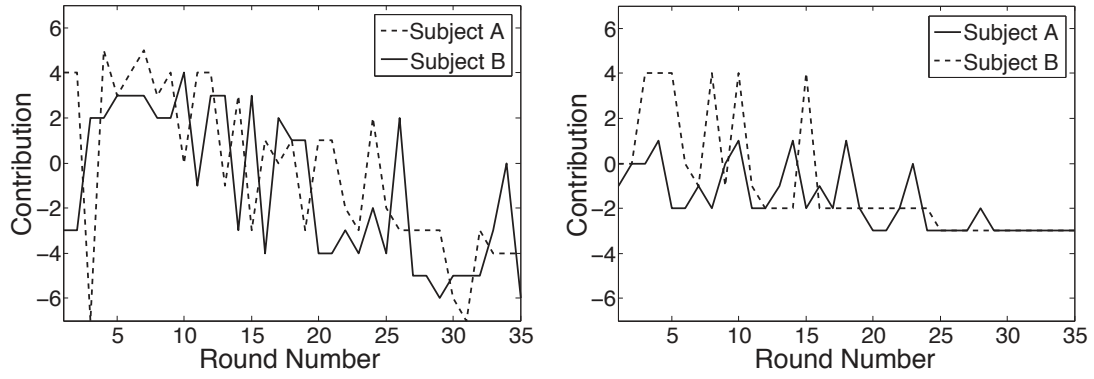


Figure 2.3: Examples of sour relationships in SYM

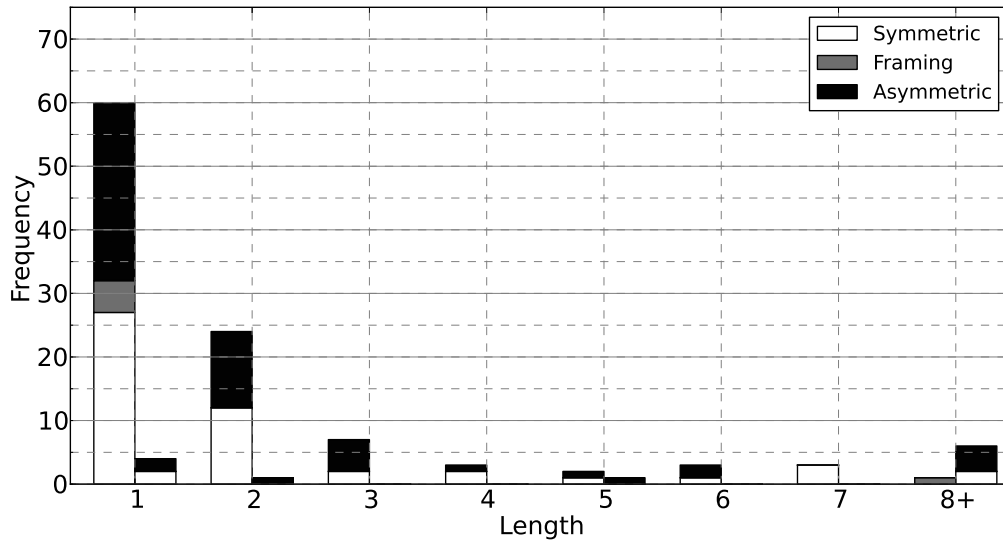


Figure 2.4: Distribution of feuds that end before (left bar) and at (right bar) the final round

¹⁰In ASYM with $p \leq 1\%$ for contributions in the first round, first 5 rounds, and the whole game, and (negatively) with the number of destructive decisions; $p \leq 5\%$ for the first round in SYM; $p \leq 5\%$ for the first 5 rounds and average contribution in FRAMING, and $p \leq 10\%$ for the first round and the overall number of destructive decisions.

Table 2.3: Probabilities of continuing a feud

| Feud length | Ratio (%) |
|-------------|----------------|
| 1 | 51/111 (45.9%) |
| 2 | 26/50 (52.0%) |
| 3 | 19/26 (73.1%) |
| 4 | 16/19 (84.2%) |
| 5 | 13/15 (86.7%) |
| 6 | 10/13 (76.9%) |
| 7 | 7/10 (70.0%) |
| 8+ | 75/76 (98.7%) |

The most restrictive definition of a sour relationship is to only consider instances of mutual destruction by both members of a dyad at the same time. Figure 2.4 shows these instances and separates them by the number of rounds that they survive, linking up with the feud-literature (Nikiforakis and Engelmann, 2011; Nikiforakis et al., 2012). A majority of such relationships only survives for one period, but a total of 51 develops past the first round. Since the experiment ended after 35 rounds we report the cases in which conflict is cut off at the end separately. Table 2.3 further shows the probability with which mutual destructiveness proceeds after different numbers of rounds, now denoted as feuds. We see that the cases which survive past the first round have increasing probabilities of proceeding further. Table 2.4 reports the number of distinct dyads that face at least one feud. We also ran a regression of the probability of mutual punishment in the aggregate data using a dyad-level logit model with random effects and using dummies indicating a past series of exactly one, two, or three – but not more –, and four or more consecutive cases of mutual destruction in each dyad, together with two similar variables for only one subject being destructive as explanatory variables. All these coefficients are highly significant, those for both having been destructive are bigger than those for only one having been destructive,

and they increase in size from two consecutive rounds onwards¹¹.

Table 2.4: Number of distinct dyads with at least one feud of a certain length

| Feud length | SYM | FRAME | ASYM | Total |
|---------------------|---------|-------|---------|----------|
| Feuds of any length | 16 (52) | 4 (6) | 18 (57) | 38 (115) |
| Length = 1 | 14 (29) | 3 (5) | 16 (30) | 33 (64) |
| Length = 2 | 9 (13) | 0 (0) | 7 (12) | 16 (25) |

Note: First number reflects distinct dyads that are involved in feuds. Overall number of feuds in parentheses.

2.4 Conclusion

This study shows that substantial destructive behavior can occur even in a public good environment once the opportunity to do so is present. Our baseline Fragile Public Good game – offering players equal room to take from or contribute to a public good, against symmetric marginal costs – showed more than 10% destructive decisions. While, unexpectedly, positive framing had significant negative effects on contributing in the early rounds of the game, players compensated for that later on, such that on average fewer destructive decisions were observed. Introducing a slight asymmetry by separating the Nash outcome from the initial status quo towards taking one token sharply increased the share of destructive decisions to more than 20% (even in the last round). Finally, we show that feuds can occur in a setting without separate punishment and it becomes increasingly difficult to exit them as they last longer.

¹¹Coefficients between 4.4 and 6.9 (s.e. between 0.47 and 0.65). This specification is problematic due to the dynamic nature of the regressors, which we try to tackle by using variables that are unique within a feud. Alternative approaches, including a subject-level based random effects model with additional explanatory variables and with subjects nested in dyads as random effects produced qualitatively similar results in that longer running periods of mutual destruction generally led to higher increases in the probability of feud progression.

2.5 Appendix: Instructions and Test Questions

2.5.1 Instructions Ring Test

Introduction

Welcome to this session, which will take approximately 1.5 hours. We kindly ask you to refrain from any communication with other participants from this point onwards. If you have a question please raise your hand and wait until an experimenter comes to your computer to help you.


You will be asked to make several decisions which will influence the amount of money you will be earning.

Throughout the session, your earnings are expressed in an Experimental Monetary Unit (MU). Your earnings will be converted into euros at the end of the session when you will be paid out confidentially and in private. The exchange rate is 700 MU = 1 euro.

Your earnings and your decisions are completely anonymous.

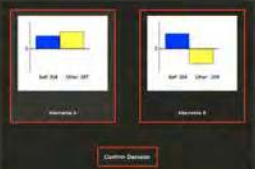
[Click here to proceed](#)

Before we start with the experiment, you are asked to perform a simple task which will earn you money and takes about 15 minutes. We will now show you the instructions for this task.

1/2 



TASK INSTRUCTIONS

In this task you will have to make 32 decisions. Each decision concerns a choice between two alternatives: Alternative A and Alternative B. Each alternative consists of two amounts of money expressed in monetary units (MU): an amount you will get (+) or lose (-) yourself and an amount another person will get (+) or lose (-). Below you see an example of two alternatives:



You can choose alternative A or alternative B by selecting them with your mouse. When you are satisfied with your choice, click on the "Confirm Decision" button and proceed to the next question.

The "Other" in these alternatives is a randomly and anonymously selected other participant in this experiment. During and after the experiment you will never learn who this participant is and this other participant will never learn who you are. You will **not** interact with this participant again after this task.


 2/3 

Every decision you make determines the number of MU that you allocate to yourself and the amount of MU that you allocate to this other participant. In the example on the previous page, if you would choose Alternative A, you gain 304 units and the other gains 397 units. If you choose Alternative B, you will gain 354 units while the other loses 354 units. Choices are confidential. The other participant will never get any information about your choices, nor will you get any information about his or her choices.

Your total earnings are determined by your own choices and the choices of the other participant. You will get paid the amount you allocate to yourself plus the amount that is allocated to you by this other participant. Your earnings will only be revealed to you at the end of the entire session, when they are paid out in cash (recall that the exchange rate is 700 MU per euro).

Please let us know if you have any questions regarding these instructions. When you are confident that you understand everything, click to proceed.

[Click here to proceed](#)

 3/3

2.5.2 Instructions and Test Questions Main Part, SYM

EXPERIMENT

The experiment consists of two parts, that are completely independent of each other. Your decisions in part 1 will not influence your potential earnings in part 2.

In each part you will be able to earn money, where it again holds that your earnings are expressed in Monetary Units with an exchange rate of 700 units = 1 Euro.

We will now start with the instructions for part 1.

1/8 ➡

Instructions Part 1

In this part you will be matched with a randomly selected other participant, which is different from the "Other" in the task you just finished.

This part will consist of 4 rounds. In all these rounds you will be paired with the same other participant.

← 2/8 ➡

Decisions and Earnings

In each round you will be asked to make a decision.

This decision always concerns the allocation of tokens over two accounts: a Private Account and a Common Account.

Tokens in the Private Account lead to earnings for yourself only.

Tokens in the Common Account lead to earnings not only for yourself but also for the participant you are paired with.

The participant you are paired with has to take a similar decision; tokens in his or her Private Account lead to earnings for her- or himself only, while tokens in the Common Account lead to earnings not only for this participant but also for you.

← 3/8 ➡

Instructions Part 1

At the beginning of each round there will be 7 tokens in your Private Account and 14 tokens in the Common Account (that you share with the other participant). The other participant will also have 7 tokens in his/her private account.

Each token in your Private Account earns you 10 MU, while every token in the Common Account earns you as well as the participant you are paired with 10 MU.

In every round you have the possibility to either take tokens from or contribute tokens to the Common Account. Tokens that you contribute to the Common Account are subtracted from your Private Account, while tokens that you take from the Common Account are added to your Private Account.

You can both take and contribute a maximum of 7 tokens. Note, however, that taking or contributing tokens comes with a transfer cost that changes with the number of tokens. The first token that you take or contribute costs you 2 MU. The second token costs you 4 MU (so, 2 tokens cost in total: $2+4=6$ MU), the third token costs you 4 MU extra (so, 3 tokens cost in total: $2+4+6=12$ MU), etc.

The same holds for the contribution of tokens to the Common Account, with the first token costing you 2 MU, the second token 4 MU, the third 6 MU, and so forth.

The next table shows the extra transfer costs and total transfer costs for tokens contributed or taken::

← 4/8 ➡

Transfer Costs of Taking/Contributing

| Token Number | Transfer Cost | Total Transfer Costs |
|--------------|---------------|----------------------|
| 1 | 2 | 2 |
| 2 | 4 | 6 |
| 3 | 6 | 12 |
| 4 | 8 | 20 |
| 5 | 10 | 30 |
| 6 | 12 | 42 |
| 7 | 14 | 56 |

Thus, for example:

- Contributing 3 tokens to the Common Account will increase your earnings from the Common Account by $3 \times 10 = 30$ MU. You will lose a total of 12 MU for the cost of transferring the tokens and lose $3 \times 10 = 30$ MU because the tokens are no longer in your private account. In total your earnings decrease by 12 MU, while the other participant's earnings increase by $3 \times 10 = 30$ MU.
- Taking 3 tokens from the Common Account will decrease your earnings from the Common Account by $3 \times 10 = 30$ MU. You will lose a total of 12 MU for the cost of transferring the tokens and earn $3 \times 10 = 30$ MU because the tokens are now in your private account. In total your earnings decrease by 12 MU, while the other participant's earnings decrease by $3 \times 10 = 30$ MU.

In short: both contributing and taking only one token will decrease your earnings by 2 MU. If you take or contribute another (second) one this will cost you 4 extra, etc. However, for every token you take the other participant loses 10 MU, while for every token you contribute s/he will gain 10 MU.

← 5/8 ➡

The figure below shows the effect of taking or contributing extra tokens on your own and the other's earnings. Each arrow represents the contribution or taking of one additional token, with the effect of this on you being shown on top of the arrow. The effect on the other is shown below the arrow. During the experiment a hard copy of this figure will be available to you.

For example:

- If you decide to contribute 4 instead of 3 tokens you will lose an extra 8 MU. The total effect of contributing 4 on your earnings would be equal to a loss of 20 MU (namely: $-2-4-6-8=-20$), whereas the total effect on the other would be a gain of 40 MU ($10+10+10+10=40$).
- If you decide to take 3 instead of 2 tokens you will lose an extra 6 MU. The total effect of taking 3 units would be equal to a loss of 12 MU (namely: $-2-4-6=-12$), whereas the total effect on the other would be a loss of 30 MU ($-10-10-10=-30$).

← 5/8 ➡

The table below shows your earnings (in MU) for all the possible choices you and the participant you are paired with can make. This table will also be available to you during the experiment.

We will explain on the next page how you should read the table.

| | | Other's possible choice | | | | | | | | | | | | | | |
|-------------|---------------|-------------------------|-----|-----|-----|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | Take | | | | | Contribute | | | | | | | | | |
| Your choice | Your earnings | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | 7 | 84 | 96 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 |
| 6 | 96 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | |
| 5 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | |
| 4 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | |
| 3 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | |
| 2 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | |
| 1 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | |
| 0 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | |
| 1 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | |
| 2 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | |
| 3 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | |
| 4 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | |
| 5 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | |
| 6 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | 396 | |
| 7 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | 396 | 408 | |



If you take 2 and the other contributes 1 your earnings are shown at the intersection of row 'take 2' and column 'contribute 1', which is 214 MU.

The other participant has the same table. Therefore, to see his/her earnings you have to switch between you and other. Thus, if you take 2 and the other contributes 1, the earnings of the other are at the intersection of row 'contribute 1' and column 'take 2' and is equal to 188 MU. To enable you to check your understanding of these instructions we will now ask you to complete a few exercises.

In case you have a question, please raise your hand, and we will come to your desk.

[Click here to proceed](#)



Exercises

If you take 5 tokens, compared to choosing 0, how would YOUR earnings change (assuming the other participant's choice stays the same)? Please use a minus sign in case of a loss.

Enter your response below using the keyboard and confirm to proceed.



Question 1/8

[Click here to proceed](#)

Exercises

If you take 5 tokens, compared to choosing 0, how would THE OTHER'S earnings change (assuming the other participant's choice stays the same)? Please use a minus sign in case of a loss.

Enter your response below using the keyboard and confirm to proceed.



Question 2/8

[Click here to proceed](#)

Exercises

If you contribute 1 and the other participant takes 2: how many MU would YOU get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

| | | Other's possible choice | | | | | | | | | | | | | | |
|-------------|---------------|-------------------------|-----|-----|-----|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | Take | | | | | Contribute | | | | | | | | | |
| Your choice | Your earnings | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | 7 | 84 | 96 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 |
| 6 | 96 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | |
| 5 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | |
| 4 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | |
| 3 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | |
| 2 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | |
| 1 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | |
| 0 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | |
| 1 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | |
| 2 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | |
| 3 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | |
| 4 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | |
| 5 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | |
| 6 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | 396 | |
| 7 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | 396 | 408 | |

Question 3/6

[Click here to proceed](#)

Exercises

If you contribute 1 and the other participant takes 2: how many MU would THE OTHER get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

| | | Other's possible choice | | | | | | | | | | | | | | |
|-------------|---------------|-------------------------|-----|-----|-----|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | Take | | | | | Contribute | | | | | | | | | |
| Your choice | Your earnings | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | 7 | 84 | 96 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 |
| 6 | 96 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | |
| 5 | 108 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | |
| 4 | 120 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | |
| 3 | 132 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | |
| 2 | 144 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | |
| 1 | 156 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | |
| 0 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | |
| 1 | 168 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | |
| 2 | 180 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | |
| 3 | 192 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | |
| 4 | 204 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | |
| 5 | 216 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | |
| 6 | 228 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | 396 | |
| 7 | 240 | 252 | 264 | 276 | 288 | 300 | 312 | 324 | 336 | 348 | 360 | 372 | 384 | 396 | 408 | |

Question 4/6

[Click here to proceed](#)

Advice

To calculate your or the other participant's earnings in a round it may be helpful to use the following procedure:

1. Number of tokens in Private Account \times 10 MU = ... MU (if you don't take or contribute anything there are 7 tokens in your Private Account)
2. Number of tokens in Common Account \times 10 MU = ... MU (if neither you nor the other takes or contributes anything there are 14 tokens in the Common Account)
3. Sum the MUs from steps 1 and 2
4. Calculate the transfer cost of taking/contributing. Recall that the transfer cost of taking or contributing is 2 MU for the first token, and increases with 2 MU for each extra token (thus, as of the second token, becoming: 4, 6, 8, 10, 12, 14).
5. Subtract the MUs from step 4 from those in step 3

Use this procedure to answer the next two questions, for which we will neither show you the table nor the figure.

[Click here to proceed](#)

Exercises

If you take 1 and the other participant takes 3: how many MU would YOU get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

Question 5/6

[Click here to proceed](#)

Exercises

If you take 1 and the other participant takes 3: how many MU would THE OTHER get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

Question 6/6

[Click here to proceed](#)

Decision

You finished the exercises. We will now show you how to make your decisions in the experiment

In each round you will see a screen such as the one depicted below. You are asked to select your choice with the mouse and confirm it by clicking on the button at the bottom.

Your Decision

1
 2
 3
 4
 5
 6
 7

Take

0

1
 2
 3
 4
 5
 6
 7

Contribute

[Click here to proceed](#)

You will then be asked what you think the other participant will choose. Please try to give your best guess of what the other participant will choose.

What do you think the other will decide?

1
 2
 3
 4
 5
 6
 7

Take

0

1
 2
 3
 4
 5
 6
 7

Contribute

[Click here to proceed](#)

2.5.3 Instructions and Test Questions Main Part, FRAME

EXPERIMENT

The experiment consists of two parts, that are completely independent of each other. Your decisions in part 1 will not influence your potential earnings in part 2.

In each part you will be able to earn money, where it again holds that your earnings are expressed in Monetary Units with an exchange rate of 700 units = 1 Euro.

We will now start with the instructions for part 1.

1/8 →

Instructions Part 1

In this part you will be matched with a randomly selected other participant, which is different from the "Other" in the task you just finished.

This part will consist of 4 rounds. In all these rounds you will be paired with the same other participant.

← 2/8 →

Decisions and Earnings

In each round you will be asked to make a decision.

This decision always concerns the allocation of tokens over two accounts: a Private Account and a Common Account.

Tokens in the Private Account lead to earnings for yourself only.

Tokens in the Common Account lead to earnings not only for yourself but also for the participant you are paired with.

Every round comes with a fixed cost of 56 MU. This also holds for the participant you are paired with.

The participant you are paired with has to take a similar decision; tokens in her or his Private Account lead to earnings for her- or himself only, while tokens in the Common Account lead to earnings not only for this participant but also for you.

← 3/8 →

Instructions Part 1

At the beginning of each round there will be 14 tokens in your Private Account and 0 tokens in the Common Account (that you share with the other participant). The other participant will also have 14 tokens in his/her private account.

Each token in your Private Account earns you 10 MU, while every token in the Common Account earns you as well as the participant you are paired with 10 MU.

In every round you have the possibility to contribute tokens to the Common Account. Tokens that you contribute to the Common Account are subtracted from your Private Account.

You can contribute a maximum of 14 tokens. Note that contributing tokens comes with a transfer gain/cost that changes with the number of tokens. The first token you contribute comes with a gain of 14 MU, the second with a gain of 12 MU, the third with a gain of 10 MU..... this continues until the seventh token that comes with a gain of 2 MU. The 8th token however comes with a cost of 2 MU, the 9th with a cost of 4 MU, until the 14th token that comes with a cost of 14 MU.

Remember: With every token you contribute the other gains 10MU.

The next table shows the extra transfer gains/costs and total transfer gains/costs for tokens contributed:

← 4/8 →

Transfer Gains and Costs of Contributing

| Token Number | Transfer Effect | Total Transfer Effect | Token Number | Transfer Effect | Total Transfer Effect |
|--------------|-----------------|-----------------------|--------------|-----------------|-----------------------|
| 1 | 14 | 14 | 8 | -2 | 54 |
| 2 | 12 | 26 | 9 | -4 | 50 |
| 3 | 10 | 36 | 10 | -6 | 44 |
| 4 | 8 | 44 | 11 | -8 | 36 |
| 5 | 6 | 50 | 12 | -10 | 26 |
| 6 | 4 | 54 | 13 | -12 | 14 |
| 7 | 2 | 56 | 14 | -14 | 0 |

Thus, for example:

- Contributing 3 tokens to the Common Account will lead to transfer gains of 14+12+10=36 MU for you. This is also the total effect on your earnings as tokens in the Common Account are worth as much to you as tokens in your private account. It also leads to a gain of 3x10=30 MU for the other participant as tokens in the Common Account are also worth 10 MU for him/her.

In short: Contributing becomes less profitable (increasingly costly), the more you contribute. Tokens in your Private Account earn you as much as tokens in the Common Account. However, for every token you contribute the other participant gains 10 MU.

← 5/8 →

The figure below shows the effect of contributing extra tokens on your own and the other's earnings. Each arrow represents the contribution of one additional token, with the effect of this on you being shown on top of the arrow. The effect on the other is shown below the arrow. During the experiment a hard copy of this figure will be available to you on your desk.

← 6/8 →

For example:

- If you decide to contribute 11 instead of 10 tokens you will lose an extra 8 MU. The total effect of contributing 11 on your earnings would be equal to a gain of 36 (namely: 14+12+...+2-2-4-6-8=36), whereas the total effect on the other would be a gain of 110 MU (11x10=110).

← 6/8 →

Advice

To calculate your or the other participant's earnings in a round it may be helpful to use the following procedure:

1. Number of tokens in Private Account x 10 MU = ... MU (if you don't contribute anything you have 14 tokens in your Private Account)
2. Number of tokens in Common Account x 10 MU = ... MU (if neither you nor the other contributes anything there are 0 tokens in the Common Account)
3. Sum the MUs from steps 1 and 2
4. Subtract the fixed cost of 56
5. Calculate the total transfer gains of contributing. Recall that: The first token you contribute comes with a gain of 14 MU, the second with a gain of 12 MU, the third with a gain of 10 MU,....., this continues until the seventh token that comes with a gain of 2 MU. The 8th token however comes with a cost of 2 MU, the 9th with a cost of 4 MU, until the 14th token that comes with a cost of 14 MU.
6. Add the MUs from step 5 to those in step 4

Use this procedure to answer the next two questions, for which we will neither show you the table nor the figure.

[Click here to proceed](#)

Exercises

If you contribute 6 and the other participant contributes 4: how many MU would YOU get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

Question 5/6

[Click here to proceed](#)

Exercises

If you contribute 6 and the other participant contributes 4: how many MU would THE OTHER get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

Question 6/6

[Click here to proceed](#)

Decision

You finished the exercises. We will now show you how to make your decisions in the experiment

In each round you will see a screen such as the one depicted below. You are asked to select your choice with the mouse and confirm it by clicking on the button at the bottom.

Your Decision

Contribute:

| | | | | | | | |
|---|---|----|----|----|----|----|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | 9 | 10 | 11 | 12 | 13 | 14 | |

[Confirm Decision](#)

[Click here to proceed](#)

You will then be asked what you think the other participant will choose. Please try to give your best guess of what the other participant will choose.

What do you think the other will decide?

Contribute:

| | | | | | | | |
|---|---|----|----|----|----|----|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | 9 | 10 | 11 | 12 | 13 | 14 | |

[Confirm Decision](#)

[Click here to proceed](#)

2.5.4 Instructions and Test Questions Main Experiment, ASYM

EXPERIMENT

The experiment consists of two parts, that are completely independent of each other. Your decisions in part 1 will not influence your potential earnings in part 2.

In each part you will be able to earn money, where it again holds that your earnings are expressed in Monetary Units with an exchange rate of 700 units = 1 Euro.

We will now start with the instructions for part 1.

1/8 →

Instructions Part 1

In this part you will be matched with a randomly selected other participant, which is different from the "Other" in the task you just finished.

This part will consist of 4 rounds. In all these rounds you will be paired with the same other participant.

← 2/8 →

Decisions and Earnings

In each round you will be asked to make a decision.

This decision always concerns the allocation of tokens over two accounts: a Private Account and a Common Account.

Tokens in the Private Account lead to earnings for yourself only.

Tokens in the Common Account lead to earnings not only for yourself but also for the participant you are paired with.

The participant you are paired with has to take a similar decision; tokens in his or her Private Account lead to earnings for her- or himself only, while tokens in the Common Account lead to earnings not only for this participant but also for you.

← 3/8 →

Instructions Part 1

At the beginning of each round there will be 7 tokens in your Private Account and 14 tokens in the Common Account (that you share with the other participant). The other participant will also have 7 tokens in his/her private account.

Each token in your Private Account earns you 11 MU, while every token in the Common Account earns you as well as the participant you are paired with 10 MU.

In every round you have the possibility to either take tokens from or contribute tokens to the Common Account. Tokens that you contribute to the Common Account are subtracted from your Private Account, while tokens that you take from the Common Account are added to your Private Account.

You can both take and contribute a maximum of 7 tokens. Note, however, that taking or contributing tokens comes with a transfer cost that changes with the number of tokens. You can take one token without any cost, but taking a second token costs you 2 MU. The third token costs you 4 MU extra (so, 3 tokens cost in total: $0+2+4=6$ MU), the fourth token costs you 6 MU extra (so, 4 tokens cost in total: $0+2+4+6=12$ MU), etc.

The same holds for the contribution of tokens to the Common Account, with the first token being for free, the second token costing you 2 MU, the third 4 MU, and so forth.

The next table shows the extra transfer costs and total transfer costs for tokens contributed or taken:

← 4/8 →

Transfer Costs of Taking/Contributing

| Token Number | Transfer Cost | Total Transfer Costs |
|--------------|---------------|----------------------|
| 1 | 0 | 0 |
| 2 | 2 | 2 |
| 3 | 4 | 6 |
| 4 | 6 | 12 |
| 5 | 8 | 20 |
| 6 | 10 | 30 |
| 7 | 12 | 42 |

Reminder: Tokens in your Private Account earn you 11 MU per token, tokens in the Common Account earn both you and the other 10 MU per token

Thus, for example:

- Contributing 3 tokens to the Common Account will increase your earnings from the Common Account by $3 \times 10 = 30$ MU. You will lose a total of 6 MU for the cost of transferring the tokens and lose $3 \times 11 = 33$ MU because the tokens are no longer in your private account. In total your earnings decrease by 9 MU, while the other participant's earnings increase by $3 \times 10 = 30$ MU.
- Taking 3 tokens from the Common Account will decrease your earnings from the Common Account by $3 \times 10 = 30$ MU. You will lose a total of 6 MU for the cost of transferring the tokens and earn $3 \times 11 = 33$ MU because the tokens are now in your private account. In total your earnings decrease by 3 MU, while the other participant's earnings decrease by $3 \times 10 = 30$ MU.

In short: Taking and contributing becomes increasingly costly, the more you do it. Each token in your Private Account earns you 1 MU more than if it is in the Common Account. However, for every token you take the other participant loses 10 MU, while for every token you contribute s/he will gain 10 MU.

← 5/8 →

The figure below shows the effect of taking or contributing extra tokens on your own and the other's earnings (after the 1 MU difference in the value of tokens in the two accounts has been taken into account). Each arrow represents the contribution or taking of one additional token, with the effect of this on you being shown on top of the arrow. The effect on the other is shown below the arrow. During the experiment a hard copy of this figure will be available to you.

For example:

- If you decide to contribute 4 instead of 3 tokens you will lose an extra 7 MU. The total effect of contributing 4 on your earnings would be equal to a loss of 16 (namely: $-1-3-5-7=-16$), whereas the total effect on the other would be a gain of 40 MU ($10+10+10+10=40$).
- If you decide to take 3 instead of 2 tokens you will lose extra 3 MU. The total effect of taking 3 units would be equal to a loss of 3 (namely: $+1-1-3=-3$), whereas the total effect on the other would be a loss of 30 MU ($-10-10-10=-30$).

← 6/8 →

The table below shows your earnings (in MU) for all the possible choices you and the participant you are paired with can make. This table will also be available to you during the experiment.

We will explain on the next page how you should read the table.

| | | Other's possible choices | | | | | | | | | | | | | | | |
|-----------------------|---|--------------------------|-----|-----|-----|-----|------------|-----|----|----|----|----|----|----|----|----|----|
| | | Take | | | | | Contribute | | | | | | | | | | |
| Your possible choices | 7 | 132 | 127 | 122 | 117 | 112 | 107 | 102 | 97 | 92 | 87 | 82 | 77 | 72 | 67 | 62 | 57 |
| | 6 | 128 | 123 | 118 | 113 | 108 | 103 | 98 | 93 | 88 | 83 | 78 | 73 | 68 | 63 | 58 | 53 |
| | 5 | 124 | 119 | 114 | 109 | 104 | 99 | 94 | 89 | 84 | 79 | 74 | 69 | 64 | 59 | 54 | 49 |
| | 4 | 120 | 115 | 110 | 105 | 100 | 95 | 90 | 85 | 80 | 75 | 70 | 65 | 60 | 55 | 50 | 45 |
| | 3 | 116 | 111 | 106 | 101 | 96 | 91 | 86 | 81 | 76 | 71 | 66 | 61 | 56 | 51 | 46 | 41 |
| | 2 | 112 | 107 | 102 | 97 | 92 | 87 | 82 | 77 | 72 | 67 | 62 | 57 | 52 | 47 | 42 | 37 |
| | 1 | 108 | 103 | 98 | 93 | 88 | 83 | 78 | 73 | 68 | 63 | 58 | 53 | 48 | 43 | 38 | 33 |
| | 0 | 104 | 99 | 94 | 89 | 84 | 79 | 74 | 69 | 64 | 59 | 54 | 49 | 44 | 39 | 34 | 29 |
| | 0 | 100 | 95 | 90 | 85 | 80 | 75 | 70 | 65 | 60 | 55 | 50 | 45 | 40 | 35 | 30 | 25 |



If you take 2 and the other contributes 1 your earnings are shown at the intersection of row 'take 2' and column 'contribute 1', which is 227.

The other participant has the same table. Therefore, to see his/her earnings you have to switch between you and other. Thus, if you take 2 and the other contributes 1, the earnings of the other are at the intersection of row 'contribute 1' and column 'take 2' and are equal to 196. To enable you to check your understanding of these instructions we will now ask you to complete a few exercises.

In case you have a question, please raise your hand, and we will come to your desk.

[Click here to proceed](#)



Exercises

If you take 5 tokens, compared to choosing 0, how would YOUR earnings change (assuming the other participant's choice stays the same)? Please use a minus sign in case of a loss.

Enter your response below using the keyboard and confirm to proceed.

| | | | | | | | | |
|-----------------|-----|-----|-----|-----|-----|-----|-----|------------|
| Effect on you | +1 | -1 | -3 | -5 | -7 | -9 | -11 | Take |
| Effect on other | -10 | -10 | -10 | -10 | -10 | -10 | -10 | |
| Effect on you | -1 | -3 | -5 | -7 | -9 | -11 | -13 | Contribute |
| Effect on other | +10 | +10 | +10 | +10 | +10 | +10 | +10 | |

Question 1/6

[Click here to proceed](#)

Exercises

If you take 5 tokens as compared to choosing 0, how would THE OTHER'S earnings change (assuming the other participant's choice stays the same)? Please use a minus sign in case of a loss.

Enter your response below using the keyboard and confirm to proceed.

| | | | | | | | | |
|-----------------|-----|-----|-----|-----|-----|-----|-----|------------|
| Effect on you | +1 | -1 | -3 | -5 | -7 | -9 | -11 | Take |
| Effect on other | -10 | -10 | -10 | -10 | -10 | -10 | -10 | |
| Effect on you | -1 | -3 | -5 | -7 | -9 | -11 | -13 | Contribute |
| Effect on other | +10 | +10 | +10 | +10 | +10 | +10 | +10 | |

Question 2/6

[Click here to proceed](#)

Exercises

If you contribute 1 and the other participant takes 2: how many MU would YOU get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

| | | Other's possible choices | | | | | | | | | | | | | | | |
|-----------------------|---|--------------------------|-----|-----|-----|-----|------------|-----|----|----|----|----|----|----|----|----|----|
| | | Take | | | | | Contribute | | | | | | | | | | |
| Your possible choices | 7 | 132 | 127 | 122 | 117 | 112 | 107 | 102 | 97 | 92 | 87 | 82 | 77 | 72 | 67 | 62 | 57 |
| | 6 | 128 | 123 | 118 | 113 | 108 | 103 | 98 | 93 | 88 | 83 | 78 | 73 | 68 | 63 | 58 | 53 |
| | 5 | 124 | 119 | 114 | 109 | 104 | 99 | 94 | 89 | 84 | 79 | 74 | 69 | 64 | 59 | 54 | 49 |
| | 4 | 120 | 115 | 110 | 105 | 100 | 95 | 90 | 85 | 80 | 75 | 70 | 65 | 60 | 55 | 50 | 45 |
| | 3 | 116 | 111 | 106 | 101 | 96 | 91 | 86 | 81 | 76 | 71 | 66 | 61 | 56 | 51 | 46 | 41 |
| | 2 | 112 | 107 | 102 | 97 | 92 | 87 | 82 | 77 | 72 | 67 | 62 | 57 | 52 | 47 | 42 | 37 |
| | 1 | 108 | 103 | 98 | 93 | 88 | 83 | 78 | 73 | 68 | 63 | 58 | 53 | 48 | 43 | 38 | 33 |
| | 0 | 104 | 99 | 94 | 89 | 84 | 79 | 74 | 69 | 64 | 59 | 54 | 49 | 44 | 39 | 34 | 29 |
| | 0 | 100 | 95 | 90 | 85 | 80 | 75 | 70 | 65 | 60 | 55 | 50 | 45 | 40 | 35 | 30 | 25 |

Question 3/6

[Click here to proceed](#)

Exercises

If you contribute 1 and the other participant takes 2: how many MU would THE OTHER get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

| | | Other's possible choices | | | | | | | | | | | | | | | |
|-----------------------|---|--------------------------|-----|-----|-----|-----|------------|-----|----|----|----|----|----|----|----|----|----|
| | | Take | | | | | Contribute | | | | | | | | | | |
| Your possible choices | 7 | 132 | 127 | 122 | 117 | 112 | 107 | 102 | 97 | 92 | 87 | 82 | 77 | 72 | 67 | 62 | 57 |
| | 6 | 128 | 123 | 118 | 113 | 108 | 103 | 98 | 93 | 88 | 83 | 78 | 73 | 68 | 63 | 58 | 53 |
| | 5 | 124 | 119 | 114 | 109 | 104 | 99 | 94 | 89 | 84 | 79 | 74 | 69 | 64 | 59 | 54 | 49 |
| | 4 | 120 | 115 | 110 | 105 | 100 | 95 | 90 | 85 | 80 | 75 | 70 | 65 | 60 | 55 | 50 | 45 |
| | 3 | 116 | 111 | 106 | 101 | 96 | 91 | 86 | 81 | 76 | 71 | 66 | 61 | 56 | 51 | 46 | 41 |
| | 2 | 112 | 107 | 102 | 97 | 92 | 87 | 82 | 77 | 72 | 67 | 62 | 57 | 52 | 47 | 42 | 37 |
| | 1 | 108 | 103 | 98 | 93 | 88 | 83 | 78 | 73 | 68 | 63 | 58 | 53 | 48 | 43 | 38 | 33 |
| | 0 | 104 | 99 | 94 | 89 | 84 | 79 | 74 | 69 | 64 | 59 | 54 | 49 | 44 | 39 | 34 | 29 |
| | 0 | 100 | 95 | 90 | 85 | 80 | 75 | 70 | 65 | 60 | 55 | 50 | 45 | 40 | 35 | 30 | 25 |

Question 4/6

[Click here to proceed](#)

Advice

To calculate your or the other participant's earnings in a round it may be helpful to use the following procedure:

1. Number of tokens in Private Account $\times 11$ MU = ... MU (if you don't take or contribute anything there are 7 tokens in your Private Account)
2. Number of tokens in Common Account $\times 10$ MU = ... MU (if neither you nor the other take or contribute anything there are 14 tokens in the Common Account)
3. Sum the MUs from steps 1 and 2
4. Calculate the transfer cost of taking/contributing. Recall that the transfer costs of taking or contributing is 0 MU for the first token, and increases with 2 MU for each extra token (thus, as of the second token, becoming: 2, 4, 6, 8, 10, 12).
5. Subtract the MUs from step 4 from those in step 3

Use this procedure to answer the next two questions, for which we will neither show you the table nor the figure.

[Click here to proceed](#)

Exercises

If you take 1 and the other participant takes 3: how many MU would YOU get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

188

Question 5/6

[Click here to proceed](#)

Exercises

If you take 1 and the other participant takes 3: how many MU would THE OTHER get at the end of this round?

Enter your response below using the keyboard and confirm to proceed.

204

Question 6/6

[Click here to proceed](#)

Decision

You finished the exercises. We will now show you how to make your decisions in the experiment

In each round you will see a screen such as the one depicted below. You are asked to select your choice with the mouse and confirm it by clicking on the button at the bottom.

Your Decision

| | | | | | | | |
|---|---|---|---|---|---|---|------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | Take |
| 0 | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | Contribute |

[Confirm Decision](#)

[Click here to proceed](#)

You will then be asked what you think the other participant will choose. Please try to give your best guess of what the other participant will choose.

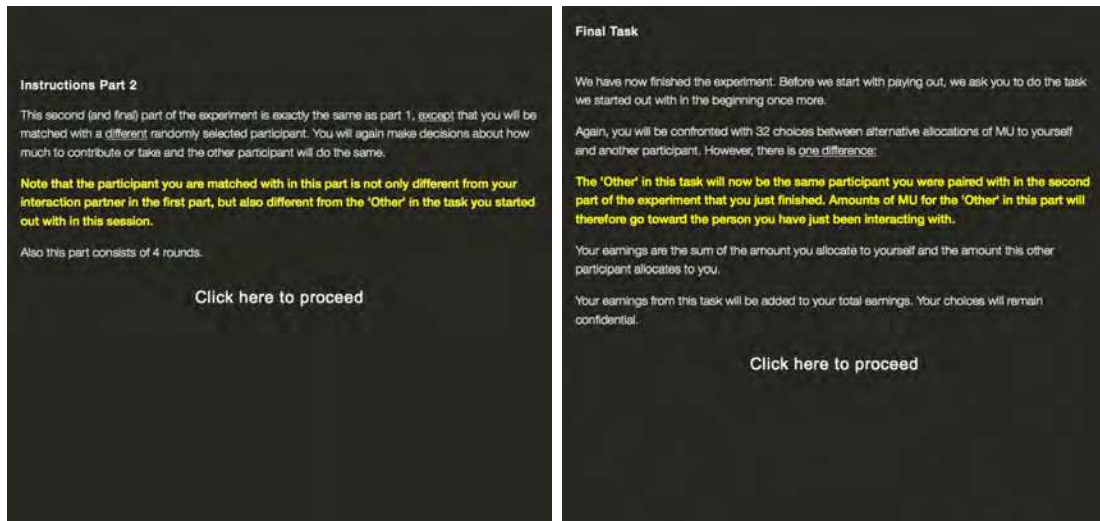
What do you think the other will decide?

| | | | | | | | |
|---|---|---|---|---|---|---|------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | Take |
| 0 | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | Contribute |

[Confirm Decision](#)

[Click here to proceed](#)

2.5.5 Interstitial Instruction Screens



Chapter 3

Asymmetry in the Development of Cooperative and Antagonistic Relationships. A Model-Based Analysis of a Fragile Public Good Game Experiment¹

¹This chapter is based on work with Ben Loerakker, Nadège Bault, and Frans van Winden, see Loerakker et al. (2016). Financial support by the Research Priority Area Behavioral Economics of the University of Amsterdam is gratefully acknowledged.

3.1 Introduction

Public good experiments typically display more cooperation than predicted by rational and selfish preferences. The cooperation levels, though, vary depending on the exact design, per individual, and are often diminishing over time.² Various explanations have been offered, among which: altruism (Andreoni and Miller, 2002), reciprocity (Falk and Fischbacher, 2006), learning (Roth and Erev, 1995), and conditional cooperation (Fischbacher et al., 2001). Recently, both direct behavioral (van Dijk et al., 2002; Bault et al., 2017) as well as neurobiological (Bault et al., 2014) evidence has been provided for an alternative explanation involving the affective tie-mechanism introduced by van Dijk and van Winden (1997). Their model is characterized by affective interpersonal ties. Simply put, a person (i) takes the welfare of another person (j) into account according to the behavior of this other person. This is a dynamic process whereby every action of j is compared to a reference action, if this action is more beneficial to i than the reference action a positive affective impulse is generated that could change the importance i attaches to the welfare of j . This means that this process is not only dynamic but also allows for an asymmetry between the development of the weight i attaches to the welfare of j and the weight j attaches to the welfare of i .

Although this model seems rather successful in tracking behavior in the public good games examined, there are several issues that need to be addressed. First, work on the tie mechanism so far has focused on cooperative interpersonal relationships, leaving negative (hate) relationships underexplored. This is important because there exists a lot of evidence that negative behaviors and (hate) relationships exist.³ Second, it is not examined whether people react differently to positive versus negative behavior of others. The difference in direction is obvious, but how about the size of the action? Third, the tie model has

²See Chaudhuri (2011) and Plott and Smith (2008) for an overview.

³Abbink and Sadrieh (2009) and Bosman and van Winden (2002) show that people are willing to destroy other people's earnings even if it does not lead to higher earnings for themselves. In the same experiment Bosman and Van Winden also show that negative emotions are involved if money is taken away from participants and when players engage in the destruction of the others' earnings by destructing their own earnings. More recently Bolle et al. (2013) show that participants are willing to decrease the chances of others to win a prize. Furthermore they find that in a repeated game setting players retaliate harmful actions and that this retaliation is driven by negative emotions caused by these harmful actions.

not been investigated in a horse race with other models using out-of-sample predictions regarding the same game⁴. Fourth, it is not clear whether the tie model would be helpful in explaining behavior across different contexts. For example could it integrate behavior observed in public good games with behavior in repeated prisoner's dilemma games?⁵

In this chapter we will address each of these four issues. To address the first two, a novel game design is used: the fragile public good game (FPG) game. In this game there is as much room for antagonistic behavior as there is for cooperative behavior. The third issue is addressed by designing an experiment with two independent parts, which allows us to predict behavior in the second part using parameter estimates from the first part. This creates proper out-of-sample predictions for the different social preferences as well as learning models that will be explored. Finally, we show that a simple two parameter tie model can mimic observed behavioral rules like tit-for-tat and is able to explain why and how players switch rules as the parameters of a prisoner's dilemma (PD) game change. Next, the parameters estimated on the behavioral data from the FPG game are investigated to see what kind of behavior they would predict in different repeated PD settings.

Furthermore, earlier findings in terms of the relative importance of the different tie parameters are confirmed. We find evidence that people react stronger to positive behavior of others than to negative behavior. This might be one of the driving factors of repeated cooperation. Another important finding is, that the tie model predicts significantly better than other models, including of social preferences and the reinforcement learning model of Roth and Erev (1995). Finally, a tie model with just two parameters seems well able to explain results found in different repeated PD environments.

Section 2 introduces the FPG game, provides a theoretical analysis using the tie model and presents our hypothesis. Section 3 describes the experimental and estimation methods used, while section 4 presents our results. Section 5 applies the tie model to the repeated prisoner's dilemma game, and section 6 concludes.

⁴See Bault et al. (2017)

⁵See for instance Dal Bó and Fréchette (2011) and Fudenberg et al. (2012)

3.2 Theory

3.2.1 Fragile Public Good Game

In order to experimentally address the questions raised regarding negative ties we designed the Fragile Public Good (FPG) game, a two-player game that allows players to financially hurt as well as help the other player. A key feature of the FPG game is that it gives as much room for destructive behavior (taking) as for constructive behavior (contributing) regarding a public good. This is achieved by having both the (standard) Nash equilibrium and the status quo - i.e., the initial allocation of tokens to the common account - in the middle of the action space. With, in addition, full symmetry in the marginal cost of taking and contributing, this leaves substantial leeway for the development of negative as well as positive ties. There are relatively few public good experiments with an interior Nash equilibrium (Laury and Holt, 2008), and, to the best of our knowledge, no such experiments that allow for as much destructive as cooperative behavior. This game enables us to estimate the parameter values of our model. Furthermore, by maintaining comparability with ordinary (non-linear) public goods games, we can compare our results with existing studies of such games. By using a repeated game where in the first part players interact with a fixed partner, but are then rematched randomly with a new partner for playing in the second part, we can investigate the out-of-sample predictive performance of our estimated model.

More specifically, both players in our FPG game are endowed with 7 tokens in their private account, while sharing a common account containing 14 tokens at the beginning of every round. Each token stored in the private account generates 10 MU for the player concerned, whereas a token in the common account generates 10 MU for both players. Each round, both players simultaneously decide whether to contribute tokens to the common account or to take tokens from the common account. They can transfer up to 7 tokens per round from the common account to their private account, or the other way around.

Transferring tokens in either direction comes at a marginal cost, that increases with 2 MU per token. The transfer of the first token thus costs 2 MU, transferring a second

one costs an additional 4 MU (for a total of $2+4=6$ MU), transferring a third token leads to a total cost of 12 MU ($2+4+6$), and so forth. The effect of contributing the first token is, thus, that the other player receives 10 MU while the contributing player gets 2 MU less. By contributing a second token a player generates another 10 MU for the other player at a cost of 4 MU, et cetera, until the seventh token which earns the other player still 10 MU while it costs the contributing player 14 MU to transfer. Taking tokens has the exact same effect on the transferring player as contributing the same amount of tokens would have. For the other player, however, the effect is the exact opposite: He or she will now lose 10 MU per token instead of gaining 10 MU. Because the only difference between taking and contributing concerns the development of, respectively, a negative and a positive externality, this game allows us to study, in a clean way, whether an asymmetry exists in the impact of hurting behavior (taking) and helping behavior (contributing).

Making no transfer may be seen as a reference point as it accords with the status quo as well as the standard Nash best response. Moreover, it may easily attract a player's attention in the payoff matrix of the game (see appendix D). We will return to this below. The game is a non-linear public good game with an internal social optimum, where both players contribute either 4 or 5 tokens. The similarities with a more conventional public good game become even clearer when one sees taking seven tokens as the starting point, so one can only contribute. In that case the stage game becomes similar to a public good game with diminishing returns to contributing, albeit with an internal standard Nash equilibrium and an internal social optimum.

3.2.2 Model

We use an adapted version of the tie model of Bault et al. (2017). In this model α_{ijt} captures the tie that i has with j at time t , and formally expresses the weight that i attaches to the utility (payoff) of j . These ties are personal, dynamic and do not necessarily have to be symmetric.

We start from the basic model in which players have the following interdependent utility

function:

$$V_{it} = U_{it} + \alpha_{ijt}U_{jt} \quad (3.1)$$

Here V_{it} denotes the (extended) utility function of player i at time t , while U_{it} and U_{jt} indicate the payoffs of i and j , respectively, at time t .

Players do not only take the current period into account, but also the subsequent one (one-period forward looking behavior). Empirical evidence suggests that players are rather myopic (see e.g. Bone et al. (2003) and Bone et al. (2004)). This leads us to the following simple extension of (3.1):

$$V_{it} = U_{it} + \alpha_{ijt}U_{jt} + (U_{it+1} + \alpha_{ijt}U_{jt+1}) \quad (3.2)$$

For the FPG game, letting C_{it} stand for i 's contribution to the common account, i 's expected payoff (U_{it}^e) can be written as :

$$U_{it}^e = 210 - C_{it}^2 - |C_{it}| + 10C_{jt}^e$$

With (3.3)

$$-7 \leq C_{it} \leq 7$$

Including future periods in the player's utility function does not affect C_{it} if i does not expect to be able to influence j 's next period contribution. Players may believe, however, that (some) other players are imitators or conditional cooperators (Fehr and Fischbacher, 2003). Therefore, we assume the following relationship until the last round (as there is no future left in the final round):

$$C_{jt+1}^e = \gamma_i C_{it} + (1 - \gamma_i) C_{jt}^e$$

With (3.4)

$$0 \leq \gamma \leq 1$$

From these equations it follows that the optimal contribution for player i depends on both the parameter γ_i , indicating how strong agent i believes he can influence agent j , and α_{ijt} ,

which expresses the weight i assigns to the payoff for j . We specify the latter as:

$$\alpha_{ijt} = \delta_{1i}\alpha_{ijt-1} + \delta_{2i}I_{ijt-1} \quad (3.5)$$

With I_{ijt-1} standing for an impulse determined by the difference between j 's last round contribution and a reference contribution. In this chapter, though, we will use the next, more general specification, which differentiates between positive and negative impulses:

$$\alpha_{ijt} = \begin{cases} \delta_{1i}\alpha_{ijt-1} + \delta_{2Ni}I_{ijt-1} & I_{ijt-1} \leq 0 \\ \delta_{1i}\alpha_{ijt-1} + \delta_{2Pi}I_{ijt-1} & I_{ijt-1} > 0 \end{cases} \quad (3.6a)$$

$$(3.6b)$$

Here, we assume that I_{ijt-1} equals the difference between the other player's contribution and the one-shot Nash equilibrium choice (0), based on the discussion above regarding the reference point (see also the estimation results in appendix B).

3.2.3 Model analysis and hypotheses

An equilibrium is defined by a situation where both players have no incentive to change their contribution. We will now discuss the conditions and nature of potential equilibria. To that purpose, we start by comparing the (expected) utility of two adjacent choices. Due to the fact that $V_{it}(C_{it})$ is concave, C_{it} is a best response if $V_{it}(C_{it}) \geq V_{it}(C_{it} + 1)$ and $V_{it}(C_{it}) \geq V_{it}(C_{it} - 1)$. Assuming here, for convenience, that both C_{it} and C_{jt}^e are greater than or equal to zero⁶, and omitting the subscripts of α and γ , the difference in utility equals:

$$V_{it}(C_{it} + 1) - V_{it}(C_{it}) = 10\alpha + 10\gamma - (2C_{it} + 2) - \gamma\alpha(2\gamma C_{it} + 2(1 - \gamma)C_{jt}^e + \gamma + 1) \quad (3.7)$$

Eq. (3.7) shows the costs and benefits of contributing an extra token (in the positive domain). If players are not playing strategically the costs are simply $2C_{it} + 2$, while the benefits are 10α . If players expect to be able to influence their counterpart the cost-

⁶Appendix A also addresses the case where C_{it} is smaller than zero.

benefit analysis becomes more complicated. There are benefits of 10γ , from the expected imitation or positive reciprocity by the other, as well as new costs of $\gamma\alpha(2\gamma C_{it} + 2(1 - \gamma)C_{jt}^e + \gamma + 1)$, as players with $\alpha > 0$ care about the fact that the other faces a cost of reciprocating or imitating.

This model leads to a number of propositions: first of all, it turns out that contributions outside of the interval $[-5, 5]$ can never be part of any equilibrium for conventional values of α between 1 and -1. This result is important as it shows that the bounds of the decision space are not part of any equilibrium in that case, which is helpful for estimating the model. For instance, suppose we would like to estimate α in the myopic model, then, if a player repeatedly made boundary decisions ($C = 7$ or $C = -7$) we would only have information about respectively, the lower bound and the upper bound of the α parameter.

Proposition 1. *Contributions outside of the interval $[-5, 5]$ can never be part of any equilibrium if $-1 \leq \alpha \leq 1$.*

Proof in appendix A, section 3.7.1

Next, focusing first on symmetric equilibria ($C_{it} = C_{jt}$), we arrive at the following proposition:

Proposition 2. *Any contribution level where $C_{it} = C_{jt} \in [-5, 5]$ can be part of an equilibrium.*

Proof in appendix A, section 3.7.2.

An asymmetric equilibrium is less likely as one player is then always worse off than the other player. Theoretically, asymmetric equilibria are not impossible, though, but the parameter constraints are more restrictive than for symmetric ones. The farther the different contributions are apart the more extreme the conditions for these equilibria become. The following proposition refers to their existence:

Proposition 3. *Asymmetric equilibria exist if either $-5 \leq C_{it}, C_{jt} < 0$ or $0 \leq C_{it}, C_{jt} \leq 5$.*

Proof in appendix A, section 3.7.3.

Our last proposition establishes a parameter restriction for efficient cooperation. For convenience, we restrict ourselves here to the myopic model as this will turn out to be the most relevant model in our study. Similar restrictions including γ could be derived for the model that allows for forward looking behavior (see Appendix A, section 3.7.2).

Proposition 4. *For the socially optimal choices to be part of an equilibrium under the myopic model, the parameters of both players tie mechanism should satisfy the restriction:*

$$0.2 \leq \frac{\delta_{2i}}{1-\delta_{1i}} \leq 0.25, \text{ which is a necessary but not a sufficient condition.}$$

Proof in appendix A, section 3.7.4.

The intuition for this results is that if players have a $\frac{\delta_2}{1-\delta_1}$ ratio that is below 0.2 they built insufficiently strong ties. If the ratio is larger than 0.25 the opposite happens: the ties become so strong that α will grow larger than 1 implying that players will overinvest in this relationship.

Based on the propositions 1,2, and 3, our first hypothesis is:

Hypothesis 1. *If both players in a dyad do not change their contribution for multiple consecutive rounds, both contribute an equal amount and this contribution lies between -5 and 5 (inclusive)*

Our second hypothesis is motivated by the earlier mentioned work of Baumeister et al. (2001) and Baumeister and Leary (1995). They collect evidence that indicates that negative experiences that coincide with negative emotions have a stronger and longer lasting impact on someone's wellbeing and behavior than positive experiences and emotions. Furthermore, Kuhnen (2015) finds that investors weigh negative news more than positive news in an experimental setting.

Hypothesis 2. *Negative impulses have a bigger impact on the weight a player allocates to the payoff of a counterpart (the social tie) than positive ones, or in the context of our model: $\delta_{2N} > \delta_{2P}$.*

The third hypothesis concerns the performance of the model, specifically its predictive accuracy. Bault et al. (2017) already investigated the comparative performance of a

ties model with $\delta_{2N} = \delta_{2P}$ within sample, where the number of parameters could be an issue. Here we apply a true out-of-sample test and compare with alternative models, now including a learning model. Therefore the final hypothesis reads:

Hypothesis 3. *When calibrated on the first FPG game our ties model gives more precise estimates of a subject's behavior in the second FPG game, as compared to competing models that are calibrated on the same data.*

Our final hypothesis is based on survey papers by Chaudhuri (2011) and Kagel et al. (1995). They find that in most public good games contribution levels are declining. If we relate these results with proposition 4, we hypothesize that most subjects will not fulfill the restrictions outlined in this proposition:

Hypothesis 4. *For the majority of the subjects $0.2 \leq \frac{\delta_{2i}}{1-\delta_{1i}} \leq 0.25$ will not hold.*

3.3 Methods

3.3.1 Experiment

In this paper we are using the data from the SYM treatment that was already covered in the analysis in chapter 2. The experiment took place in November 2012 and April 2013. It consisted of 3 sessions with 130 (65 female, 2 unreported) participants. All subjects were recruited through the recruitment system of the CREED laboratory of the university Amsterdam. Students who had participated in previous public good experiments or power-to-take experiments (as recorded in the CREED recruitment system) were excluded.

The entire experiment was held in the CREED laboratory and completely computerized. In the experiment we used Monetary Units (MUs) to express the earnings of the participants, which were converted to euros by a rate of 700 MU to one euro. The average earnings in the experiment were €25.65 (there was no show-up fee as the theoretical

minimum earnings exceeded €10. No participant earned less than €15) and the sessions took about two hours.

During the experiment the participants were first asked to perform a Social Value Orientation (SVO) test (Liebrand and McClintock, 1988b), where we use the version of van Dijk et al. (2002). In this test participants decide on payoff allocations between Self and an anonymous Other. MUs allocated to Other affected the earnings of a random other participant in the experiment. Participants were informed that all their choices in the SVO test remained confidential and only learned their earnings at the end of the experiment. An example of the choices made in an SVO test can be found in appendix D, figure 3.6.

Every question of the SVO test concerns a choice between two payoff allocations. Each allocation represents a point on a circle around the origin, where the payoff to self is on the x-axis and the payoff to Other is on the y-axis. In total the participants had to make 32 of these choices in the SVO test. An angle is constructed by aggregating all the vectors spanned up by the 32 chosen payoff allocations. An individual's distributional preferences can be expressed by this angle. For example, an angle of zero degrees means that one is completely selfish, a 45 degree angle indicates that one maximizes the sum of the payoffs to Self and Other, and an angle of 90 degrees would indicate that one only cares about the payoff of Other. The size of the vector tells us how consistent the choices are. If all choices are consistent with a certain preference the size of the vector will be 1000. In the examples given above, the tangent of the angle is always positive. However, just as with the alpha in our theoretical model, also negative values are possible. The interpretation of these values is analogous to a negative α , as they indicate that a person is willing to give something up in order to decrease the payoff for Other.⁷ The tangent that results from the test can be interpreted as an indication of the initial alpha (α_0) and will be used as such later on.

After this SVO test the participants played 35 rounds of the Fragile Public Good (FPG)

⁷Situations where individuals prefer negative payoffs over positive payoffs for themselves are not taken into consideration here and very seldom observed.

game, explained above, in a partner setting (with a different partner than the "other" from the SVO task). In the introduction of the FPG game it was made clear that both taking and contributing came at a cost. In order to check if players understood the game, they had to answer quiz questions and played three trial rounds. In these trial rounds they could also get acquainted to the feedback they would receive during the actual game. After every round they saw the choice of the other player, their own payoff in the round they just played and the payoff of the other, both of which were represented using numbers as well as bars so as to visualize the difference between the payoffs.

After the first FPG game the participants were informed that a second one would follow, again with a randomly matched partner but not the one from the previous game. Also this game consisted of 35 rounds. The final task of the experiment task was another SVO test, where the other in the test was now the same as in the final FPG game. This final test will not be used here.

3.3.2 Estimation

For our estimation procedure we follow Bault et al. (2017). To close the model and to enable us to estimate it we introduce a random variable ϵ_{ik}/θ_i as a noise term, where θ_i represents the rationality or choice intensity of player i . If we now assume ϵ_{ik} to be i.i.d. and double-exponentially distributed, we arrive at a multinomial logit model. Now let π_{ikt} be the probability that a player chooses contribution k in period t , then if we multiply all these probabilities we obtain our likelihood measure:

$$\prod_t \pi_{ikt} = \prod_t \frac{e^{\theta_i V_{ikt}^e}}{\sum_h e^{\theta_i V_{iht}^e}} \quad 0 < \theta < \infty \quad (3.8)$$

Estimation requires a value for α_0 , the tie parameter prior to any interaction with the other player. In the estimation results shown in the subsequent sections we used the measure taken from the SVO test. For those participants with an inconsistent tie measure (the tie measure is considered inconsistent when the length of the vector is smaller than 600) we use $\alpha_0 = 0$. The model is estimated on the first FPG game.

In the group level estimations we set $\delta_{1i}=\delta_{1j}$, $\delta_{2i}=\delta_{2j}$ and $\theta_i=\theta_j$ for all i and j and estimate the model using Matlab’s `fmincon` optimization procedure based on the likelihood described in equation (3.8). When calculating standard errors we clustered all observations from the same individual.

3.4 Results

3.4.1 Descriptive statistics

The average angle of 128 participants in the SVO test was 6.03 degrees, which corresponds to an α value of 0.11. The observations of 2 participants were lost due to technical problems, while the choices of 8 participants were considered inconsistent because their vectors were below 600 out of 1,000 in length (see Liebrand and McClintock (1998)). The SVO tests concerning those participants are, therefore, deleted from the analyses.

A summary of the behavior during the FPG games is given below in table 3.1. We observe that average contributions are noticeably higher in the second FPG game than in the first FPG game:

Table 3.1: Descriptive statistics

| Game | FPG1 (n=130) | FPG2 (n=130) |
|------------------------------|--------------|--------------|
| Average contribution | 2.28 | 2.86* |
| Avg contribution first round | 1.26 | 2.53* |
| Avg contribution last round | 0.68 | 0.70 |
| % negative contributions | 11.3% | 4.6% |

Note: * indicates significance at the 1% level, using a Wilcoxon sign-rank test with contributions on the pair level.

This is also illustrated in figure 3.1, that shows the average contributions per round:

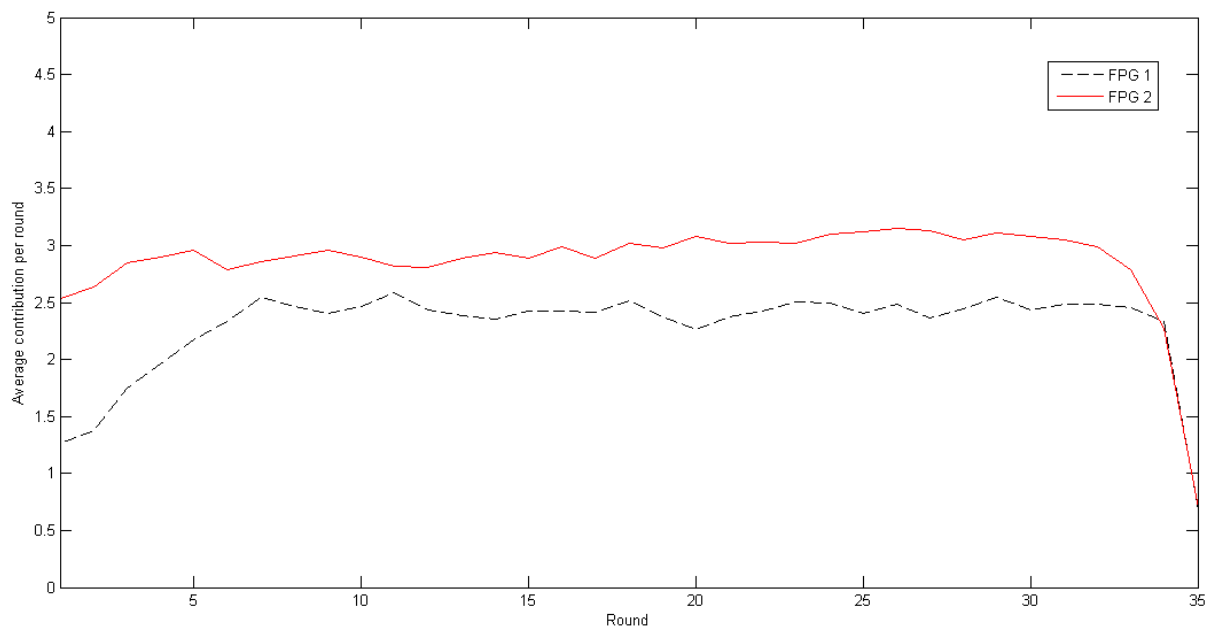


Figure 3.1: Average contributions per round

The figure above suggests that behavior in the second game is definitively influenced by the first game, as participants start off with much higher contributions. Furthermore, we observe that between rounds 7 and 33 the difference between the games is fairly constant at around 0.5, until the decline in the last couple of rounds (end-effect) leads to almost identical contributions in the end.

Another difference between the two games concerns the number of destructive decisions. While taking seems to play an important role in the first FPG game, its role in the second one is diminished noticeably. The first game, though, shows that destructive behavior can be relatively frequent even if there are plenty of opportunities to stay away from it. For illustration, figure 3.2 shows two pairs that, respectively, establish a cooperative and a sour relationship.

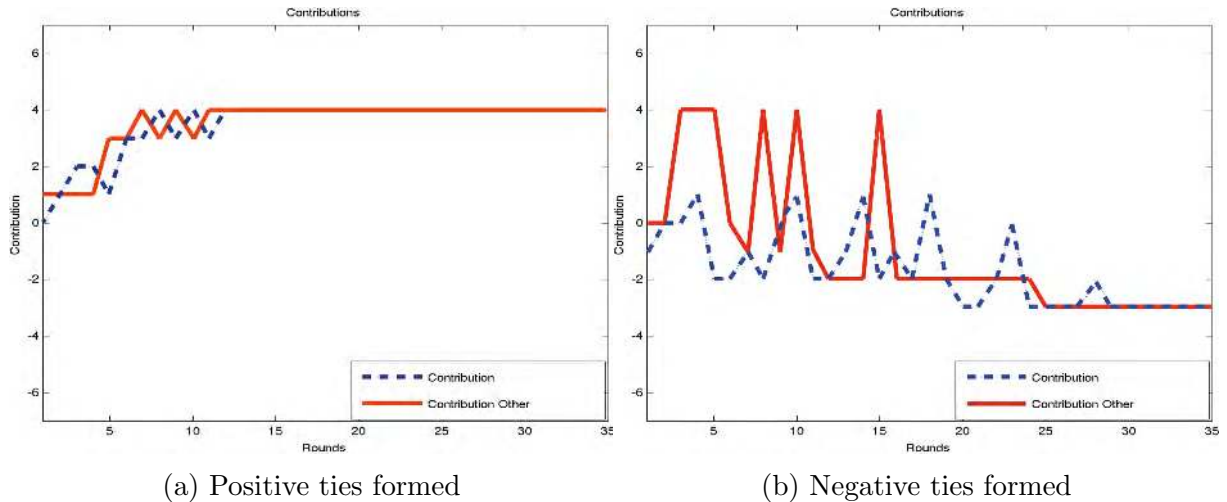


Figure 3.2: An example of a successfully built cooperative relationship and a relationship that has turned sour

We find that only 1.5% (135 out of 9100) of the contributions are either larger than 5 or smaller than -5. Moreover, we do not find any instance of two players contributing an unequal but constant amount for 3 rounds or more. This confirms our first hypothesis (H1).

The rest of this section is organized as follows: Section 3.4.2 investigates the group level estimation results, 3.4.3 is devoted to individual level estimates, while section 3.4.4 studies the predictive performance of the model.

3.4.2 Group level results

We start by estimating the myopic version of the model (labeled Myopic), represented by (3.1), neglecting for the moment forward looking behavior introduced in (3.2). This leaves a model with only 3 parameters. Despite its simplicity, this model has been found quite successful in explaining public good contributions (Bault et al., 2017).

Psychological studies (Baumeister et al., 2001) suggest that people react to negative experiences and emotions differently than to positive ones. Therefore, in our next model, we will allow for differences between the impact of negative and positive impulses, again using the myopic version of the model (M.NP), as formulated by equations (3.6a) and (3.6b).

The third model that we will investigate is the forward-looking model (FL), represented by equations (3.2) and (3.4). Finally, we estimate a model that allows for the aforementioned difference between the strength of negative and positive impulses as well as forward looking behavior (FL.NP).

The estimation results regarding these four models are found in table 3.2.

Table 3.2: Group level estimations

| | γ | θ | δ_1 | δ_2 | δ_{2P} | δ_{2N} |
|--------|------------|-------------|-------------|-------------|---------------|---------------|
| Myopic | | 0.16*(0.02) | 0.54*(0.09) | 0.10*(0.02) | | |
| M.NP | | 0.17*(0.02) | 0.49*(0.10) | | 0.12*(0.02) | 0.08*(0.02) |
| FL | 0.06(0.06) | 0.16*(0.02) | 0.54*(0.10) | 0.10*(0.02) | | |
| FL. NP | 0.03(0.06) | 0.17*(0.03) | 0.49*(0.12) | | 0.11*(0.02) | 0.08*(0.02) |

Note: standard errors are between brackets;

* indicates significance at the 1% level.

The myopic model, as well as the other models, estimates θ to be around 0.16. To give an idea about its interpretation, the predicted chance that a player with an α -value of zero contributes zero is estimated to be about 30%. The chance that a player contributes one as well as the chance that a player takes one is estimated to be about 20%, the chance of contributing and taking 2 is about 10%, while all the other contributions together take up the remaining 10% probability. For comparison, if θ would be 0 all choices have a probability of $\frac{1}{15}$ ($< 7\%$), while if θ goes to infinity the probability of a player choosing zero goes to 1. δ_1 is estimated to be close to $\frac{1}{2}$, so if the other contributes zero the valuation of the payoff of the other is halved. δ_2 is estimated to be 0.1 which means that a contribution of 5 would lead to an α of around 0.5, if the initial α -value is zero.

If we allow for a dichotomy between positive and negative impulses (M.NP) we find similar θ and δ_1 values. However, δ_{2P} seems larger than δ_{2N} . The improvement in the likelihood is significant at the 10% level ($p \approx 0.09$) even if we take just the average improvement per subject as a single observation (and significant at the 1% level if all rounds from all

players are taken into account).⁸

We next consider the forward-looking model (FL), but neglect the potential difference between positive and negative impulses for the moment. We find that γ is insignificant. Only a modest improvement in the likelihood is obtained when compared with the improvement caused by an extra impulse parameter in the myopic model ($p > 0.50$ if we take the average improvement per individual, $p \approx 0.10$ if we take all contributions into account). The other parameters, θ , δ_1 and δ_2 are very similar to the values found for the myopic model.

The full model (FL.NP) shows results that can be seen as a combination of the results found for M.NP and FL. Positive impulses are again stronger in impact than negative impulses, with parameter values similar to the ones of the myopic model, while δ_1 is again around 0.5.

An interesting result is that δ_1 is estimated to be close to 0.5 in all model specifications, which is consistent with earlier findings for two-player public good games (Bault et al., 2017). The fact that this finding is replicated could indicate that in this type of environment people, at least, weigh their history about as much as new information. It is also noteworthy that at the group level $0.2 < \frac{\delta_{2P}}{1-\delta_1} < 0.25$ always holds, while $0.2 < \frac{\delta_{2N}}{1-\delta_1} < 0.25$ does not. This indicates that our fourth hypothesis (H4) might not be correct and suggests that (most, not all) people are able to form stable cooperative relationships, but do not sustain long destructive relationships.

Furthermore, it is worth noting that we find that positive impulses seem to have a stronger effect on the α parameter than negative impulses. This seems to contrast with the findings summarized by Baumeister et al. (2001). However, there are studies that find that the influence of a positive signal might weigh stronger than that of a negative signal, see for instance King-Casas et al. (2005) and Rand et al. (2009). Another aspect could be the

⁸It should be noted that the difference between δ_{2P} and δ_{2N} reported in table 3.2 is a conditional result. Not all participants in our experiment were exposed to negative contributions and the ones that were exposed to them are not an exogenously chosen or created group. The players that at any point in time are faced with negative contributions are often also the ones that made negative contributions themselves. In other words, this behavior shows up pairwise and pairs that make negative contributions are likely to have different characteristics than pairs that do not make such contributions.

earlier mentioned finding that people are not only more affected by negative experiences but that they are also more motivated to get out of a negative situation. This result suggests that we should reject our second hypothesis (H2).

Note, furthermore, that the difference between δ_{2P} and δ_{2N} could be influenced by the fact that the myopic model does not allow for any forward-looking behavior. If (some) players are in fact forward looking, this might be partially captured by the δ_2 -parameter(s). If this was the case it would lead to a bias in the δ_2 -parameter(s), where the effect on δ_{2N} would be negative while the effect on δ_{2P} would be positive. The reason is that if a player is forward looking he or she wants to contribute more than if a player is not (as it is assumed that the other will positively react). This positive effect on the contributions will lead to δ_{2P} being higher, while δ_{2N} will be estimated to be lower (so that the effect will be less negative). However, we see that also in the full model (FL.NP) this difference between positive and negative impulses still exists. This directly opposes our second hypothesis (H2).

Moreover, the parameter γ , which is to capture the forward-looking behavior, is never significantly different from zero at the 5% level.

3.4.3 Individual level results

From the individual level estimates we can see how many of our participants are able to maintain stable cooperative or destructive relationships. We start by evaluating the myopic model. We find that 104 out of 130 participants meet the conditions mentioned in proposition 4. This means that 80% of our subjects are able to build sufficient ties to sustain cooperation. Of the remaining 26 participants, 10 have a ties mechanism that is too strong to be efficient (they are not able to sustain an efficient cooperative relationship), while the other 16 have an insufficiently strong tie mechanism. We find that in total 72 (36 pairs) out of these 104 participants are in fact cooperating efficiently in the final 10 rounds (that is, in more than 6 out of the last 10 rounds both contribute equally and either 4 or 5 tokens). This is in line with proposition 4, stating that $0.2 < \frac{\delta_{2i}}{1-\delta_{1i}} < 0.25$ is

a necessary but not a sufficient condition for stable and efficient cooperation. It, however, contradicts our fourth hypothesis (H4).

When we allow for different parameters for positive and negative impulses – as we did on the group level – we find that for most participants the positive impulse parameter is larger than the negative one: for 37 out of 130 participants $\delta_{2P} > \delta_{2N}$, for 15 participants $\delta_{2N} > \delta_{2P}$, and for two participants $\delta_{2N} = \delta_{2P} = 0$. Two participants did not receive any impulses, one received only negative impulses, and the remaining 73 participants encountered just positive impulses.

Shifting our attention to the forward looking-model now, we first investigate how many individuals have an estimate of γ that is significantly different from zero. It turns out that 79 out of 130 participants are indeed forward looking (γ positive at the 5% level). This seems to contrast with our previous finding at the group level, which might suggest a large heterogeneity among subjects in their forward-looking behavior.

When we compare the results of the full model, however, the same pattern as found for the group level is observed. Now, only 47 out of the 130 participants show forward looking behavior⁹. Moreover, we still find that for 88 out of the 130 participants $0.2 < \frac{\delta_{2Pi}}{1-\delta_{1i}} < 0.25$, while most players do not seem to be able to sustain negative relationships, as for only three participants we find that $0.2 < \frac{\delta_{2Ni}}{1-\delta_{1i}} < 0.25$. There are also ten participants that build excessively strong negative ties (i.e., α values smaller than -1). These findings might explain why we observe some prolonged intervals of negative contributions and the existence of sour relationships, as illustrated in figure 2¹⁰. Looking at the evidence presented at the group as well as at the individual level we must reject H4.

⁹Not all participants encounter many negative impulses though, so this result is conditional on encountering enough negative impulses

¹⁰For more on this see Hoyer et al. (2014), where the occurrence of such relationships is further analyzed with different experimental designs.

3.4.4 Predictive performance out-of-sample and model comparison

Now that we have estimated the Ties model both at the group level and the individual level, we put it to a more difficult test. We investigate if our model is not only able to explain behavior after the fact, but also to predict behavior in independent future rounds. Moreover, we will compare its predictive performance with the performance of three other models: the inequity-aversion model by Fehr and Schmidt (1999), the reinforcement learning model of Roth and Erev (1995), and a model with a fixed weight attached to the payoff of the other player (i.e., a fixed α).

In order to get true out-of-sample predictions, we use the following procedure. We first estimate the myopic model at the group-level, allowing for different positive and negative impulse parameters. We choose the myopic model because the other models we compare our model with do not allow for forward-looking behavior either. Moreover, this does not affect the performance of our model too much as the additional parameter capturing this effect is insignificant. First we estimate the model on the group-level, then we take the contributions of the new other in the second game to calculate the α -values (for α_0 the values from the SVO test are used again), using the estimated parameters of the first game. The predicted action is the choice that generates the highest likelihood according to (3.8). Note that we do not re-estimate the model after every round and also do not readjust α on the basis of choices made by the participants themselves in the second game. This allows predictions to run away from the realized values. The fact that contributions in the second game were generally higher in the second game should make forecasting harder.

We use this procedure not only for the ties model but also for the other models of social preferences and the basic reinforcement learning model referred to above. To make forecasts for these models we use a similar procedure as described for the ties model.

For the fixed alpha model this means that we estimate the α parameter at the group-level on the behavioral data of the the first FPG game and then use this estimates this to predict the choices made in the second FPG game.

For the Fehr-Schmidt model we estimate, again at the group-level, the α and β parameters of the following expression for the expected utility of a particular choice k) V_{ikt}^e , using the behavioral data of the first FPG game and (3.8):

$$V_{ikt}^e = X_{ikt}^e - \alpha(X_{jt}^e - X_{it}^e)_{X_{it} < X_{jt}^e} - \beta(X_{it}^e - X_{jt}^e)_{X_{it} > X_{jt}^e} \quad (3.9)$$

Where X_{ht}^e ($h = i, j$) denotes the expected payoffs calculated using either the expected contribution of the other in the same round (participants were asked for this after every choice made by themselves) or by the actual contribution of the other in the previous round. When estimating this model, we find β to be larger than α for both specifications of X_{ht}^e ($h = i, j$), which is contrary to the predictions and findings by Fehr and Schmidt, but more often found in the literature (Yang et al., 2016). What is more problematic, though, is that both α and β are estimated to be larger than 1, again violating the assumptions of the model. Note that $\beta > 1$ implies that one would prefer to throw away a dollar to diminish inequality with one dollar. Because of the clear lack of support for this model we will not further consider it below.

In the Roth and Erev model of reinforcement learning players learn the value of certain actions by playing them. The higher the payoff after playing a certain action the more this action gets reinforced, meaning that the probability of choosing this (or a similar) action increases. In the three parameters version of the model used here: s denotes the strength (or speed) of learning, indicating how much the chosen action is reinforced, ϕ stands for a decay effect that captures the speed by which the attraction of an action diminishes over time, and E denotes an experimentation effect that represents the reinforcement of adjacent choices. To get to estimates we use a similar procedure as explained in Erev and Roth (1998), meaning that all probabilities of an individual i choosing an action k at time t (π_{ikt}) are initially the same, as the attraction of each choice q_{ikt} is assumed to be the same q at the beginning of the game. After a choice is made (zero is chosen as the first prediction in this exercise) the distance (R) between the realized payoff and the minimal payoff, combined with the effect of the parameters, determine the new attraction

of choices and thereby the choice probability distribution in the round thereafter. More precisely:

$$\begin{aligned}
\pi_{ikt} &= q_{ikt} / \sum_h q_{iht} \\
q_{ik1} &= q_{ih1} = q \\
q_{ikt+1} &= (1 - \phi_i)q_{ikt} + E_{ikt} \\
E_{ikt} &= s_i R_{jt}(1 - \epsilon_i) \text{ if } k = j \\
E_{ikt} &= s_i R_{jt}(\epsilon_i/2) \text{ if } k = j \pm 1 \\
E_{ikt} &= 0 \text{ otherwise}
\end{aligned} \tag{3.10}$$

Another interesting model for explaining behavior in dynamic settings was introduced by Camerer and Ho (1999). Their Experience-Weighted Attraction (EWA) learning model not only allows for learning via payoffs, but also that players may learn over time what other players are likely to do. Although this kind of belief learning is undoubtedly important in many economic settings, it should not affect behavior in our game, as in our game the net return on a contribution vis-à-vis that of another contribution of a player does not depend on the contribution of the other player but only on his or her own contribution. Finally, we mention the 'types' model of Levine (1998), a social preference model that may appear similar behaviorally, but is conceptually quite different from the ties model. In this model the weight an individual i attaches to the utility (u_j) of another individual j is dependent on one's own (constant) altruism parameter (α_i), the belief about the altruism parameter of the other (α_j) and a parameter λ that weighs both, such that i 's utility (u_i) gets transformed into an extended utility, v_i :

$$v_i = u_i + \frac{\alpha_i + \lambda\alpha_j}{1 + \lambda} u_j \tag{3.11}$$

Although this model assumes unexplained fixed altruistic parameters, there is some similarity with the ties model. If the contributions of the other player are seen as signals of that player's altruism level, these signals would then change the belief of the other's

altruism level and thereby the weight one attaches to his or her utility (α). Because the model does not specify the belief updating process, let alone how to apply it in our setting, we do not further consider it here.

Our discussion of social preferences models and the learning model of Roth and Erev, reflects that only a few social preference models available are able to make predictions for the dynamic behavior in our games. This is not surprising, as most theoretical models are not designed to explain dynamics.

For the predictions regarding the second game, we again use the choices with the highest likelihood of being chosen, given the parameters estimated on the behavioral data of the first game. Table 3.3 presents the mean absolute error and the mean squared error of the predictions:

Table 3.3: Out of sample prediction with group-level estimates

| Model | Mean Absolute Error | Mean Squared Error |
|---------------|---------------------|--------------------|
| Ties Model | 0.51 (0.53) | 1.65 (2.39) |
| Fixed Alpha | 1.92 (2.41) | 6.97(14.05) |
| Roth and Erev | 1.64 (1.07) | 4.39 (4.59) |

Note: standard errors using average errors per individual between brackets

From the results in table 3.3 we can conclude that the Ties model seems to perform best, supporting our fourth hypothesis. This is confirmed by Wilcoxon signed rank tests with the average error (for both squared and absolute errors) per individual as observations. These tests show that this model outperforms the other models when it comes to predicting ($p < 0.01$ for all tests). Furthermore it is interesting to note that especially the reinforcement learning model by Roth and Erev does not do a good job when it comes to predicting behavior out of sample as it's mean absolute prediction error is more than three times higher than that of the Ties model.

As an alternative test we check how individual-level estimations perform. We use the same procedure as with the group-level estimates, but now each player's predicted choice is calculated using individual estimates. Table 3.4 shows the results.

Table 3.4: Out-of-sample prediction with individual estimates

| Model | Mean Absolute Error | Mean Squared Error |
|---------------|---------------------|--------------------|
| Ties Model | 1.00 (1.11) | 3.58 (5.61) |
| Fixed Alpha | 1.69 (1.72) | 6.11(12.68) |
| Roth and Erev | 3.01 (1.47) | 11.63 (7.21) |

Note: standard errors using average errors per individual between brackets

Again it turns out that the Ties model performs significantly better than the learning model and the fixed alpha model, supporting our fourth and final hypothesis. Note, though, that with this specification both the Ties and especially the reinforcement learning model, perform worse than when group-level estimates are used. This may seem surprising, but is caused by the fact that some individuals experience very little variation in impulses in the first FPG game, making it difficult to estimate their individual parameters precisely.

3.5 Applying the Ties Model to the Repeated Prisoner’s Dilemma

The Ties model gives an explanation for the development of cooperation or antagonism that is quite different from the rest of the literature, as it focuses on changes in social preferences generated by interaction experiences rather than on given (fixed) social preferences or simple heuristics represented by automata. In this section we will explore a connection to another strand of research focusing on the evolution of behavior in repeated games, specifically a series of studies by Dal Bó and Fréchet (2011) and Fudenberg et al. (2012). To understand the strategies people use when placed in environments that are either well- or ill-suited to generate cooperation, they have subjects play multiple repeated Prisoner’s Dilemma (PD) games with continuation probabilities between $1/2$ and $7/8$. Using maximum likelihood estimation procedures, they estimate the share of a series of simple strategies, or automata, such as tit-for-tat (TFT), always defect (AD), and tit-

for-two-tats (TF2T). Fudenberg et al. (2012) find that people become 'slower to anger' and 'faster to forgive' – i.e. are more willing to allow a defection and pick up cooperation after only a few cooperative choices of the counterpart –, if cooperation becomes more profitable. This is reflected in the presence of a strategy like TF2T in these environments. This section serves to illustrate how different parameter combinations of the Ties model can generate (or mimic) these strategies as well as a the shift towards more forgiving strategies as cooperation becomes more attractive. In appendix C these arguments are worked out in more mathematical detail.

We start our Ties model-based analysis of the PD game by introducing other-regarding preferences. We do this by adding the α -weighted payoff of the other to a player's payoff. Starting from a general representation of a PD game without any other-regarding payoffs (see table 3.5a), we apply these other-regarding preferences to two specific games with benefit/cost (b/c) ratios of 2 (table 3.5b) and 4 (table 3.5c). These examples are chosen for comparability with the games found in Fudenberg et al. (2012). What stands out from these new payoff matrices is that defecting is now no longer necessarily the dominant action. If α is larger than 1/2 (in table 3.5b) or 1/4 (in table 3.5c), cooperation becomes dominant. If we now define the impulse generated by a cooperative choice to be of size one and the impulse from defection by the other to be of size zero (as this is the Nash equilibrium action of the stage game), we can apply a similar model as the one we introduced for the public good game (see below).

Table 3.5: Prisoner's Dilemma (with other regarding preferences)

| | (a) b/c | (b) b/c=2 | (c) b/c=4 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|--|---------------|-----------|---|---|-----|----|---|---|---|--|--|---|---|---|--------------|---------------|---|-------------|---|--|--|---|---|---|--------------|---------------|---|-------------|---|
| | <table style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 50%;"></th> <th style="width: 50%; text-align: center;">C</th> <th style="width: 50%; text-align: center;">D</th> </tr> <tr> <th style="width: 50%; text-align: center;">C</th> <td style="padding: 5px;">b-c</td> <td style="padding: 5px;">-c</td> </tr> <tr> <th style="width: 50%; text-align: center;">D</th> <td style="padding: 5px;">b</td> <td style="padding: 5px;">0</td> </tr> </table> | | C | D | C | b-c | -c | D | b | 0 | <table style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 50%;"></th> <th style="width: 50%; text-align: center;">C</th> <th style="width: 50%; text-align: center;">D</th> </tr> <tr> <th style="width: 50%; text-align: center;">C</th> <td style="padding: 5px;">1+1α</td> <td style="padding: 5px;">-1+2α</td> </tr> <tr> <th style="width: 50%; text-align: center;">D</th> <td style="padding: 5px;">2-α</td> <td style="padding: 5px;">0</td> </tr> </table> | | C | D | C | 1+1 α | -1+2 α | D | 2- α | 0 | <table style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 50%;"></th> <th style="width: 50%; text-align: center;">C</th> <th style="width: 50%; text-align: center;">D</th> </tr> <tr> <th style="width: 50%; text-align: center;">C</th> <td style="padding: 5px;">3+3α</td> <td style="padding: 5px;">-1+4α</td> </tr> <tr> <th style="width: 50%; text-align: center;">D</th> <td style="padding: 5px;">4-α</td> <td style="padding: 5px;">0</td> </tr> </table> | | C | D | C | 3+3 α | -1+4 α | D | 4- α | 0 |
| | C | D | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C | b-c | -c | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D | b | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | C | D | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C | 1+1 α | -1+2 α | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D | 2- α | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | C | D | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C | 3+3 α | -1+4 α | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D | 4- α | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Note: Table a gives the actual payoffs of player 1, while b and c give the valuation of these payoffs by a player that also cares about the other player. C stands for cooperation, and D for defection.

Both Dal Bó and Fréchette (2011) and Fudenberg et al. (2012) find experimental evidence

that many subjects in their experiments use either a tit-for-tat (TFT)¹¹ or a tit-for-two-tats (TF2T)¹² strategy although these strategies are often not evolutionary stable (in an evolutionary game theory context).¹³ Below we will show that the simple and neurobiologically underpinned Ties model (Bault et al. (2014)) can help explain the behavior observed in these experiments.

The previously mentioned studies combine the simplicity and descriptive power of strategies like tit-for-tat and grim trigger with sophisticated estimation procedures that illustrate the popularity of these strategies among experimental subjects. They do not, however, explain why and when exactly players switch to different strategies when the cost/benefit ratio in a PD environment changes. Using the Ties model we can fill this gap and predict different behavior for different b/c ratios. It also allows us to test if the behavior of subjects is consistent within a game. While the previously mentioned studies do not attempt to explain why subjects use different strategies within the same game, our method does not attach a single strategy to an individual or even to an individual in a particular interaction. Another advantage is that applying the Ties model allows us to see if the tie mechanism and resulting strategies are consistent across different social dilemma environments and games and could thus help formulate models that are also relevant outside the specific laboratory environments used for investigation.

An example of a strategy that is easily generated by the tie mechanism is the wellknown TFT strategy. According to this strategy, a player starts with cooperating (choosing C) and, subsequently, simply chooses whatever his opponent did in the previous round. Thus, after observing C (D) the player chooses C (D). If play starts with D first instead of C , the strategy is labeled DTFT. For the tie mechanism to generate such behavior, the

¹¹Dal Bó and Fréchette estimate that for their games with a continuation probability of $3/4$ between 35% (if $b/c \approx 2$) and 56% (if $b/c \approx 4$) of subjects choose TFT. Fudenberg et al., who consider many more strategies, find between 20%, for $b/c=1.5$, and 7%, for $b/c=4$, of players choosing TFT.

¹²Dal Bó and Fréchette do not consider TF2T. Fudenberg et al. find that around 12% of their players choose TF2T.

¹³TFT requires a player to start with choosing C and, subsequently, to choose whatever his opponent did in the previous round. Thus, after observing C (D) the player chooses C (D). If a player starts with D first instead of C , the strategy is labeled DTFT. TF2T requires a player to always choose C , unless his counterpart chose D in the previous two periods. If a player starts with D first instead of C , the strategy is labeled DTF2T.

following is required: First, a player should have a strong enough impulse parameter δ_2 (the exact size depends on the b/c ratio). Secondly, the memory of this player must not be too strong, as otherwise a strong tie can be built up that sustains deviations by the other player. Hence, the tie persistence parameter δ_1 must be sufficiently small. Finally α_0 determines whether play starts with C (for TFT) or D (for DTFT).

If we allow players to start with $\alpha_0 \neq 0$ we find in appendix C that, for certain parameter values of δ_1 and δ_2 , play starts to mimic often reported strategies like AD, TFT, and TF2T. However, we can also use the parameter value estimates found here and see to what strategies these parameter estimates correspond, a task to which we turn next. Besides the before-mentioned strategies we also investigate a modified strategie: 'qualified' tit-for-two-tats (QTF2T). This strategy is similar to TF2T in all but one aspect. It requires more than one cooperative choice by the other before they are willing to 'forgive' a defecting (D) decision. In terms of the Ties model this means that first the value of α has to be build up, in this case described in appendix C, section 3.9.3, until the theoretical maximum ($\frac{\delta_2}{1-\delta_1}$), but any value corresponding to any number of consecutive cooperative decisions can be chosen, before someone is forgiving. The intuition behind these 'qualifications' is that players with using TF2T are vulnerable to exploitation. Other players could exploit them by alternating between C and D. Since it seems unlikely that players would accept such exploitation we require the other player to show good intentions for a longer period, before these strategies become 'forgiving'. In appendix C the case for $b/c = 4$ is worked out.

For a sensible comparison between the parameter values found in this study and those relevant for a PD environment we need to normalize the impulse, I . This is because if we multiply the impulse by a factor i , the estimate of δ_2 will change by factor $1/i$. Therefore, we normalize by assuming a cooperative action in a PD game to be equivalent with a fully Pareto efficient action ($C=4$) in our FPG game and defining the impulse in that event to be equal to $I_n \equiv \frac{C_j - C_j^{ref}}{C^{eff} - C_j^{ref}}$. Thus, in order to translate the values we found for δ_2 to values suitable for a PD game environment, where choosing C (the efficient choice) is valued as 1, we multiply δ_2 by 4. As before, α_0 will stand for the starting value of α .

Figures 3.3 and 3.4, the first for $b/c = 2$ and the second for $b/c = 4$, which show which parameter values of the tie mechanism (δ_1 and δ_2) correspond to which strategies¹⁴. The lines mark the conditions for which the Ties model predicts the behavior of the strategies mentioned earlier. The crosses in the graph represent the different individuals in our experimental study, using normalized δ_{2P} (as only positive impulses are possible in this environment) and δ_1 values, as estimated with the myopic model, that allows for a dichotomy between positive and negative impulses.

¹⁴for characteristics of these strategies and a more elaborate analysis, see appendix C.

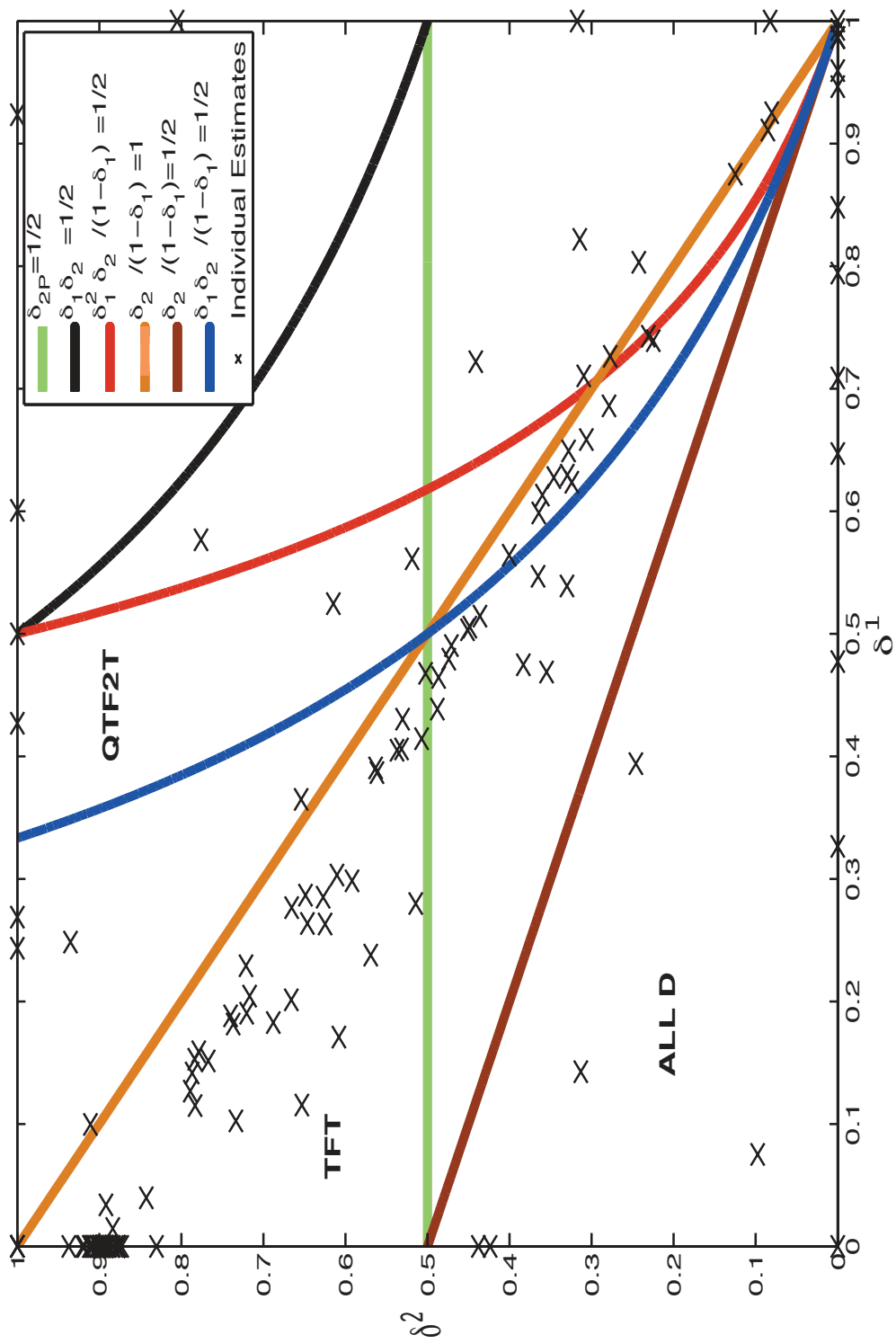


Figure 3.3: Parameter estimates and strategies for $b/c=2$

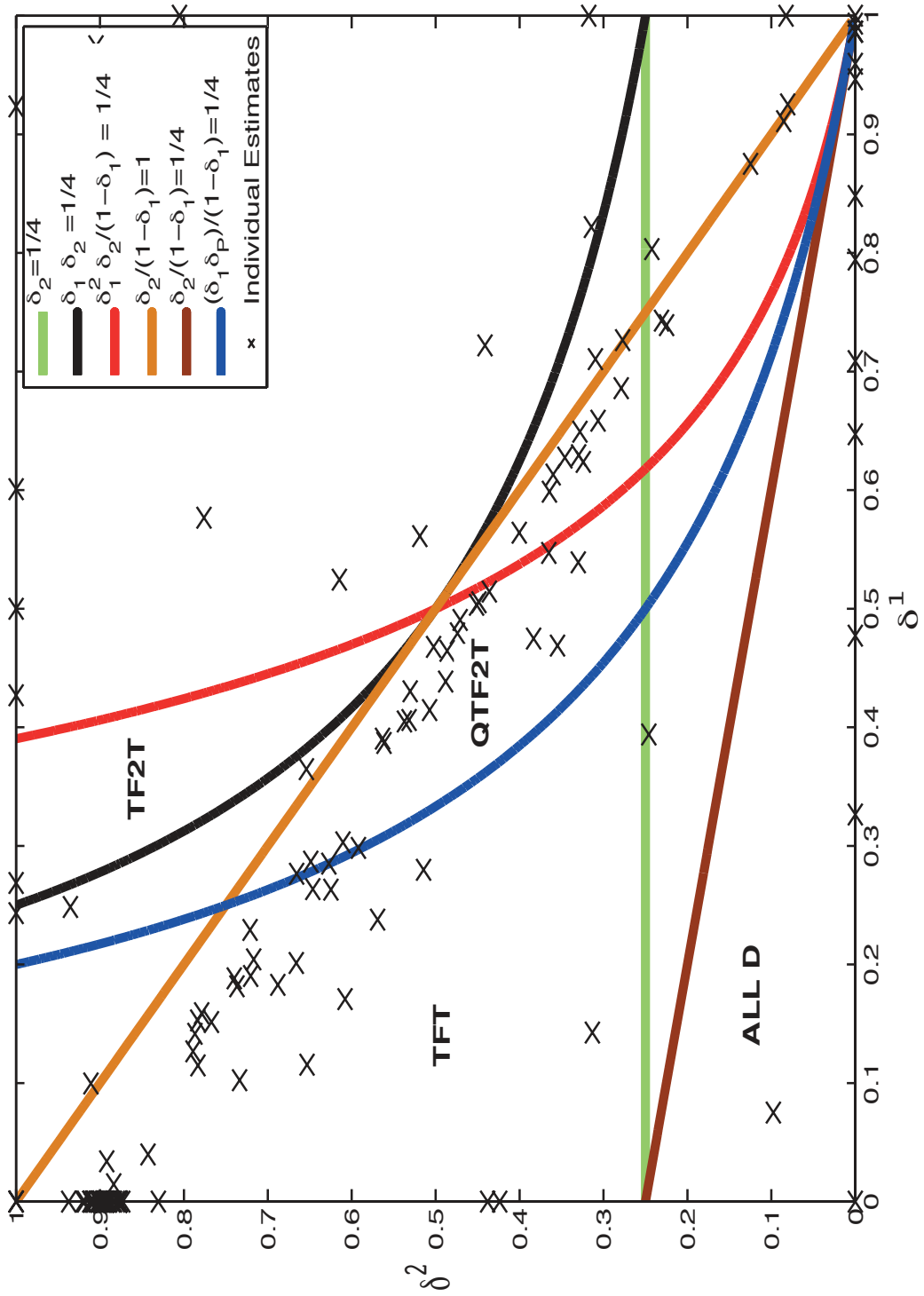


Figure 3.4: Parameter estimates and strategies for $b/c=4$

From these figures it becomes clear that for a lower b/c ratio the same players 'switch' from cooperative strategies to strategies which imply more defection. It also shows that in the $b/c = 4$ setting, (D)TFT, TF2T and related strategies are commonly found, as in Fudenberg et al. (2012). There is however also a noticeable difference, the lack of players playing the AD strategy. A potential reason for this might be the possibility to destroy the public good, which could lead players to be reluctant to play the selfish choice (as they might in our game be afraid for punishment after contributing zero). This is highlighted by the fact that if we use the estimates from the myopic model, which does not allow for a difference in impulse impact, there are some more AD players, as δ_2 is typically estimated to be lower in this case. Observations that are to the right of the top-left bottom-right diagonal represent players with a very strong tie mechanism as continuous cooperation by the other player would lead to an α -value greater than one.

Finally, it is interesting to note that the Ties model could also explain the repeated PD finding of Breitmoser (2015) that if one player chooses C and the other D, both show an about equal probability of playing C in the next round. This is presented as evidence against the existence of TFT. If one thinks in terms of a tie mechanism this finding may not be so surprising. After all, if a player played C in the previous round his or her α value must have been relatively high, while if a player played D this value must have been relatively low. Now, because the tie (α) of the former player will decay (as it gets multiplied by δ_1) the chance that this player chooses C declines. In contrast, the tie of the other player will be reinforced (with δ_2) by counterpart's cooperative action in the previous round. Consequently, the α values will move towards each other. In short, as one tie is initially relatively strong (reflecting a higher probability of playing C) and becomes weaker, while the other tie is relatively weak (reflecting a higher probability of playing D) and becomes stronger, the chances to play C for both players converges, making the finding of Breitmoser well explicable by a tie mechanism.

3.6 Conclusion

We conclude with a summary of our main findings. First of all it turns out that the results are very much in line with earlier studies using the Ties model. More specifically, we also observe that the memory component of the tie mechanism, represented by the tie persistence parameter δ_1 , is about equally important as the impulse component(s), represented by the parameter δ_2 , with the scale free δ_1 being estimated to be close to 0.50.

In contrast to our original hypothesis we find that positive impulses have a stronger impact than negative ones. Apparently, the bad is not stronger than the good (Baumeister et al., 2001) in this context. This asymmetry in the tie mechanism is helpful in getting cooperation going, while not rendering cooperators defenseless against people that are just trying to benefit from them. We also find that players do not seem to be very much forward-looking.

Our out-of-sample predictions show that the Ties model significantly outperforms both a reinforcement learning model as well as a model with constant social preferences. For both the learning as well as the Ties model we find that the predictive power improves if we use group-level instead of individual-level estimates. This appears due to the lack of behavioral variability for some of our subjects.

The Ties model also generated insights for a (repeated) Prisoner's Dilemma (PD) game context. Strategies observed in experiments can be understood with the help of the Ties model. Moreover, using the estimated parameters from our public good game, the model helps explain why people switch to different strategies when faced with a different cost-benefit ratio. Our alternative explanation for the behavior in repeated PD games does require people to switch strategies in a seemingly ad hoc way.

3.7 Appendix A: Proof of Propositions

3.7.1 Proposition 1: Contributions outside of $-5 \leq C_{it} \leq 5$ can never be part of any equilibrium if $-1 \leq \alpha \leq 1$.

We use the fact that our agents can only change their decision in discrete steps. Subtracting $V(C_{it})$ from $V(C_{it} + 1)$ we get:

$$V(C_{it}) - V(C_{it} + 1) = 2C_{it} + 2 - 10\alpha + \gamma(-10 + \alpha(2\gamma C_{it} + 2(1 - \gamma)C_{jt}^e + \gamma + 1)) \quad (3.12)$$

Where γ is between 0 and 1. For the proposition to be true this equation must be positive.

We reformulate this condition to:

$$2C_{it} + 2 + 2\alpha\gamma(\gamma C_{it} + (1 - \gamma)C_{jt}^e + \frac{1}{2}\gamma + \frac{1}{2}) > 10(\alpha + \gamma) \quad (3.13)$$

We begin by only looking at equilibria with symmetric contributions ($C_{it} = C_{jt}$).

$$2C_{it} + 2 + 2\alpha\gamma(C_{it} + \frac{1}{2}\gamma + \frac{1}{2}) > 10(\alpha + \gamma) \quad (3.14)$$

So at $C_{it} = 5$ we have:

$$\begin{aligned} 12 + 2\alpha\gamma(5 + \frac{1}{2}\gamma + \frac{1}{2}) &> 10(\alpha + \gamma) \Rightarrow \\ 12 + \alpha(\gamma(11 + \gamma) - 10) - 10\gamma &> 0 \end{aligned} \quad (3.15)$$

This last statement is always true for $-1 < \alpha < 1$, since if α is one we have:

$$12 + 11\gamma + \gamma^2 > 10(\gamma + 1) \quad (3.16)$$

Which is always the case. If α is -1 we have:

$$12 - 11\gamma - \gamma^2 > 10(\gamma - 1) \quad (3.17)$$

Now since (3.15) is a monotone function in α these results hold for the entire interval.

We can use the same method to show that $V(C_{it}) > V(C_{it}-1)$ always holds when $C_{it} \leq -5$.

At $C_{it} = -5$ the equivalent of (15) is:

$$-12 + 2\alpha\gamma\left(-6\frac{1}{2} + \frac{1}{2}\gamma\right) > 10(\alpha + \gamma) \quad (3.18)$$

This is never true for positive α 's. For α is -1 we get:

$$-12 + \gamma(13 - \gamma) > 10(\gamma - 1) \quad (3.19)$$

This cannot be for true for γ between zero and one either.

For the asymmetric equilibria we have to go back to (3.13). The left side is increasing in C_{jt}^e for $\alpha > 0$ and decreasing when $\alpha < 0$. To see if there are instances where contributing 6 is preferred to contributing less we therefore only need to check for $C_{jt}^e = 1$. This gives:

$$12 + 2\alpha\gamma\left(5\gamma + (1 - \gamma) + \frac{1}{2}\gamma + \frac{1}{2}\right) > 10(\alpha + \gamma) \quad (3.20)$$

Again this is always true for $0 \leq \alpha \leq 1$ (and if the other contributes positively, α cannot be negative in an equilibrium). A similar procedure can be used to show that no choice more negative than -5 can be part of an equilibrium.

3.7.2 Proposition 2.2: All symmetric equilibria with $-5 \leq C_{it} \leq 5$ are possible.

For a stable situation we need a value for α such that $V(C_{it} - 1) < V(C_{it}) > V(C_{it} + 1)$ holds and we need that after (infinitely) repeated play of C_{it} this still holds. We start by investigating the case in which here agents are not forward looking (γ is zero).

Myopic Agents

We first look at the case in which both contributions are positive. In this situation equation (3.14) simplifies to:

$$2C_{it} + 2 > 10\alpha \quad (3.21)$$

From (3.21) we with every increase of α by 0.2 the contribution that gives the highest value shifts one up. For an equilibrium to be sustainable we need the α -value to be stable (in a steady state) for a the given contribution. So we use (3.5), and look for:

$$\alpha = \delta_1\alpha + \delta_2I \quad (3.22)$$

For I we use the Nash equilibrium as a reference point as we did throughout the chapter. This leads to:

$$\begin{aligned} \alpha &= \delta_1\alpha + \delta_2C \\ &\text{or} \\ \alpha &= \frac{\delta_2C}{1 - \delta_1} \end{aligned} \quad (3.23)$$

From (3.21) we know that:

$$0.2C < \alpha < 0.2(C + 1) \quad (3.24)$$

Combining (3.23) and (3.24) we obtain:

$$0.2 < \frac{\delta_2}{1 - \delta_1} < 0.2 + (0.2/C) \quad (3.25)$$

If we look at the same situation ($\gamma = 0$) for negative values (more precisely for $C \leq -2$, we will discuss the situations in which C is 0 or -1 later) we change (3.21) into:

$$2C > 10\alpha \quad (3.26)$$

So also in the equation above we see that the best response changes with every increase (or drop) in α of 0.2. Following an analogous procedure to the one we used for a positive C we obtain the following condition:

$$0.2 > \frac{\delta_2}{1 - \delta_1} > 0.2 - \frac{0.2}{C} \quad (3.27)$$

If $C = 0$, then the stimulus is zero. This will lead to the value of α moving gradually towards zero as well. As the best response to an α -value of zero is to play 0 we have that the $[0, 0]$ equilibrium can always exist regardless of the δ -parameters. To the entire range of α -values for which a contribution of zero is a best response we use (3.26) and observe that as long as $\alpha > -0.2$ the value of playing zero is bigger than the value of -1. This gives us the lower bound $\alpha = -0.2$. Now for the higher bound we have to see when playing 1 is more attractive than playing 0. From (3.21) we find that this boundary is 0.2.

Forward Looking Agents

If γ is unequal to zero all values of C_{it} are still part of symmetric equilibria, but the condition on δ_1 and δ_2 becomes stricter. Just as in the previous case we start from (3.14) and fill in (3.23):

$$\begin{aligned} 2C + 2 + 2\frac{\delta_2 C}{1 - \delta_1}\gamma(C + \frac{1}{2}\gamma + \frac{1}{2}) &> 10(\frac{\delta_2 C}{1 - \delta_1} + \gamma) \\ 2C + 2 > (10 - 2\gamma(C + \frac{1}{2}\gamma + \frac{1}{2}))\frac{\delta_2 C}{1 - \delta_1} + 10\gamma \end{aligned} \quad (3.28)$$

$$\frac{2C + 1 - 10\gamma}{(10 - 2\gamma(C + \frac{1}{2}\gamma + \frac{1}{2}))C} > \frac{\delta_2}{1 - \delta_1} \quad (3.29)$$

If $10 - 2\gamma(C + \frac{1}{2}\gamma + \frac{1}{2}) < 0$ the inequality changes direction.

We also fill in the lower bound we obtain:

$$\frac{2C - 10\gamma}{(10 - 2\gamma(C - 1 + \frac{1}{2}\gamma + \frac{1}{2}))C} < \frac{\delta_2}{1 - \delta_1} < \frac{2C + 2 - 10\gamma}{(10 - 2\gamma(C + \frac{1}{2}\gamma + \frac{1}{2}))C} > \frac{\delta_2}{1 - \delta_1} \quad (3.30)$$

We can repeat this procedure in the negative domain and get the following condition:

$$\frac{2C - 2 - 10\gamma}{(10 - 2\gamma(C - 1 + \frac{1}{2}\gamma - \frac{1}{2}))C} > \frac{\delta_2}{1 - \delta_1} > \frac{2C - 10\gamma}{(10 - 2\gamma(C + \frac{1}{2}\gamma - \frac{1}{2}))C} \quad (3.31)$$

3.7.3 Proposition 3: Asymmetric equilibria exist if $C_i C_j > 0$ and

$$|C_i| \leq 5 \text{ and } |C_j| \leq 5$$

For simplicity we restrict ourselves to myopic agents. This changes (3.23) into:

$$\alpha = \frac{\delta_2 C_j}{1 - \delta_1} \quad (3.32)$$

And (3.24) changes into:

$$0.2C_i < \alpha < 0.2(C_i + 1) \quad (3.33)$$

Leading to:

$$0.2 \frac{C_i}{C_j} < \frac{\delta_{2i}}{1 - \delta_{1i}} < 0.2 \frac{C_i}{C_j} + \frac{0.2}{C_j} \quad (3.34)$$

In order for this situation to be an equilibrium we also need:

$$0.2 \frac{C_j}{C_i} < \frac{\delta_{2j}}{1 - \delta_{1j}} < 0.2 \frac{C_j}{C_i} + \frac{0.2}{C_i} \quad (3.35)$$

Looking at the extreme case of a [1,5] equilibrium this implies:

$$0.04 < \frac{\delta_{2i}}{1 - \delta_{1i}} < 0.08 \quad (3.36)$$

And:

$$1 < \frac{\delta_{2j}}{1 - \delta_{1j}} < 2 \quad (3.37)$$

While such an equilibrium is mathematically possible, for it to be maintained the two players have to be quite different.

There are no equilibria where one player contributes a negative amount while the other contributes a positive amount. Constant negative contributions eventually create a negative α in the other and contributing positively can not be an optimal choice under a negative α .

3.7.4 Proposition 4: For the socially optimal choices to be a stable equilibrium under the myopic model, both players satisfying $0.2 < \frac{\delta_{2i}}{1 - \delta_{1i}} < 0.25$ is a necessary, but not a sufficient condition.

This result is directly visible in (3.25) if we plug in 4 as the contribution level. (3.25) also shows that if a player has the characteristics to be in a socially optimal equilibrium he or she is also willing to conform with any other (non negative) symmetric equilibrium with lower contributions.

3.8 Appendix B: Reference Point

In this part we will evaluate the model fit for different reference points in our model. In this section we restrict ourselves to the myopic version of the model, allowing for different positive and negative impulse parameters (as this was the best predicting model). We have for the size of the impulse:

$$I_{ijt} = C_{jt} - C_i^{ref} \quad (3.38)$$

The definition of the reference contribution C^{ref} is not trivial. Several points are however

appealing from a theoretical standpoint. The first candidate that we consider is the point that we chose in the main part of the chapter, the contribution that an agent chooses in a one-shot Nash equilibrium ($C^{ref} = 0$). Another static option is the use of the Pareto optimal contribution as a reference point ($C^{ref} = 4$). It is also possible that the reference point is not static and depends on either an agent’s own behavior or the previous behavior of the other. We test the predictive performance for two such reference points: $C_t^{ref} = C_{it}$ if an agent’s own contribution is used and ($C_t^{ref} = C_{jt-1}$) if one looks at the contribution of the other in the previous round. In the last two cases we initialize the system using $C_1^{ref} = 0$. In table 3.6 the parameters from estimating the model using different reference points are shown.

Table 3.6: Estimates for different reference points

| Reference point | δ_1 | δ_{2P} | δ_{2N} | $\sum LL$ |
|--------------------|------------|---------------|---------------|-----------|
| Nash | 0.490 | 0.115 | 0.080 | -8545 |
| Pareto | 0.975 | 0.108 | 0.000 | -11265 |
| Own Contribution | 1 | 0.003 | -0.006 | -11598 |
| Contribution other | 0.989 | 0.179 | 0.160 | -10826 |

From the table we see that using the Nash solution as a reference point produces to the highest likelihood. It is also interesting to note that with dynamic reference points δ_1 is estimated to be very close to 1. A reason for this might be that if two players are in a positive symmetric equilibrium (where $C_{it}=C_{jt}$ for multiple rounds) the value of α goes to zero (or might even go negative in case of the pareto optimum being the reference point) as all the impulses are zero. We observe these equilibria quite regularly in our dataset. The only way for the models with these particular reference point specifications to keep α high, which is necessary for positive contributions to occur, is for δ_1 to approach zero. A side effect of this is that with $\delta_1 \approx 1$ players basically have an infinite memory and early impulses have the same effect as new ones. Judging on the basis of the likelihood, though, this is not the case.

Also the Pareto optimum does not perform well. This can be due to the fact that using

this reference point even positive contributions might lead to a negative α and thus to the strange situation that if $\delta_{2N} > 0$ positive contributions by one player would lead to (expected) negative contributions by the other. This would lead to a negative spiral, that we hardly ever observe in the data.

We thus conclude that, if we use (3.6a), (3.6b) and (3.38) to model the tie mechanism, then $C^{ref} = 0$ is the best rule to use for the reference point.

It is interesting to see that if we focus on $\Delta\alpha = (\alpha_t - \alpha_{t-1})$ in a positive and symmetric equilibrium, we get into a situation where the change in α is basically determined by the change in contributions since the initial value of α diminishes over time. If we start from the basic tie mechanism described in (3.5) we have (with $I = C_{jt}$):

$$\begin{aligned} \alpha_{ij2} &= \alpha_{ij1}\delta_{1i} + \delta_{2i}C_{j1} \Rightarrow \alpha_{ij3} = \alpha_{ij1}\delta_{1i}^2 + \delta_{1i}\delta_{2i}C_{j1} + \delta_{2i}C_{j2} \Rightarrow \\ \alpha_{ijt} &= \alpha_{ij1}\delta_{1i}^{t-1} + \delta_{1i}^{t-2}\delta_{2i}C_{j1} + \dots + \delta_{1i}\delta_{2i}C_{jt-2} + \delta_{2i}C_{jt-1} \Rightarrow \\ \alpha_{ijt} &\approx \frac{\delta_{2i}C}{1 - \delta_{1i}} \quad (\text{for } t \rightarrow \infty) \end{aligned} \tag{3.39}$$

Consequently, for $\Delta\alpha$ it holds in this situation:

$$\begin{aligned} \Delta\alpha_{it+1} &= \alpha_{it+1} - \alpha_{it} = \delta_{1i}\alpha_{it} + \delta_{2i}C_{jt} - \alpha_{it} \Rightarrow \\ \Delta\alpha_{it+1} &= \delta_{2i}C_{jt} - (1 - \delta_{1i})\alpha_{it} \approx \delta_{2i}C_{jt} - (1 - \delta_{1i})\frac{\delta_{2i}C^{eq}}{1 - \delta_{1i}} = \delta_{2i}(C_{jt} - C^{eq}) \end{aligned} \tag{3.40}$$

In the symmetric equilibria (the most commonly observed equilibria in our experiment) we have that, when the reference point with respect to α is fixed, the change in α approximates a linear function of the change in the other player's contribution.

3.9 Appendix C: Tie Model Parameters for different Repeated PD Strategies

3.9.1 Introduction

In this section we look at a number of strategies for the repeated prisoner's dilemma game and determine which combinations of parameters in the tie model are compatible with those strategies. This analysis provides the basis for the hypothetical distribution of strategies presented in section 5, which we derived based on our parameter estimates from the fragile public good game. We take a prisoner's dilemma game of the following form, which corresponds to Fudenberg et al. (2012)'s $b/c=4$ case (only the required α -values will change in the analysis below if one uses other b/c ratios):

Table 3.7: Prisoner's Dilemma with $b/c=4$

| | | |
|---|------|-------|
| | C | D |
| C | 3,3 | -1, 4 |
| D | 4,-1 | 0,0 |

We start from the basic Ties model in which agents have the following extended utility function:

$$V_{it} = U_{it} + \alpha_{ijt}U_{jt} \quad (3.41)$$

Where V_{it} denotes the extended utility of player i at time t and U_{it} and U_{jt} stand for the direct own utility (payoff) of i and j , respectively, at time t , while α_{ijt} represents the weight i attaches to the utility of j (i 's tie with j) in period t , which is updated as follows:

$$\alpha_{ijt} = \delta_{1i}\alpha_{ijt-1} + \delta_{2i}I_{t-1} \quad (3.42)$$

With α_{ij1} denoting the initial tie. We define the impulse I_{t-1} to be the scaled amount by which the other deviated in the previous period from the standard (one-shot) Nash equilibrium choice. If the other player cooperated in the previous period the impulse

equals 1, if the other defected it equals 0. In this model the choice between cooperating and defecting is fully dependent on the level of α . If α is larger than 1/4 cooperating is a dominant choice while for α smaller than 1/4 defecting is dominant, as in the standard models.

Both Dal Bó and Fréchette (2011) and Fudenberg et al. (2012) find experimental evidence that many subjects in their experiments use either tit-for-tat (TFT) or tit-for-two-tats (TF2T) as a strategy. In this exercise we will see if the simple and neurologically underpinned Ties model (Bault et al., 2014, 2017) could explain the behavior described by these strategies.

3.9.2 Tit-for-tat

The TFT strategy is simple: A player begins by playing C and simply imitates the other player's action in future periods. We can derive conditions on the ranges of parameters δ_1 , δ_2 , and α_{ij1} for which this behavior is sustained. First, we investigate what levels of α can be reached after continuous play of either C or D . Since $\delta_1 < 1$, being exposed to infinitely repeated defection leads to α going to zero. Now, what about continuous play of C ? Noting that the initial value of α vanishes as t goes to infinity, we start from the expression for α in period 2 and iterate:

$$\begin{aligned} \alpha_{ij2} = \delta_{1i}\alpha_{ij1} + \delta_{2i}I_1 & \Rightarrow \alpha_{ij3} = \delta_{1i}^2\alpha_{ij1} + \delta_{1i}\delta_{2i}I_1 + \delta_{2i}I_2 \Rightarrow \\ \alpha_{ijt} = \delta_{1i}^{t-1}\alpha_{ij1} + \delta_{1i}^{t-2}\delta_{2i}I_1 + \dots + \delta_{1i}\delta_{2i}I_{t-2} + \delta_{2i}I_{t-1} \end{aligned} \quad (3.43)$$

Furthermore, if C is played, I is always equal to one, so that:

$$\begin{aligned} \alpha_{ijt} = \delta_{1i}^{t-1}\alpha_{ij1} + \delta_{1i}^{t-2}\delta_{2i} + \dots + \delta_{1i}\delta_{2i} + \delta_{2i} & \Rightarrow \\ \alpha_{ijt} \approx \frac{\delta_{2i}}{1 - \delta_{1i}} \quad (\text{as } t \rightarrow \infty) \end{aligned} \quad (3.44)$$

This is an important result as it gives an upper limit to the alpha level that can be reached

in an infinitely repeated game. Now that we have established both the upper and the lower limit for α , we can check the conditions for which the behavior according to the ties model is identical to the TFT strategy. The first condition is simply that the first action must be to play C . This leads to the simple condition (omitting the i and j subscripts, for convenience):

$$\alpha_1 \geq 1/4 \tag{3.45}$$

We also need that, no matter how high the current level of α is, after only one period of D played by the other, a TFT player weakly prefers D over C . Using eqs. (3.42) and (3.45), it follows that the tie value of a player who experienced infinitely repeated C and one period of D is equal to $\frac{\delta_1 \delta_2}{1 - \delta_1}$. Since D can only be weakly preferred if $\alpha \leq 1/4$, we get the condition that:

$$\frac{\delta_1 \delta_2}{1 - \delta_1} \leq 1/4 \tag{3.46}$$

On the other hand, we also need that, no matter how low α is, after only one period of C played by the counterpart C is weakly preferred over D , which requires (as $\alpha \geq 0$):

$$\delta_2 \geq 1/4 \tag{3.47}$$

By combining (3.46) and (3.47) we find that $\delta_1 \leq 1/2$ should hold. The intuitive explanation for these values is that a player has to be sufficiently sensitive to a changed impulse and must not have too strong of a 'memory'.

3.9.3 Tit-for-2-tats

After having defined the Tie-model parameters for which TFT is the resulting strategy, we now repeat the same exercise for the TF2T behavior. As this strategy also starts with playing C , we need (3.45) to hold. Furthermore, even after continuous D play, after only one period of C by the counterpart, C should be weakly dominating D , so (3.47) should

hold as well.

Before we proceed with imposing further restrictions, we have to decide how strict we want to be in our interpretation of the TF2T strategy. If we take it at face value we have to assume that even after a history like $DDDDDDCD$ a player will still be patient and play C . It also implies that this player would constantly play C against a counterpart that keeps alternating between C and D . In order to account for such phenomena we evaluate two different versions of TF2T, one that takes the strategy literally and one that requires multiple periods of C before the trust in the other is restored. For convenience, the latter version will be called "qualified tit-for-two-tats" (QTF2T). For the standard TF2T we need that, even if we start out with $\alpha = 0$ and the other player cooperates in one round, only to defect immediately thereafter, a player would still reply with C to that D choice. This requires that:

$$\delta_1 \delta_2 \geq 1/4 \tag{3.48}$$

In addition, we need that, even at the highest possible level of α , after two periods of D a player wants to choose D , which requires (using eqs. (3.45)):

$$\frac{\delta_1^2 \delta_2}{1 - \delta_1} \leq 1/4 \tag{3.49}$$

Combining (3.48) and (3.49) gives:

$$\frac{\delta_1}{1 - \delta_1} \leq 1 \text{ or } \delta_1 \leq 1/2 \tag{3.50}$$

So we have:

$$\delta_2 \geq 1/2 \tag{3.51}$$

A potential problem for the result above is that the upper bound of α , $\frac{\delta_{2i}}{1 - \delta_{1i}}$, will be larger or equal to 1 (with equality only if $\delta_1 = \delta_2 = 1/2$), since if we combine (3.48) and (3.44) :

$$\frac{\delta_2}{1 - \delta_1} = \frac{\delta_2 \delta_1}{\delta_1(1 - \delta_1)} \geq \frac{1}{4(\delta_1(1 - \delta_1))} \Rightarrow \frac{\delta_2}{1 - \delta_1} \geq 1 \quad (3.52)$$

This 'problem' can be solved by using a QTF2T strategy where we assume here (for simplicity) that the value of α must be maximized for a player to play C after the other player chooses D . In this case, it is required that:

$$\frac{\delta_1 \delta_2}{1 - \delta_1} \geq 1/4 \quad (3.53)$$

3.10 Appendix D: Payoff Matrix and SVO Example

| | | Other's possible choices | | | | | | | | | | | | | | | |
|-----------------------|--------|--------------------------|-----|-----|-----|-----|-----|-----|-----|--------------|-----|-----|-----|-----|-----|-----|-----|
| | | ← Take | | | | | | | | Contribute → | | | | | | | |
| | | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Your possible choices | Take ↑ | 7 | 84 | 94 | 104 | 114 | 124 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 |
| | | 6 | 98 | 108 | 118 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 |
| | | 5 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 |
| | | 4 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 |
| | | 3 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 |
| | | 2 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | 234 | 244 | 254 | 264 | 274 |
| | | 1 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | 278 |
| | | 0 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 | 270 | 280 |
| Contribute ↓ | 1 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | 278 | |
| | 2 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | 234 | 244 | 254 | 264 | 274 | |
| | 3 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | |
| | 4 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 | |
| | 5 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | |
| | 6 | 98 | 108 | 118 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | |
| | 7 | 84 | 94 | 104 | 114 | 124 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | |

Figure 3.5: Payoff Matrix

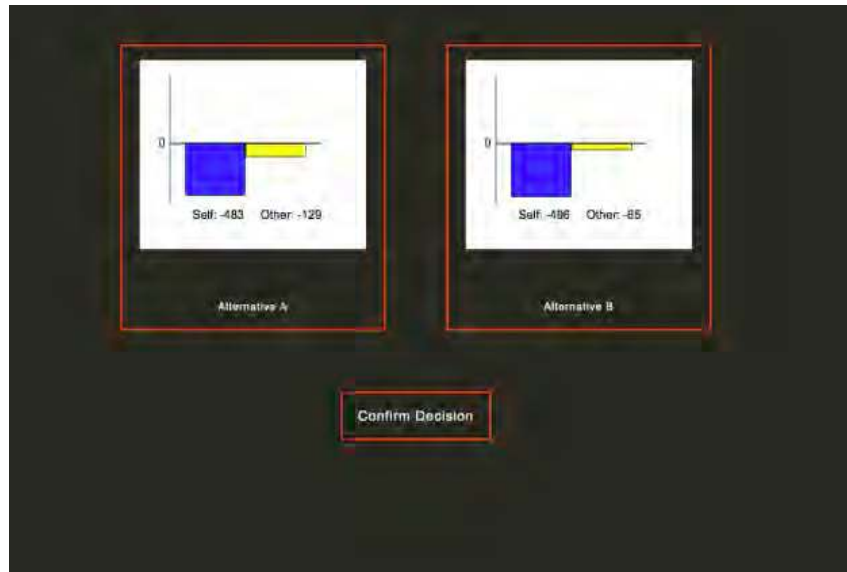


Figure 3.6: SVO Decision Screen

Chapter 4

The Processing of Positive and Negative Social Impulses¹

¹This chapter is based on work with Nadège Bault, Ben Loerakker, and Frans van Winden, which was part of the Research Priority Program Brain & Cognition at the University of Amsterdam. Financial support by the Research Priority Area Behavioral Economics of the University of Amsterdam is gratefully acknowledged.

4.1 Introduction

As part of an increasingly interconnected population of more than 7 billion individuals, one of the most important tasks that any human faces is to evaluate how to behave towards other humans. Since perfect introspection into others' intentions is impossible, we often have to rely on our own emotional reaction to their behavior to guide our decision making in relation to others. This study analyzes the influence of cooperative or destructive actions by others on a subject's affective social tie towards that other individual. In particular, we are interested in how the neural processing of positive and negative impulses to these ties differs.

There is an increasing body of literature on the neural mechanisms that govern affective states. A recent review article (Ruff and Fehr, 2014) collects evidence regarding the hypothesis that social and non-social valuation use shared neural processes, or at least are performed in very closely related regions. A slightly different view can be found in Declerck et al. (2013), which outlines a theoretical model in which social cognition is a distinct system. Previous studies of human social behavior have explored affective attachment, such as friendship (Krienen et al., 2010; Fareri et al., 2012b), sympathy (Decety and Chaminade, 2003) and romantic attachment (Aron et al., 2005; Fisher et al., 2005; Zeki, 2007). Notably, the posterior superior temporal sulcus (pSTS, see figure 4.1) is implicated in response to cooperative partners (Singer et al., 2006), friends and loved ones (Bartels and Zeki, 2000), while the medial prefrontal cortex (mPFC) is involved in making trait judgments of close friends (Heatherton et al., 2006; D'Argembeau et al., 2007; Fareri et al., 2012b), in cooperative decisions (McCabe et al., 2001; Rilling et al., 2004), in trust (Krueger et al., 2007) and in preference plasticity (Garvert et al., 2015). Several studies also point out the role of the striatum in social decision making (Bhanji and Delgado, 2014). There are also studies on the modulation of the updating of experienced social information in the context of trust games (Fareri et al., 2012a). See Li et al. (2014) and Wagner et al. (2012) for further review studies on the neuroscience of social cognition. The temporoparietal junction (TPJ) has been quoted as playing a role in modulating

the value of more or less generous decisions, relative to social distance (Strombach et al., 2015).

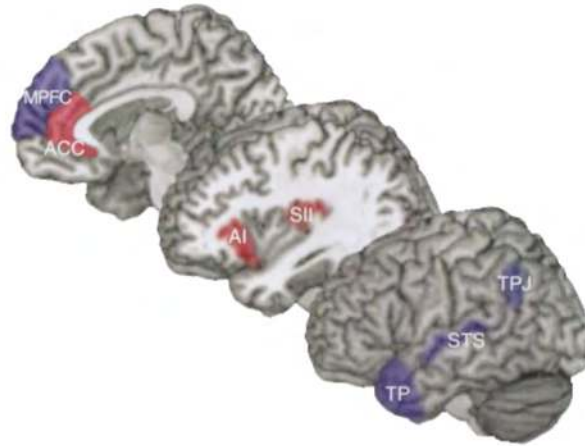


Figure 4.1: Overview of some of the mentioned areas, together with the anterior cingulate cortex (ACC), secondary somatosensory cortex (SII), and temporal poles (TP) (Singer and Tusche, 2014, Figure 27.2)

Understanding how the perception of gains and losses differs is a pivotal element of understanding decision making in general. Effects such as loss aversion (Tversky and Kahneman, 1991) imply that humans systematically differ in how they perceive and value experiences in these different domains. This raises the question whether the way in which such experiences are processed also differs (Sharot and Garret, 2016). Fundamentally the question is whether losses and gains are processed by the same system or by different systems (Rick, 2011). Additionally, possible mechanisms for valuation processing include a reaction in brain activity to experiences irrespective of the domain, or a non-linear, U-shaped reaction. A recent meta-analysis (Bartra et al., 2013) suggests that posterior cingulate cortex (PCC, to be found further in the back than the ACC in figure 4.1) and ventromedial prefrontal cortex (vmPFC) react consistently in a linear fashion, the anterior insula (AI) non-linearly, and the striatum with both types of reaction. This is in line with the idea that different structures play a role in the evaluation of the direction and salience of an effect. Most studies on positive and negative experiences look at rewards and losses that are either immediately presented to the subject or can be consumed later, but are independent of the influence of any other human. A number of studies have also looked

into the processing of social rewards, mostly pointing towards the role of the striatum (Fliessbach et al., 2007; Izuma et al., 2008; Kohls et al., 2013; Zink et al., 2008).

In order to make any comparison between positive and negative effects on a subject's affective state toward another it is imperative to disentangle the income effects of an action from its social meaning. To this end, we use the fragile public good game (FPG) from chapters 2 and 3. The variant we use has the feature that both the effects on one's own income and on the other's income change symmetrically with different decision in the cooperative and competitive domain. This enables us to compare the neural processing of such behaviors. As was demonstrated behaviorally in a number of studies (Abbink and Herrmann, 2009; Hoyer et al., 2014; Nikiforakis and Engelmann, 2011), repeated interaction with an option of destructive behavior can lead to feuds, as expressed by repeated destructive behavior by two participants.

In this study we compare the neural processes underlying the building of positive versus negative ties using functional Magnetic Resonance Imaging (fMRI, see box in 4.2.3 for a quick introduction). We take as a basis a model of social choice that explicitly models the development of social ties over a repeated interaction (van Dijk and van Winden, 1997). In this model the decisions of a specific partner may have a positive or negative *impulse* on the *tie value* by which an individual weights that partner's welfare. It was first tested empirically by van Dijk et al. (2002) and Sonnemans et al. (2006) (see also Bault et al. (2017)). In line with Hoyer et al. (2014) and Loerakker et al. (2016) we adapt the model to a game in which the range of positive and negative decisions is identical and directly comparable. In these studies the model has been demonstrated to track behavior closely and to outperform models of fixed social preferences in certain situations, supporting the idea that the underlying tie value as defined in the model does indeed track the development of subjects' preferences for the monetary earnings of others.

In a previous study (Bault et al., 2014), the neural dynamics of the formation of social ties were analyzed. It was shown, among other findings, that the level of social ties, as defined by the van Dijk and van Winden (1997)-model, was encoded in the pSTS and the TPJ. In addition, the behavior of other participants – the impulse, using the model's

terminology – modulated activity in the pSTS. These results aligned well with a number of studies that find these regions to play a relevant role in social interactions. Examples include the liking of others (Fahrenfort et al., 2012), but also processing of the effects of an action on an other and their future behavior (Behrens et al., 2008; Hampton et al., 2008; Hill et al., 2017; Singer et al., 2006).

More generally speaking, pSTS and TPJ are thought to be part of a system that serves to mentalize the needs, beliefs and emotions of others (Singer and Tusche, 2014). That process is distinct from the processing of both monetary and social rewards, which is generally thought to be represented by processes in the striatum and the vmPFC (Rangel and Clithero, 2014). Disentangling these processes requires the use of model-based analysis, see for example Hampton et al. (2008) and Hill et al. (2017).

Building on the findings in Bault et al. (2014), we are using a similar design, but modify it to match the symmetric design from Hoyer et al. (2014) in a fMRI experiment. Doing so we extend the literature on the processing of positive and negative experiences into the realm of social ties. Based on the hypothesis that there are indeed systematic differences between the way that these different types of information are processed in general, we may expect a similar observation in the specific case of human interaction regarding changes in social ties. Since our new design also allows for negative tie values, the question arises whether positive and negative tie values modulate activity in different regions, and if in the same regions, whether they are encoded linearly or quadratically.

In line with common practice in the neuroscience literature we start out with a Materials and Methods section, section 3, which describes the design of the experiment, including technical details. After the subsequent Results section (4) we move to the Discussion (5) and end with a Conclusion (6) of the results of this experiment.

4.2 Materials and Methods

Each experimental session included 3 participants, one of which was scanned in a fMRI scanner. We first present the two types of tasks that participants performed, followed by

an outline of the general procedure of the experiment. In total 25 subjects participated while in a fMRI scanner ("scanned subject") and a further 47 subjects participated only in the role of interacting with the scanned subjects through a computer system ("non-scanned subject")². We first present the two types of tasks that participants performed: a Social Value Orientation (SVO) test in subsection 3.1, and a Fragile Public Good game (FPG) in subsection 3.2, followed by an outline of the general procedure of the experiment in subsection 3.3.

4.2.1 Social Value Orientation

We used the same combination of 32 questions as used in van Dijk et al. (2002) to determine the subjects' initial level of SVO, based on the SVO test found in Liebrand and McClintock (1988a). The questions are identical to the questions found in chapter 2 of this dissertation.

4.2.2 Fragile Public Good Game

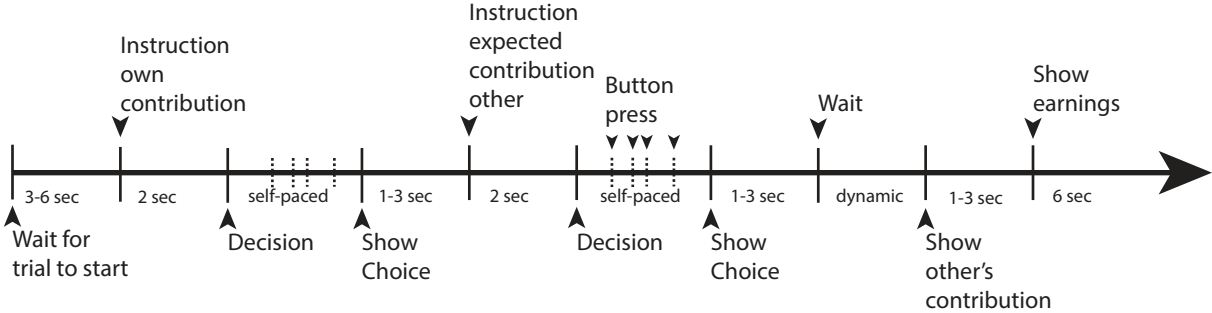


Figure 4.2: Timeline of a Round of the Fragile Public Good Game

The public good game that was used was strategically equivalent to the game used in the "symmetric" treatment in chapter 2, albeit not programmed using the same software. Apart from minor differences in the graphical presentation of the game the main difference is that our scanned subjects could not use a computer mouse to enter their commands but used a controller with 4 buttons to make their choices. Additionally, participants in

²In three cases organizational delays made it impossible to run the second of the two games discussed in 2.2, which is why there are not exactly twice as many non-scanned subjects as scanned subjects.

the experiment presented in chapters 2 and 3 were also given a print-out of the payoff matrix (figure 4.16, identical to figure 3.5). To prevent visual distractions and focus the subjects there was no digital equivalent of that in the scanner. For an equal playing field non-scanned subjects also were not given a printout of the matrix. Non-scanned subjects could use a mouse. Figure 4.2 provides a detailed timeline of one round of the game. It shows the different screens that the subjects saw during one round, together with the time that those screens were visible. Some screens were shown for a fixed amount of time, while others were either self-paced or depended on the speed at which the other subject made their decision. It is worth pointing out that the payoffs of both the individual subject and his or her partner were shown as illustrated in figure 4.3.

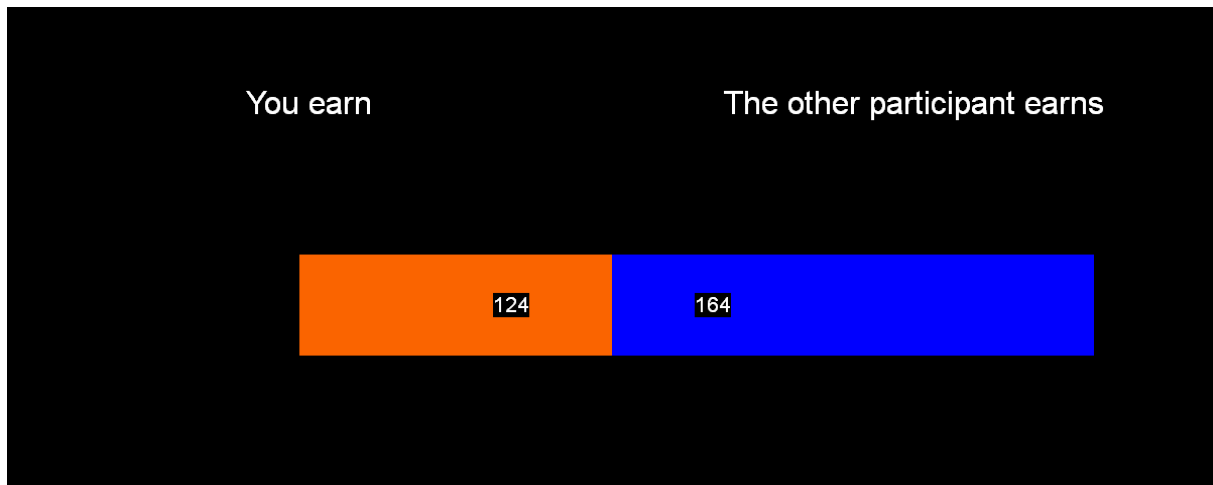


Figure 4.3: Payoff screen

4.2.3 Experimental Procedure

The experiment was conducted throughout 2013 at the Amsterdam Center for Brain and Cognition at the University of Amsterdam. We invited scanned subjects and non-scanned subjects to show up to the lab at different times and briefed them in separate rooms to ensure full anonymity. After signing their consent forms subjects were presented with the Social Value Orientation (SVO) test. This test was conducted using a separate computer outside of the scanning room. Subjects were told that their decisions affected their own earnings and the earnings of a unknown other participant, which was implemented by

assigning their earnings to the other subject without informing them about their earnings until the end of the session. Afterwards the scanned subject was brought to the scanner room and placed inside the scanner.³ Here the scanned subject interacted sequentially with two different non-scanned subjects, who were seated in front of a computer in a separate room that was connected to the control room. The anatomical scan was performed in the break that was necessary while switching the non-scanned subjects, who did not personally meet the scanned subject. In the main part of the experiment the scanned subject and the non-scanned subject played the (FPG) game for 35 rounds. After the 35 rounds were over, the scan was interrupted, but the scanned subject remained in the scanner. The first non-scanned subject was taken out of the lab to a separate room to answer the exit-questionnaire and get paid. During that time the second non-scanned subject was placed in front of the computer for non-scanned participants. Then the scanned subject played another 20 rounds of the FPG with the second non-scanned subject. Afterwards both remaining subjects were taken to separate rooms to fill out their exit questionnaires and be paid their earnings, still maintaining full anonymity.

We recruited scanned subjects at different campuses of the University of Amsterdam using posters and flyers. Since non-scanned subjects did not require any screening or particular information about the experiment, the majority of them was recruited using the recruitment system of the Center for Research in Experimental Economics and political Decision Making (CREED). Scanned subjects were remunerated with a show-up fee of 25 euro and earned between 38.36 and 44.38 euro in total. Non-scanned subjects, for whom the experiment was shorter and noticeably less physically uncomfortable, received a show-up fee of either 7 (first non-scanned subject) or 12 euro (second non-scanned subject)⁴, earning between 16.20 and 20 euro in total. Sessions took approximately 2 hours including briefing and exit questionnaires for scanned subjects and approximately 90 minutes for non-scanned subjects.

³Scanned subjects were also connected to a skin conductance measurement system, but we do not use that data in this analysis.

⁴The show-up fees for non-scanned participants differed because they played a different number of rounds and only the first subject participated in a ring test. This led to different earnings from the experiment itself, despite similar time investment, which we compensated through the show up fee.

All computerized tasks in this experiment were programmed in Neurobs Presentation software package.

Box: Short introduction to fMRI analysis

In functional magnetic resonance imaging the goal of the analysis is to generate a three-dimensional representation of the activity in the brain at a certain point in time. To this end, subjects are placed in a scanner during the experiment. The scanner can pick up on small magnetic field changes caused by differences in the blood oxygen level in a region over time, giving it the name Blood Oxygen Level Dependent (BOLD) analysis. These changes serve as a proxy for the activity in that region. The resulting data is a series of images containing a number of so-called voxels, or points in three-dimensional space, with a particular activation level for each point. In our case these voxels have a size of 3x3x3 millimeters and images are recorded every 2 seconds.

The first step in the analysis process, which is only very broadly summarized here, is preprocessing. This step includes a number of separate procedures that correct for anatomical differences of the brains of the different participants in the experiment, unavoidable tiny head movements, and the fact that the brain is scanned in "slices" rather than at once, which leads to temporal disparities. For technical reasons certain smoothing procedures are also necessary.

After preprocessing we have a time series of images for each subject. In order to be able to analyze this data, it is then combined with timecodes that tell us what part of the experiment a subject was in at any given time. Specifically, these are the events mentioned in figure 4.2. In this case we model each event using a boxcar function, meaning that the event is assumed to take a specific amount of time (generally until the next event starts), as opposed to an impulse at the onset of a time period. We first run a first level analysis for each individual subject. In order to get a predicted signal that tracks the level of activity during that time period, we convolve the binary variable that indicates whether we are within a period or outside of it, with the

canonical Hemodynamic Response Function (HRF, see figure 4.4), which describes the shape of the response in a region if it becomes active. Parametric regressors are interaction terms with these dummies and code the different levels of variables such as contribution and return. Using GLM regression we then fit the predicted signal to the measured signal, estimating coefficients for the regressors in each individual voxel.

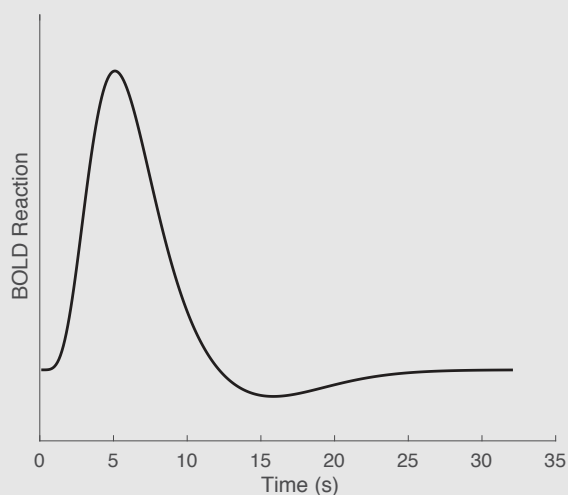


Figure 4.4: HRF as used in SPM. Time = 0s is the time of the event in question.

Finally we need to combine the different results from the first level analysis into a group result. This is done by running t-tests on the regressors found for the different participants at each voxel. At this stage we can also test for the correlation of activation levels with subject-level variables such as SVO or the Social Ties model's parameters. The outcome of this step is a three-dimensional image, which represents the t-test results of the group. The images found in the results section are such group-level t-tests, projected onto a two-dimensional image, where the colored area indicates that results in that area are significant at a certain significance level mentioned. The location of an activated region is usually described by three parameters (X, Y and Z), which describe the position of its peak activation voxel in three-dimensional space in millimeters. Due to the large number of voxels that are being analyzed, correction for multiple comparisons is common when calculating the significance level of each

voxel. In addition to so-called whole-brain analysis a researcher can use regions of interest in order to analyze a specific area. In this experiment we are using this approach to investigate changes in the BOLD signal in particular areas for different levels of some variables.

For further information the interested reader may consult sources such as Poldrack et al. (2011).

fMRI data acquisition

What follows is a technical discussion. The reader is referred to box 4.2.3 for a brief non-technical introduction to fMRI analysis.

Images were acquired using a Philips3T Achieva scanner. Sessions began with the acquisition of a phase image. Functional images were acquired using 2 separate T2*-weighted sequences for the two different partners a participant interacted with (37 coronal slices; flip angle (FA) 76.1°; echo time (TE), 27.63ms; repetition time (TR), 2s; slice thickness, 3mm; field of view (FOV), 240x240mm; in plane voxel resolution, 3x3mm). In between the two functional scans a high-resolution anatomical image was collected (220 coronal slices; FA, 8°; TE, 3.8ms; slice thickness, 1mm; FOV, 240x188mm; in-plane voxel resolution, 1x1 mm).

Pre-processing and data analysis was performed using the Wellcome Trust Centre for Neuroimaging's Statistical Parametric Mapping (SPM) 12 (<http://www.fil.ion.ucl.ac.uk/spm/>). Functional images were corrected for differences in slice acquisition time. Images were then realigned and unwarped to correct for motion artifacts. For each participant, the structural image was segmented and bias corrected based on six tissue probability maps, and coregistered to the mean functional image. Structural data were spatially normalized to the standardized Montreal Neurological Institute space using the deformation field generated by the segmentation routine. The transformation parameters estimated in this step were applied to all functional images. Functional images were spatially smoothed with an 8-mm full width at one-half maximum Gaussian kernel.

fMRI model

The first level fMRI model includes separate regressors for the following types of events (see figure 4.2): a subject's deliberation period, a validation period (during which the decision was entered into the system), displaying the choice that was made, deliberation over the other's expected contribution, validation of the other's expected contribution, the display of the expectation for the other's contribution, the display of the other's contribution, and the display of a positive payoff difference or a negative payoff difference. The level of a participant's contribution, with taking coded as a negative contribution of the respective size, and the expected level of the other's contribution were added as parametric regressors during the two respective validation periods. The actual contribution of the other participant was added as a parametric regressor during the display of it and the value of the participant's tie parameter (α) was added during the first deliberation period. On the group level we also introduced a participant's SVO, as well as the different parameters from the estimation of the behavioral model as parametric regressors during the period immediately after a subject confirmed their decision. We combine the two different FPG sessions, in which subjects interacted with different others, in the analysis of our data.

Questionnaires and demographic data

The ratio of female subjects was 13 out of 25 and average age was 22.6 years. Their fields of study were diverse, Law and Political Science being the biggest single categories with four members each. The average value for the extend to which they believed that the other person was real was 4.7 out of 7, their average "liking" of the partners was 4.4 out of 7⁵

⁵The ex post control question on the game itself proved ineffective, as a noticeable number of subject reported confusion as to its interpretation.

4.3 Results

4.3.1 Behavior

SVO

Among those who participated as subjects in the scanner, 21 out of 25 reported their SVO with a vector length of 600 or more.⁶ After excluding answers with a shorter vector length for inconsistency, we are left with an average angle of 18.7° . Two subjects reported negative angles below -5° . We also collected the SVO of the non-scanned participants who were active in the first game of each session. 21 out of 25 subjects answered the questions with a vector length of more than 600, with an average angle of 4.1° .⁷ 3 subjects reported negative angles below -5° .

FPG

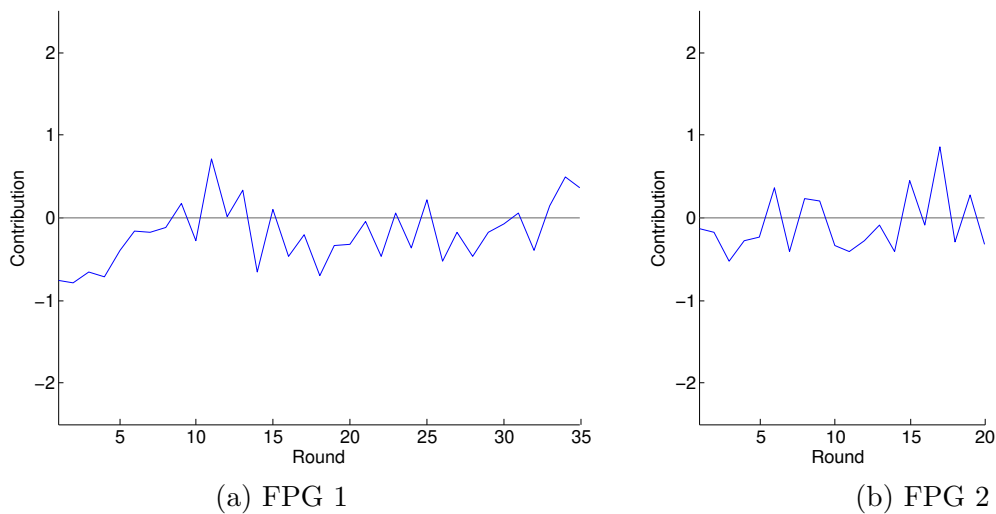


Figure 4.6: Contribution development over time, scanned subjects

The contributions made by the subjects in this experiment are generally close to 0 on average (see table 4.5 for an overview of all data mentioned in this section). There is

⁶Vector length refers to the distance of the endpoint described by the combination of a subject's choices from the center of the circle. It can be interpreted as a measure of consistency.

⁷The difference between average angles among scanned and non-scanned participants is not significant. It is most likely an artifact from a difference in group composition among the two types of subjects, which is a result of using different recruiting methods (see section 4.2.3).

| | Contributions | | Ratio of Contributions >=0 | | Decision Time Contribution | | Expected Contribution Other | | Ratio of Expected Contributions Other >=0 | | Decision Time Expected Contribution Other | | Contribution Other | | Ratio of Positive Contributions Other | | Payoff | |
|------|---------------|--|----------------------------|--|----------------------------|--|-----------------------------|--|---|--|---|--|--------------------|--|---------------------------------------|--|---------|--|
| | tokens | | percent | | seconds | | tokens | | percent | | seconds | | tokens | | percent | | ECU | |
| Mean | -0.12 | | 59.77 | | 4.77 | | 0.21 | | 61.26 | | 4.38 | | -0.27 | | 55.89 | | 191.09 | |
| (SD) | (2.24) | | (27.65) | | (2.79) | | (1.76) | | (24.30) | | (1.53) | | (2.04) | | (28.02) | | (20.68) | |
| Mean | -0.06 | | 60.00 | | 3.89 | | 0.58 | | 65.91 | | 3.81 | | -0.10 | | 57.95 | | 192.55 | |
| (SD) | (2.47) | | (27.60) | | (1.30) | | (2.19) | | (26.35) | | (0.90) | | (2.57) | | (32.94) | | (24.56) | |
| Mean | -0.27 | | 55.89 | | 3.79 | | -0.38 | | 51.43 | | 3.38 | | -0.12 | | 59.77 | | 194.59 | |
| (SD) | (2.04) | | (28.02) | | (1.23) | | (2.28) | | (32.18) | | (0.53) | | (2.24) | | (27.65) | | (20.31) | |
| Mean | -0.10 | | 57.95 | | 4.04 | | 0.49 | | 65.68 | | 3.71 | | -0.06 | | 60.00 | | 193.40 | |
| (SD) | (2.57) | | (32.94) | | (1.55) | | (2.14) | | (29.49) | | (2.05) | | (2.47) | | (27.60) | | (21.40) | |

Note that the scanned subjects in games 1 and 2 are the same individuals, whereas the non-scanned subjects change in between the two games. Negative contribution indicates that subject took more than they contributed.

Figure 4.5: FPG game, average behavioral data

no discernible trend or endgame effect (see figure 4.6). Average contributions are not discernibly different in the two different games, as is the ratio of positive and negative decisions. Decision times decrease by 18 and 13 percent for the participants' own decisions and their guess of the other's decision, but the differences are not significant. Compared to previous studies (Hoyer et al., 2014; Bault et al., 2014) a relatively small number of participants ever experienced a convergence to either zero contributions by both players or to the social optimum of both contributing 4 or 5.⁸ A vast majority of subjects experienced both positive and negative contributions from their peer. Defining a zero contribution to fall into the positive category, 88% experienced both in at least 5 out of the 35 rounds. This is also reflected in their own behavior, with 84% contributing positively at least 5 times. However, only 10 subjects, or 40%, of the participants experienced both positive and negative contributions in each of the two games. Generally speaking only a fraction of the dyads exhibited behavioral patterns in line with previous experiments in this series, such as dyad 2543 in figure 4.7. In many cases we observed fairly random seeming behavior, such as in the case of the second dyad shown in figure 4.7, number 2546.

⁸Approximately 5 scanned subjects, see appendix 4.6.3, dyads 2260, 2489, 2490, 2543, 2544.

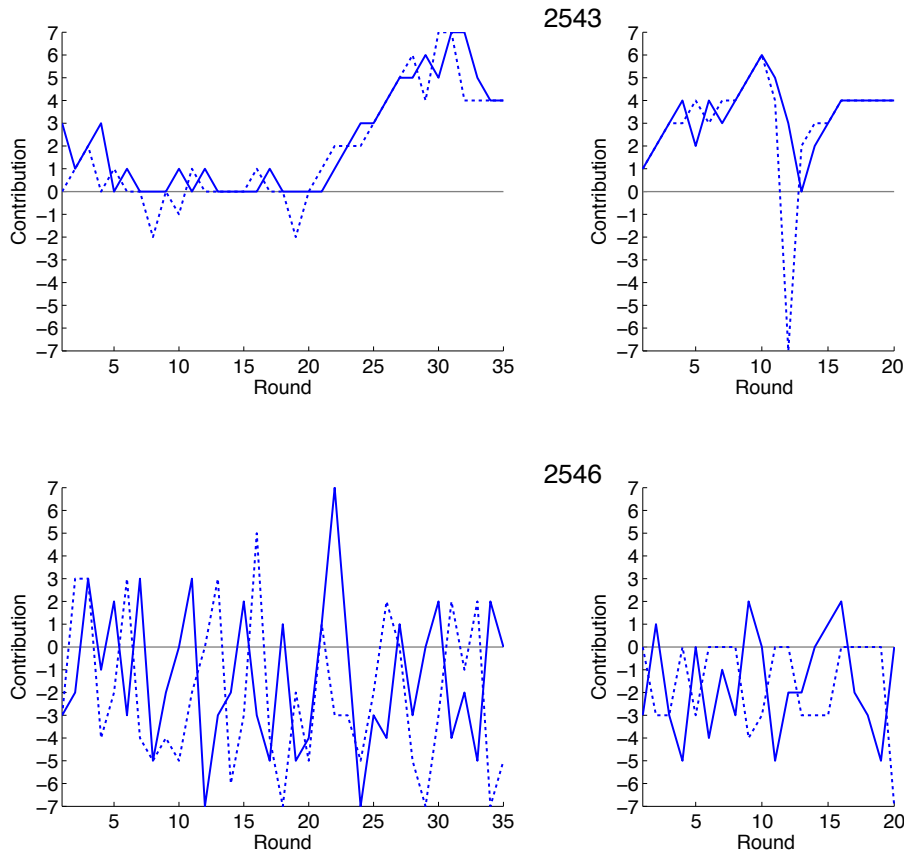


Figure 4.7: Example dyads. Solid line: scanned subject’s contribution; striped line: non-scanned subject’s contribution

4.3.2 Model estimates

Based on the findings in chapter 3 and to facilitate comparability with an earlier experiment (Bault et al., 2014) we use the tie model model using its non-forward-looking version.⁹ We pool all our observations from scanned subjects to construct a group-level estimate. Using only the first game, we reach the results listed in the first row of table 4.1, whereas we pool the 35 rounds from the first game and the 20 rounds from the second game in row 2. The α value is reset to zero between games, as the second game is played with a different participant. Generally we observe that, just as in chapter 3, positive impulses seem to carry a greater weight than negative ones. The most notable difference is that we observe lower θ -estimates, a result of the greater degree of variance in behavior

⁹See chapter 3 for more details on the different variations of the model.

(see appendix 4.6.3 for an overview of the decisions in the separate dyads).

Table 4.1: Estimations group level

| | θ | δ_1 | δ_{2P} | δ_{2N} |
|-------------|-------------|-------------|---------------|---------------|
| FPG 1 | 0.03 (0.02) | 0.66 (0.05) | 0.12 (0.02) | 0.09 (0.01) |
| FPG 1 and 2 | 0.03 (0.00) | 0.61 (0.04) | 0.13 (0.01) | 0.10 (0.01) |

Note: standard errors are between brackets

4.3.3 fMRI results

Effect of contribution

The parametric regressor of the contribution parameter does not show significant results at a threshold of $p = 0.001$. Lowering the voxel-level threshold to $p = 0.005$ (uncorrected) reveals, next to some activity in the occipital lobe, a 12-voxel sized cluster in the left dlPFC correlated positively with contribution (see figure 4.8, peak at -24, 23, 41). Lowering the threshold even further to $p = 0.01$ also reveals some activity in the left TPJ (not shown).

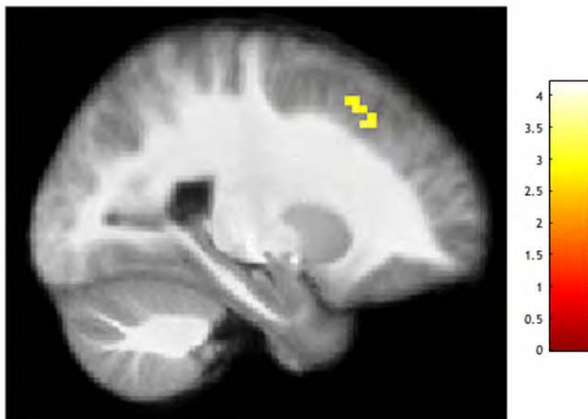
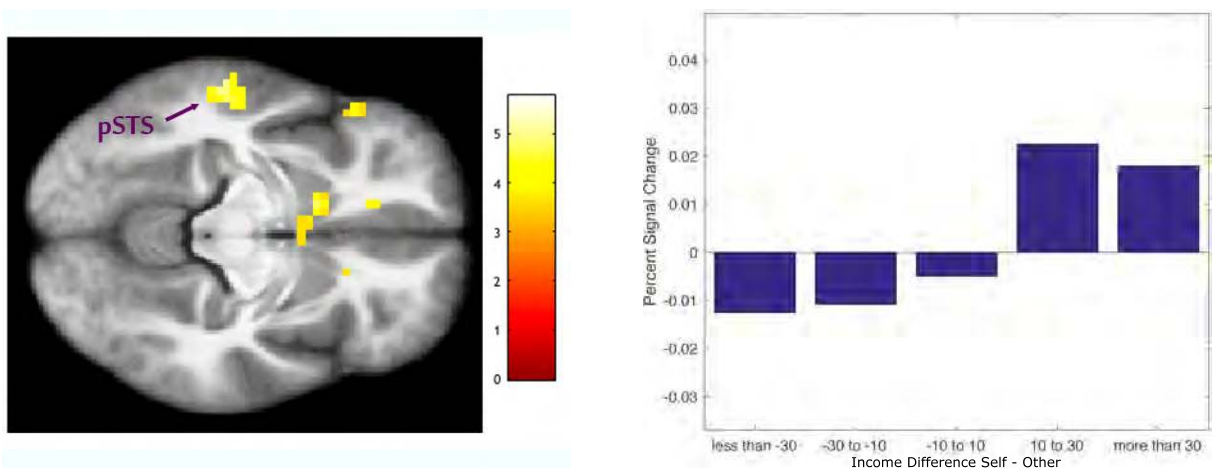


Figure 4.8: Parametric effect of own contribution. T-map projected on averaged brain of all participants. Sagittal view at $X = -24$, uncorrected threshold $p = 0.005$

Effect of the other's behavior and income differences

The parametric regressor of the impulse parameter does not show any significant results. Because the graphical representation of the results of a round of play during the experi-

ment placed a lot of focus on the relative size of the two participants' earnings, this result could be driven by subjects focusing their attention on payoff instead (see section 4.4 for a discussion): A contrast of periods in which a participant earned more than the other relative to periods in which the other earned more (figure 4.9a) reveals a cluster at the left pSTS that is significant at a p-value of 0.01 (uncorrected, peak at -63, -25, -4) with a weaker correlation in the right pSTS (single voxels above uncorrected $p=0.001$ threshold, no notable clustering). In addition, clusters in the prefrontal cortex (-45, 26, 13) and the ventral striatum (-12, 11, -10) are notable (uncorrected p-values between 0.1 and 0.13).



(a) T-map projected on averaged brain of all participants. Axial view at $Z = -10$, uncorrected threshold $p = 0.001$.

(b) Percent signal change, recorded in activation left pSTS cluster with peak at -63, -25, -4.

Figure 4.9: Payoff self minus payoff other.

Additionally, we analyzed the percent signal change in the activated cluster in the left pSTS (figure 4.9b), showing a jump in activation between earning roughly the same as the other and earning more.

Parametric effect of SVO during the decision validation phase

Moving on group level effects, we find a cluster in the dorsolateral prefrontal cortex that is negatively correlated with SVO across subjects during the decision phase (uncorrected p -value = 0.001 at the cluster level, peak at -30, 32, 38, see figure 4.10).

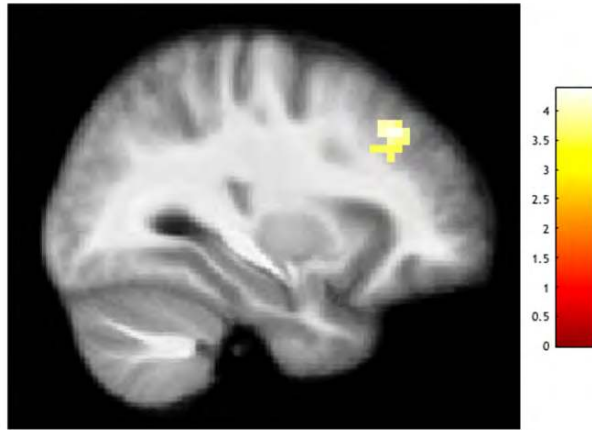


Figure 4.10: Parametric effect of SVO during contribution decision. T-map projected on averaged brain of all participants. Sagittal view at $X = -30$, uncorrected threshold $p = 0.001$

Effects of the different model parameters

The parametric regressor of α , the parameter that models the subjects' tie value, does not show any significant results. We can detect a cluster that is positively correlated with δ_1 in the vicinity of the superior frontal gyrus (peak at -18, 2, 56, cluster level uncorrected p-value 0.13, see figure 4.11a) and a negatively correlated cluster of the same p-value near the anterior insula (27, 26, -2, figure 4.11b). δ_{2N} (the δ_2 parameter that is in effect when encountering negative impulses) shows some singular negatively correlated voxels close to the anterior insula, but no consistent clusters. There is no notable effect of δ_{2P} .



(a) Sagittal view at $X = -18$, uncorrected threshold $p = 0.001$



(b) Sagittal view at $X = 26$, uncorrected threshold $p = 0.001$

Figure 4.11: Parametric positive and negative effect of δ_1 parameters during confirmation of a decision. T-map projected on averaged brain of all participants.

We also test the effect of the activation of different tie parameters in a cluster in the right pSTS, where Bault et al. (2014) found a negative trend in increasing tie values, in order to investigate the question of whether the coding of tie levels behaves linearly in both negative and positive domains. Results of this are shown in figure 4.12.¹⁰

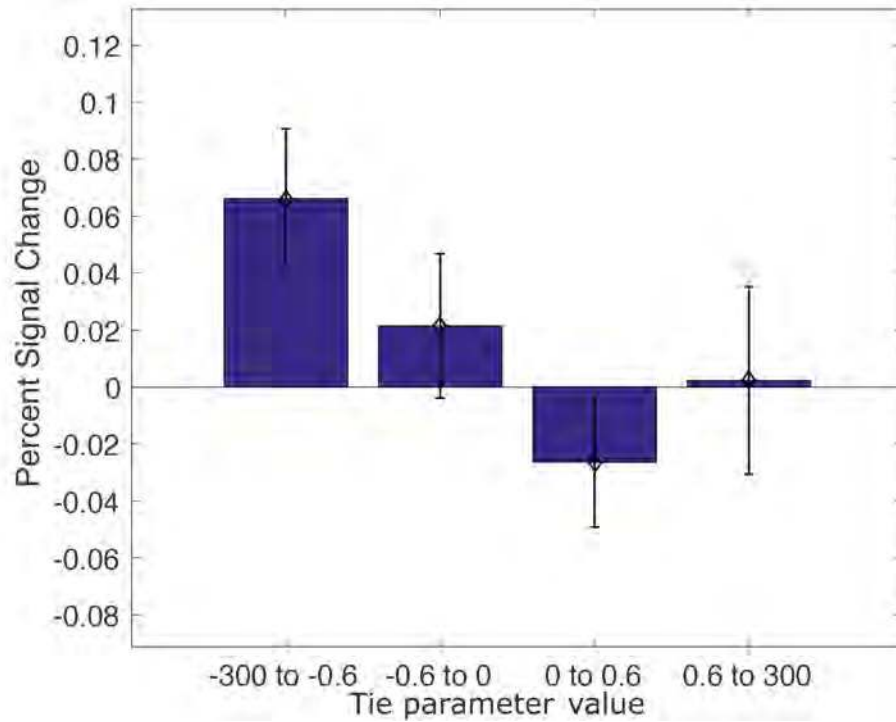


Figure 4.12: Percent signal change in a right pSTS cluster at 46, -40, 0. Tie values binned into 4 bins, limited at a maximum absolute value of 300.

4.4 Discussion

Subjects show more randomness in their choices as compared to previous experiments with a similar design (Bault et al., 2014, 2017; Loerakker et al., 2016), making the estimation of proper subject-level model parameters notably more difficult. On the group level this is noticeable in the weaker estimates for the model’s θ -parameter compared to earlier experiments (see table 4.1). There are two possible explanations for this: One is that the changes in experimental design compared to Bault et al. (2014) – primarily

¹⁰Because of the problems estimating the model we had to exclude a number of values that came from subjects with tie values that reached, and stayed at, rather extreme values, but this required a rather arbitrary choice as to what value would still make sense in the context of the model (here chosen at 300). Bins were chosen so as to be equally wide in range, while also attempting to keep the number of observations in each bin as similar as possible at the same time.

allowing for destructive behavior in addition to different degrees of cooperative behavior – fundamentally changed the way that participants interpreted the experiment and the actions available to them. Seeing as earlier experiments (Hoyer et al., 2014; Loerakker et al., 2016) test this exact difference in design and do not find behavior that is fundamentally different from that in Bault et al. (2014), this explanation seems unlikely to be the main reason behind this behavior. A second explanation is that the increased complexity introduced by a game that allows both cooperative and destructive behavior was too much to deal with without access to the payoff matrix. In difference to both Bault et al. (2014) and Hoyer et al. (2014), subjects used an interface in which they had no continuous access to the payoff matrix of the game (figure 4.16), a decision that was made in order to reduce the amount of visual information participants were exposed to during their decision making process and prevent spurious visual cortex activity.

We observe that the coefficients of parametric regressors introduced in the fMRI analysis differ from the findings in Bault et al. (2014). This is likely a result of the fact that many subjects varied their behavior steadily in a fairly random fashion, compared to the frequent convergence to positive and negative equilibria found earlier. This behavior suggests that they differed notably in their interpretation of the meaning and impact of different decisions that they and their partners took. It would therefore be surprising if their processing of the resulting experiences were found to be comparable to what was found in earlier experiments. This makes it inherently difficult to investigate the role of the underlying social tie value α in decision making and its representation as regarding to brain activity. It appears that, since the dynamic analysis of ongoing play was rather difficult for participants, they reverted to simpler behavioral traits and more general ways of analyzing the game to determine their decisions in the experiment than those formulated by the model.

Together with the fact that, unfortunately, only very few sessions turned out to provide enough data to estimate separate parametric impulse regressors in the positive and the negative domain, as shown in section 4.3.1¹¹, our main hypothesis, namely whether pos-

¹¹Running a model in which two different parametric regressors are used for positive and negative

itive and negative impulses are processed differently, is difficult to measure directly. We do however find an interesting result when looking at a region that was found to correlate with tie-values (α -parameter) in Bault et al. (2014): activations found here are consistent with the hypothesis that also in the negative (destructive) domain, right pSTS activation decreases in tie value, as opposed to a quadratic reaction curve (see figure 4.12). Looking at a subset of subjects for which the estimation of the behavioral models performed relatively well (see appendix 4.6.7) we even see an almost perfectly linear trend across both domains (see figure 4.13). The difference between this result and figure 4.12 could be explained by the fact that our overall sample contains many subjects for which the estimation procedure for the behavioral model struggled to produce meaningful results. Notably, both versions of this analysis show a clear downward trend within the negative tie scores. This is consistent with the statement that tie values are processed linearly, as a previous study already established a negative relationship between right pSTS and tie value in an environment where almost no destruction was possible (Bault et al., 2014). Our data shows a similar relationship in an environment where the option of destructive behavior allows for the development of truly negative ties.

contributions decreases the available number of usable sessions from 47 (2 per participant, minus 3 cases in which the second game was not run) to 33. Three candidates have to be excluded completely because they did not experience sufficiently varied impulses in either game. The reason for this is that in order to estimate a parametric regressor properly we require at least 2 different cases of both cooperative and destructive play by the other player within a session.

The lower amount of usable data and the larger number of regressors makes it more difficult to investigate any part of the experiment, such as payoff contrasts. When contrasting the parametric regressors for positive and negative impulses nonetheless, the only notable area is a higher activation of the left thalamus for *negative* impulses (peak value at -21, 28, 5, 7 voxels above $p = 0.001$ uncorrected threshold. See appendix for figure 4.24).

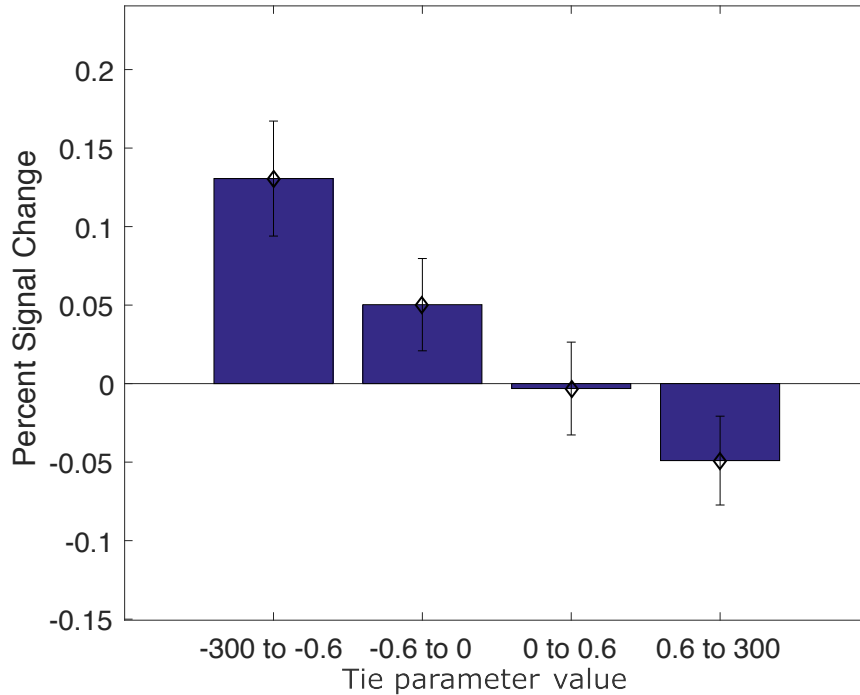


Figure 4.13: Percent signal change in a right pSTS cluster at 46, -40, 0 (taken from Bault et al. (2014)). Tie values binned into 4 bins, limited at a maximum absolute value of 300, filtered for a θ value of at least the median value.

A characteristic that provides information about potentially relevant behavioral general traits is a participant’s level of SVO, as measured prior to the main game. The contributions made by participants were positively correlated with SVO (see appendix 4.6.4). In the case of their average contribution during the first five rounds, a measure found to be fairly reliably connected to SVO (Hoyer et al., 2014), this relationship is significant at the 10% level (one-sided). Consistent with the (ex-post formulated) hypothesis that participants reverted to their general cooperative or destructive disposition in their decision making, we do see modulation of activity in regions associated with social decision making. Our observations regarding dlPFC activity levels is consistent with recent findings that this region is more active in proselves during decision making in a Prisoner’s Dilemma game (Fermin et al., 2016).

Looking at the results phase of the experiment, when the other’s decision and own and other’s payoffs were presented, it appears that subjects focused more on payoff than on the impulse (the other’s decision). This can be explained by the rather short time window

during which the other's decision was displayed (as short as one second), combined with the illustrative nature of the payoff screen (see figure 4.3). The timing of these two phases was designed to keep both types of information relevant, as otherwise the payoff would have been presented after the game has already been largely processed. Combined with the way the figure used to illustrate the results was designed (figure 4.3), however, it seems to have triggered an unexpected focus on comparing own income and the income of the other participant. Alternative models using own payoff and efficiency, i.e. joint income, as parametric regressors did not yield any meaningful results. Unfortunately, this made a full analysis of differences in the processing of positive and negative impulses, one of the main goals of this experiment, largely impossible. Ultimately this lead us to select a model that, other than in earlier experiments in this series, focuses on this binary difference rather than absolute payoff. Using the fairly simple data point of differences in earnings rather than the other's decision, the interpretation of which requires fully understanding and analyzing the game, is also consistent with the interpretation that participants used payoff as a proxy for the intention behind the other participants decision.¹² Another reason in favor of this interpretation is that, in difference to previous experiments, subjects did not see the payoff matrix during the experiment, a decision made to reduce the visual complexity of the decision making screen. It did however also further complicate the understanding of the relationship between other's (and own) contributions and payoffs. Taken together, these arguments could explain why we find fairly clear activity in the pSTS for payoff, whereas an earlier experiment (Bault et al., 2014) found similar patterns when looking at the other's contribution, a variable that is completely uninformative for activity in our analysis. Activity found in the ventral striatum is in line with known results relating to more general social comparisons in the domain of monetary earnings (Fliessbach et al., 2007; Tricomi et al., 2010).¹³ Results in the PFC are typically found

¹²In fact, while the correlation between the other's contribution and own payoff is not necessarily perfectly correlated, as payoff also depends on own contribution, we observe in practice that in our sample other's contribution and own payoff are indeed highly correlated ($\rho = 0.90$, and $p < 0.01$, using each round as a separate observation). Differences in payoff are strongly correlated with other's contribution as well ($\rho = 0.51$, $p \leq 0.001$).

¹³In (Tricomi et al., 2010) this case is only observed for participants with relatively low endowments, a distinction that does not have an equivalent in our experiment.

in more medial areas (Pessiglione and Delgado, 2015), whereas we found activity closer to the dlPFC. A potential explanation for this difference might once more lie in the relatively high degree of complexity of our paradigm. Since the exact interplay between contributions and income is not trivial, it is possible that treating experiences the same as simple and precise positive or negative experiences is an oversimplification. The dlPFC has been shown to be involved with tasks that involve working memory and reasoned calculation (Miller and Cummings, 2007), consistent with the idea that our tasks requires some processing of experiences.¹⁴

On to the topic of individual model parameters we follow the line of argumentation in (Bault et al., 2014) that different values for δ_1 , δ_{2P} , and δ_{2N} can be interpreted as personality trait, namely tie persistence and positive and negative proneness to changing a tie in reaction to impulses. The fact that δ_{2P} and δ_{2N} do not show any significant correlations is hardly surprising and determining such was not the goal of the experiment: They are estimated on a subset of experienced impulses, namely the positive and negative contributions by the other. However, since our first level model does not distinguish between positive and negative impulses, also the group level variable cannot be separated between these two different cases. Hence, in lieu of any other options, we are forced to introduce the two variables during the contribution event without selectively applying them to only positive or negative impulses.

As outlined above, the estimation of correlations with model-based regressors suffers from the fact that behavioral patterns in this experiment did not align quite as well with the model as in earlier experiments with fairly similar designs, in particular Bault et al (2014) and Hoyer et al (2014). For the underlying mechanism of the model to be visible one would expect participants to move in the same direction – either more cooperative or more destructive play – for at least parts of a game. Many dyads in this experiment did not follow that behavioral pattern, but displayed more volatile decision making. To

¹⁴The dlPFC is also frequently mentioned in the context of overriding selfish impulses (Rilling and Sanfey, 2011). It should also be noted that it has been found to react to the experience of *unfair* treatment, which could be interpreted as being similar to less contribution by the other in our experiment (Sanfey et al., 2003), while parts of the dlPFC also react the other way around (Glimcher and Fehr, 2014).

gain some more insight on the way in which this might have affected results, we selected a subset of the participants for further analysis. Results did not change meaningfully. While we had speculated to see clearer results in this subset, in particular in regards to contribution and the tie parameter α , this was not the case. For details see appendix 4.6.7.

4.5 Conclusion

As shown in sections 4.3 and 4.4, results of this experiment proved challenging to analyze due to more volatile behavior than expected. This was despite the fact that a similar game, designed to apply the Social Ties model, had already been tested in a fMRI environment (Bault et al., 2014). On top of that the extensions made here, most notably the introduction of an additional destructive domain of decisions, had been successfully tested in an extensive behavioral experiment (Hoyer et al., 2014; Loerakker et al., 2016). It appears that the combination of these extensions with the added stress of a fMRI experiment relative to the purely behavioral experiment and the missing payoff matrix in the interface increased the number of subjects who only possessed a superficial understanding of the game. An additional reason for the differences between Hoyer et al. (2014) and this experiment could lie in the different composition of subject groups, where the former relied more heavily on economics students, who can be assumed to have a greater than average understanding of game theory.

The results of this experiment suggest a number of potential improvements for the future implementation of interactive games designed for the dynamic analysis of cooperative and/or destructive behavior. While experimenters always strive to design experiments that are simple enough to be fully grasped by a subject before the start of the experiment, care should be taken that any misunderstandings can at least be unlearned during the experiment. In this experiment this could have been facilitated by displaying a full payoff matrix when illustrating the result of each round, a feature that was cut from the design compared to Bault et al. (2014) in order to reduce any confounding factors from the fMRI

analysis. Without this repeated feedback about the underlying mechanisms of the game subjects might have been prevented from improving their understanding of the design over time. Potentially this might even have lead to increasingly misguided interpretations of observed events. A stronger focus in presentation on the contribution of the other participant vis-à-vis the subject's own payoff would have been more in line with the underlying model and has the potential to support findings about the processing of the other's contribution.

Despite the challenges presented by the high degree of randomness observed in the subjects' behavior, some main takeaways stand out. First, in line with Hoyer et al. (2014), we do observe a high amount of destructive decisions, despite the fact that it is costly to make these decisions. Second, a group level estimates of the parameters for the Tie Model are comparable with the results found in Loerakker et al. (2016) and (Bault et al., 2014). Third, contribution and SVO correlate with activation in the dlPFC. The latter is consistent with the interpretation that α_0 , the initial Social Tie, plays a relevant role in social decision making under the above described circumstances, namely that our subjects had trouble dynamically interpreting the events that unfolded during the course of the experiment. α_0 can be considered a fallback tool for deciding how cooperative to act in such a situation. Fourth, in line with Bault et al. (2014) higher tie (alpha) values go together with a decrease in activation in a right pSTS region, an effect that we can now observe also for more negative tie values than before, consistent with the hypothesis of a linear reaction. Finally, experiencing relatively high payoffs compared to the partner's payoff elicited activity in the pSTS among our subjects.

4.6 Supplementary Material

4.6.1 Instructions

These slides show the part of the instructions which differed from the instruction presented in chapter 2.

INSTRUCTIONS

You will now receive the instructions for the experimental parts you will complete inside the scanner.

You will interact with another participant, sitting in a small room next to the scanner room. This person is different from the "Other" in the task you just finished.

It will consist of two parts, that are completely independent of each other. Your decisions in part 1 will not influence your potential earnings in part 2.

Each part will consist of 35 rounds. In all the rounds of part 1 you will be paired with the same other participant. Then another participant will take his/her place and you will go through part 2.

In each part you will be able to earn money, where it again holds that your earnings are expressed in Monetary Units with an exchange rate of 700 units = 1 Euro.

← Use arrow left to go to the next screen →

INSTRUCTIONS

Decisions and Earnings

In each round you will be asked to make a decision. This decision always concerns the allocation of tokens over two accounts: a Private Account and a Common Account.

Tokens in the **Private Account** lead to earnings for yourself only.

Tokens in the **Common Account** lead to earnings not only for yourself but also for the participant you are paired with.

The participant you are paired with has to take a similar decision; tokens in his or her Private Account lead to earnings for her- or himself only, while tokens in the Common Account lead to earnings not only for this participant but also for you.

← Use arrow left to go back or arrow right to go to the next screen →

INSTRUCTIONS

At the beginning of each round there will be 7 tokens in your Private Account and 14 tokens in the Common Account (that you share with the other participant). The other participant will also have 7 tokens in his/her private account.

Each token in your Private Account earns you 10 MU, while every token in the Common Account earns you as well as the participant you are paired with 10 MU. In every round you have the possibility to either take tokens from or contribute tokens to the Common Account. Tokens that you contribute to the Common Account are subtracted from your Private Account, while tokens that you take from the Common Account are added to your Private Account.

You can both take and contribute a maximum of 7 tokens. Note, however, that taking or contributing tokens comes with a transfer cost that changes with the number of tokens. The first token that you take or contribute costs you 2 MU. The second token costs you 4 MU (so, 2 tokens cost in total: 2+4=6 MU), the third token costs you 4 MU extra (so, 3 tokens cost in total: 2+4+6=12 MU), etc.

The same holds for the contribution of tokens to the Common Account, with the first token costing you 2 MU, the second token 4 MU, the third 6 MU, and so forth.

← Use arrow left to go back or arrow right to go to the next screen →

INSTRUCTIONS

This table shows the extra transfer costs and total transfer costs for tokens contributed or taken:

Transfer Costs of Taking/Contributing

| Token number | Transfer Cost | Total Transfer Costs |
|--------------|---------------|----------------------|
| 1 | 2 | 2 |
| 2 | 4 | 6 |
| 3 | 6 | 12 |
| 4 | 8 | 20 |
| 5 | 10 | 30 |
| 6 | 12 | 42 |
| 7 | 14 | 56 |

← Use arrow left to go back or arrow right to go to the next screen →

INSTRUCTIONS

Thus, for example:

- Contributing 3 tokens to the Common Account will increase your earnings from the Common Account by 3x10 = 30 MU. You will lose a total of 12 MU for the cost of transferring the tokens and lose 3x10=30 MU because the tokens are no longer in your private account. In total your earnings decrease by 12 MU, while the other participant's earnings increase by 3x10=30 MU.
- Taking 3 tokens from the Common Account will decrease your earnings from the Common Account by 3x10 = 30 MU. You will lose a total of 12 MU for the cost of transferring the tokens and earn 3x10=30 MU because the tokens are now in your private account. In total your earnings decrease by 12 MU, while the other participant's earnings decrease by 3x10=30 MU.

In short: both contributing and taking only one token will decrease your earnings by 2 MU. If you take or contribute another (second) one this will cost you 4 extra, etc. However, for every token you take the other participant loses 10 MU, while for every token you contribute s/he will gain 10 MU.

← Use arrow left to go back or arrow right to go to the next screen →

INSTRUCTIONS

The table below shows your earnings (in MU) for all the possible choices you and the participant you are paired with can make. If you take 7 and the other contributes 1 your earnings are shown at the intersection of row 'take 7' and column 'contribute 1', which is 214 MU.

| | | Other's possible choices | | | | | | | | | | | | | | | |
|--------|----|--------------------------|-----|-----|-----|-----|-----|-----|------------|-----|-----|-----|-----|-----|-----|-----|-----|
| | | Take | | | | | | | Contribute | | | | | | | | |
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Take 0 | 0 | 70 | 60 | 50 | 40 | 30 | 20 | 10 | 0 | 70 | 80 | 90 | 100 | 110 | 120 | 130 | 140 |
| Take 1 | 10 | 60 | 50 | 40 | 30 | 20 | 10 | 0 | -10 | 80 | 90 | 100 | 110 | 120 | 130 | 140 | 150 |
| Take 2 | 20 | 50 | 40 | 30 | 20 | 10 | 0 | -10 | -20 | 90 | 100 | 110 | 120 | 130 | 140 | 150 | 160 |
| Take 3 | 30 | 40 | 30 | 20 | 10 | 0 | -10 | -20 | -30 | 100 | 110 | 120 | 130 | 140 | 150 | 160 | 170 |
| Take 4 | 40 | 30 | 20 | 10 | 0 | -10 | -20 | -30 | -40 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 |
| Take 5 | 50 | 20 | 10 | 0 | -10 | -20 | -30 | -40 | -50 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 |
| Take 6 | 60 | 10 | 0 | -10 | -20 | -30 | -40 | -50 | -60 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 |
| Take 7 | 70 | 0 | -10 | -20 | -30 | -40 | -50 | -60 | -70 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 |

← Use arrow left to go back or arrow right to go to the next screen →

INSTRUCTIONS

| | | Other's possible choices | | | | | | | | | | | | | | |
|-----------------------|------|--------------------------|-----|-----|-----|-----|-----|-----|------------|-----|-----|-----|-----|-----|-----|-----|
| | | Take | | | | | | | Contribute | | | | | | | |
| Your possible choices | Take | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | 7 | 84 | 94 | 104 | 114 | 124 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 |
| 6 | 98 | 108 | 118 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | |
| 5 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | |
| 4 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 | |
| 3 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | |
| 2 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | 234 | 244 | 254 | 264 | 274 | |
| 1 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | 278 | |
| 0 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 | 270 | 280 | |
| 1 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | 278 | |
| 2 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | 234 | 244 | 254 | 264 | 274 | |
| 3 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | |
| 4 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 | |
| 5 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | |
| 6 | 98 | 108 | 118 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | |
| 7 | 84 | 94 | 104 | 114 | 124 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | |

The other participant has the same table. Therefore, to see his/her earnings you have to switch between you and other. Thus, if you take 2 and the other contributes 1, the earnings of the other are at the intersection of row 'contribute 1' and column 'take 2' and is equal to 188 MU. To enable you to check your understanding of these instructions we will now ask you to complete a few exercises.

← Use arrow left to go back or arrow right to go to the next screen →

INSTRUCTIONS

To enable you to check your understanding of these instructions we will now ask you to complete a few exercises.

You will not be able to go back to these instructions once you press the space bar.

← Use arrow left to go back to the instructions or press the SPACE bar to start the exercises

The buttons will be used such as:
 Blue = left
 Yellow = right
 Green = up and down
 Red = confirmation button

4.6.2 Payoff Matrix

| | | Other's possible choices | | | | | | | | | | | | | | |
|-----------------------|------|--------------------------|-----|-----|-----|-----|-----|-----|------------|-----|-----|-----|-----|-----|-----|-----|
| | | Take | | | | | | | Contribute | | | | | | | |
| Your possible choices | Take | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | 7 | 84 | 94 | 104 | 114 | 124 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 |
| 6 | 98 | 108 | 118 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | |
| 5 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | |
| 4 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 | |
| 3 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | |
| 2 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | 234 | 244 | 254 | 264 | 274 | |
| 1 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | 278 | |
| 0 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 | 270 | 280 | |
| 1 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | 278 | |
| 2 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | 234 | 244 | 254 | 264 | 274 | |
| 3 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | 248 | 258 | 268 | |
| 4 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | 260 | |
| 5 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 | |
| 6 | 98 | 108 | 118 | 128 | 138 | 148 | 158 | 168 | 178 | 188 | 198 | 208 | 218 | 228 | 238 | |
| 7 | 84 | 94 | 104 | 114 | 124 | 134 | 144 | 154 | 164 | 174 | 184 | 194 | 204 | 214 | 224 | |

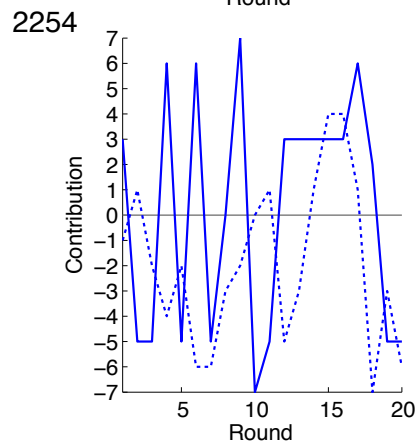
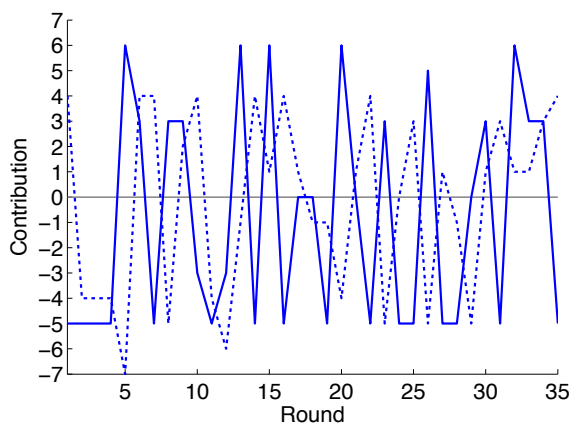
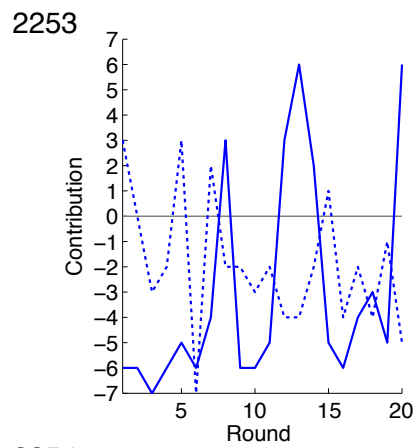
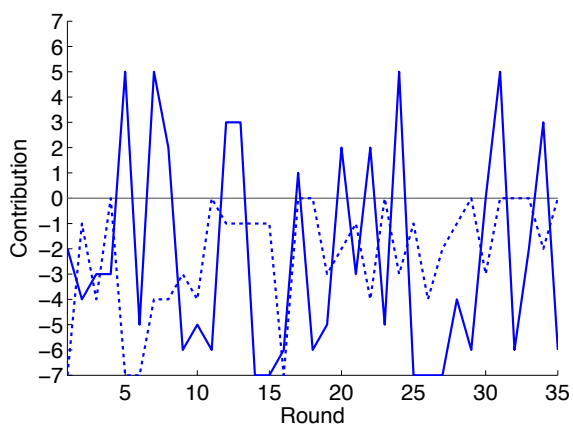
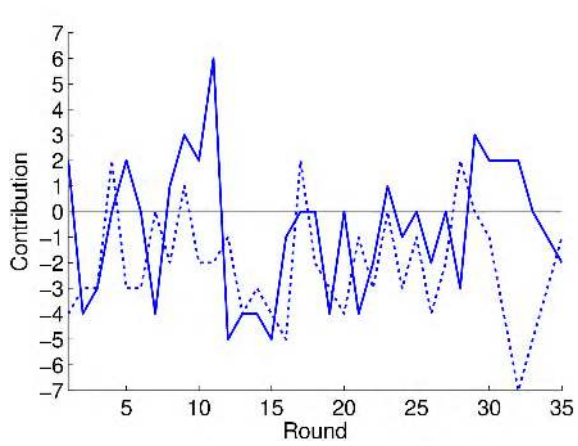
Figure 4.16: Payoff Matrix

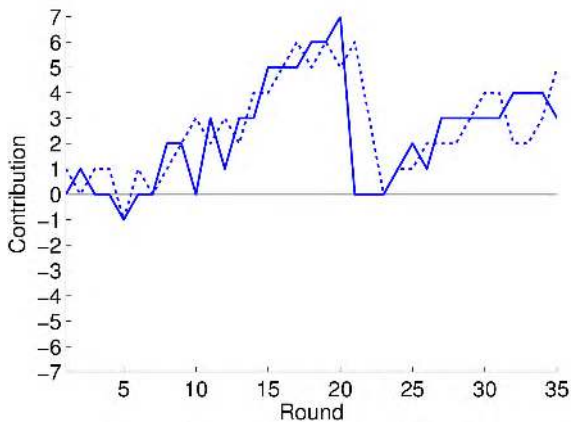
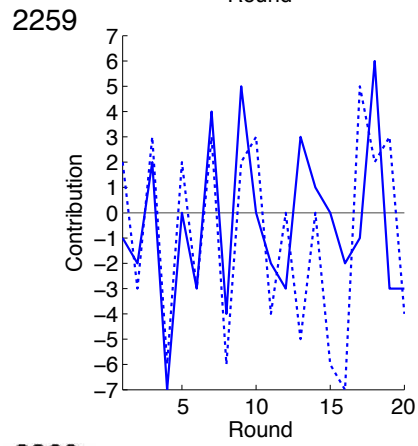
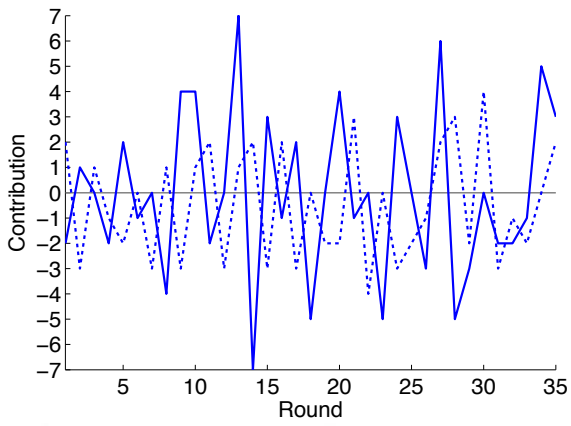
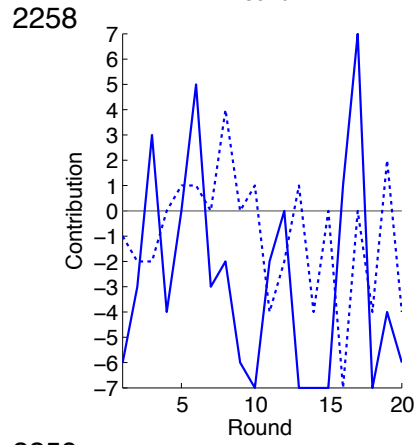
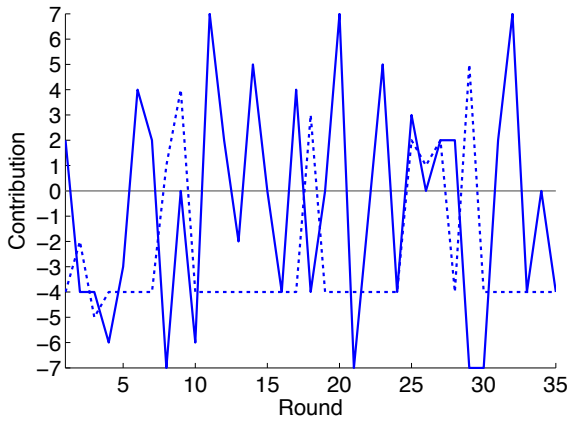
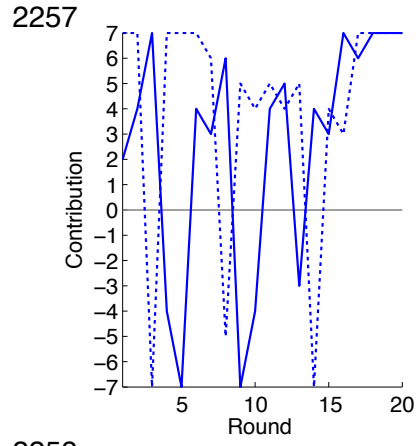
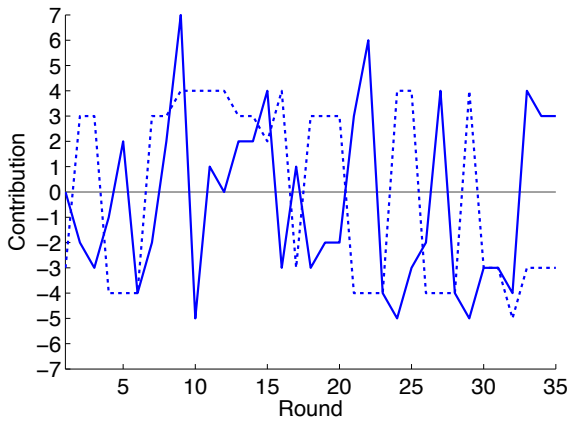
4.6.3 Contributions in individual Dyads

Legend:

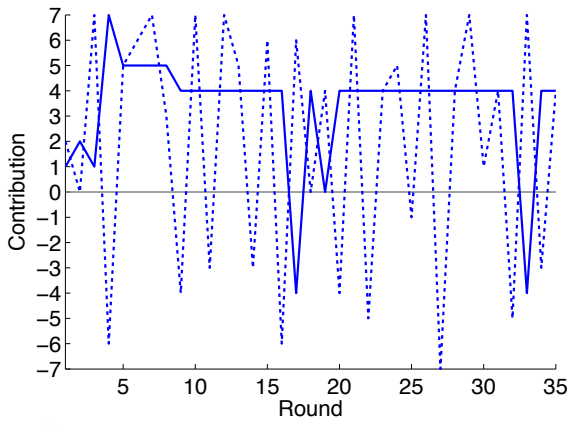
Scanned subject's contribution: _

Non-scanned subject's contribution: ..

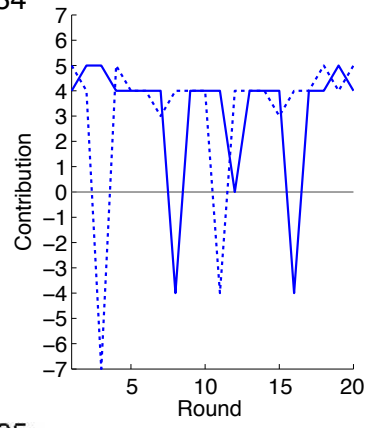




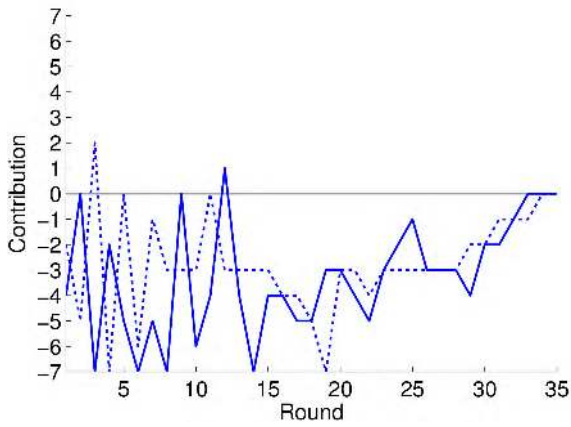
2260



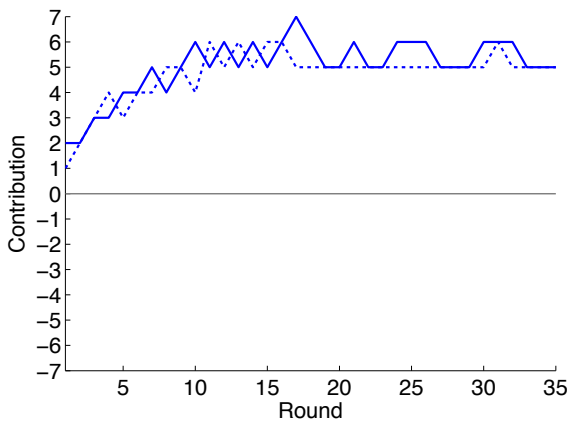
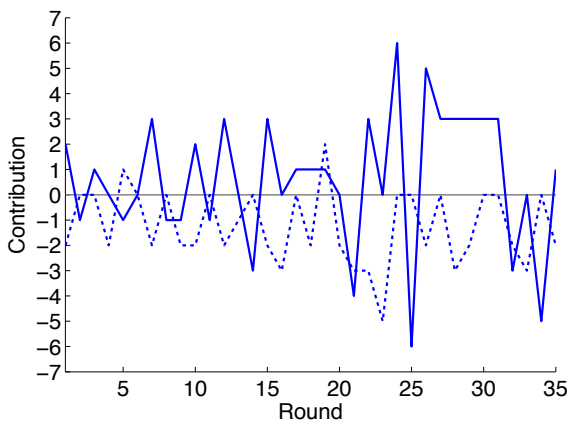
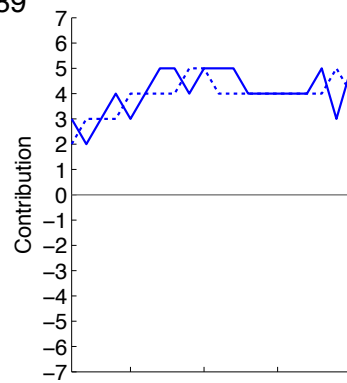
2484



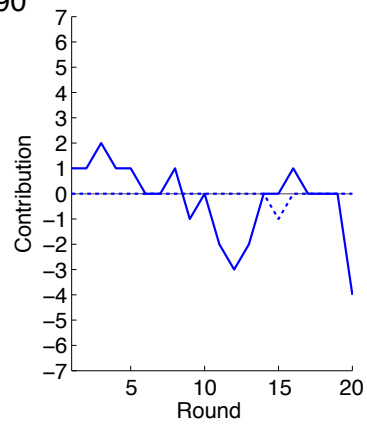
2485

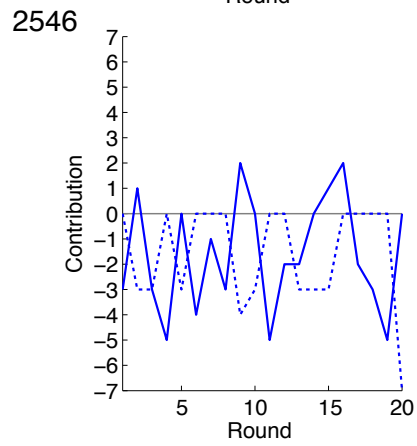
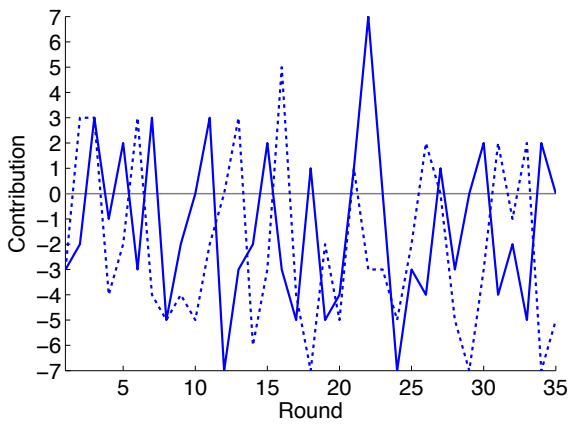
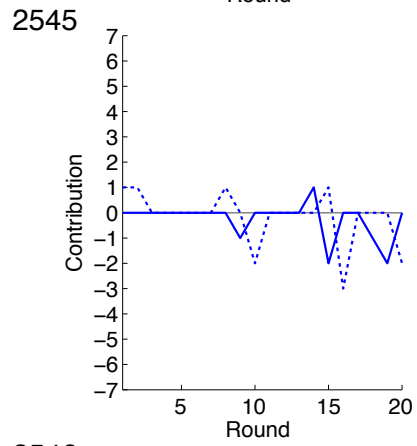
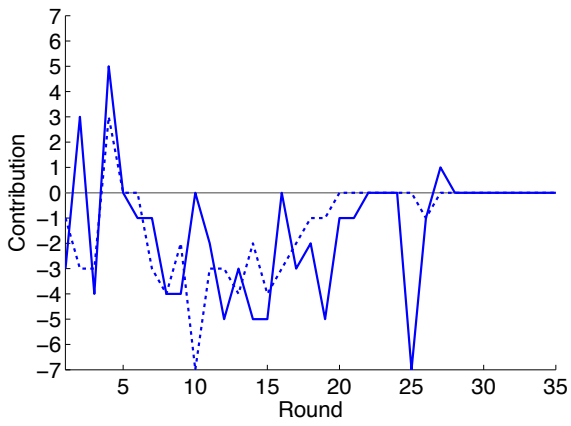
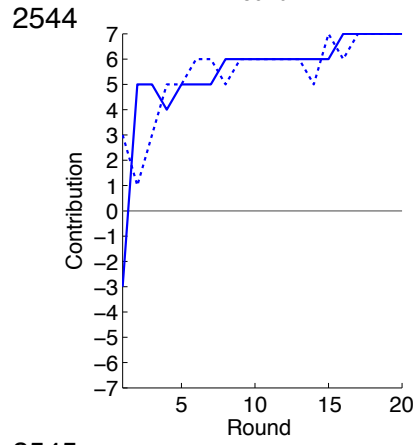
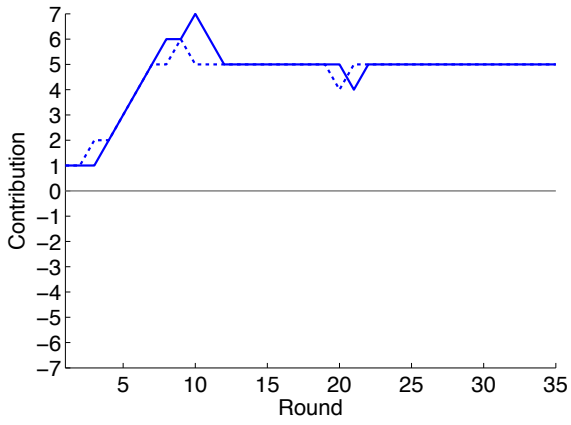
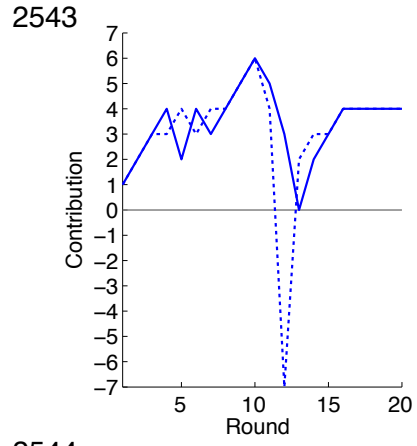
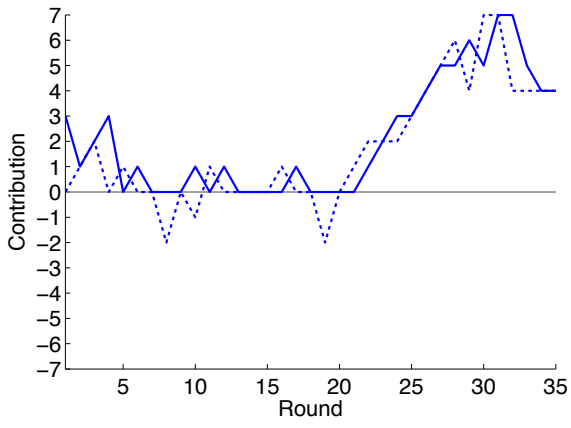


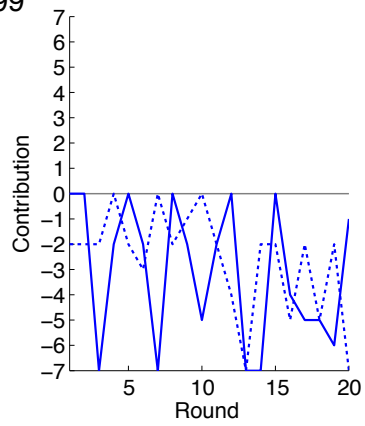
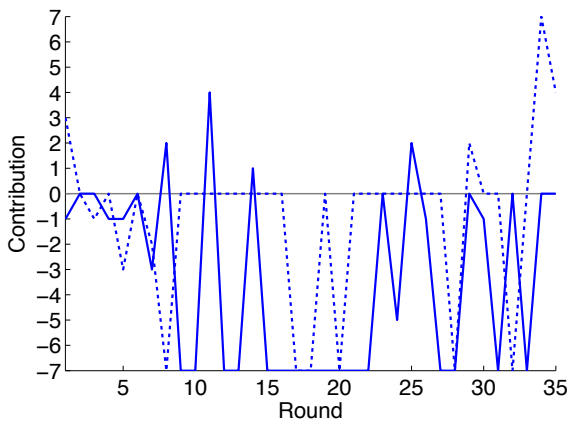
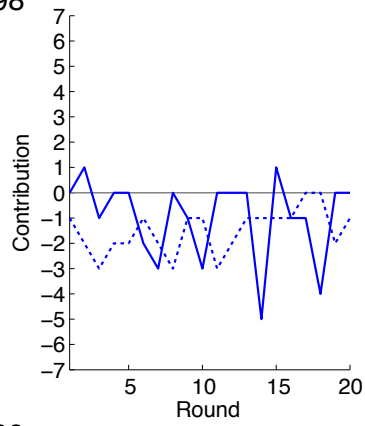
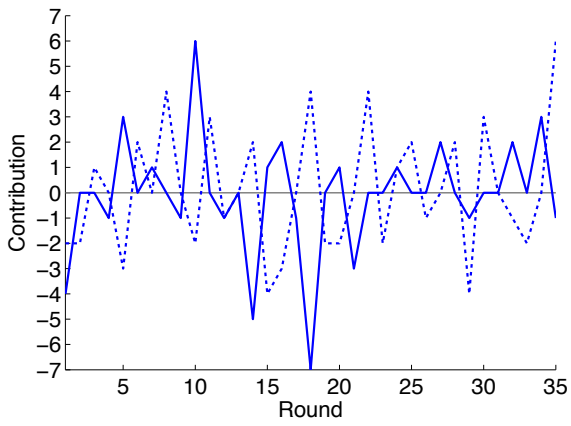
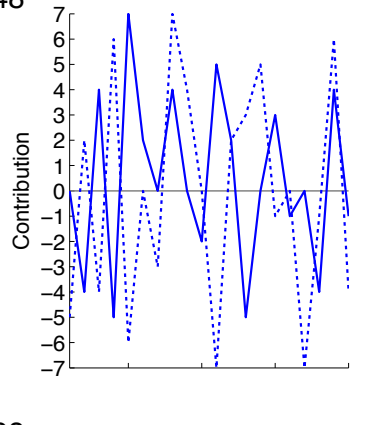
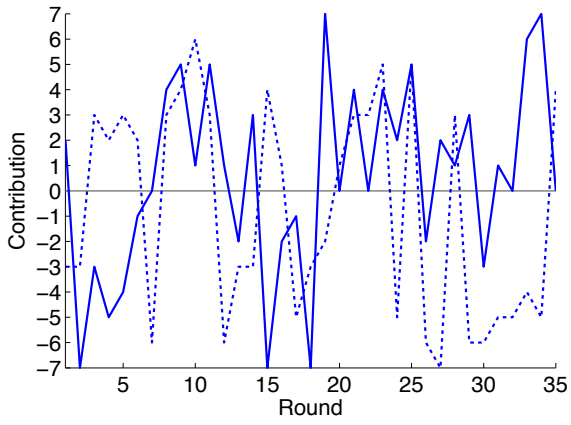
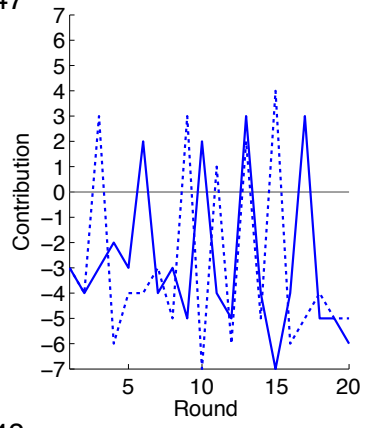
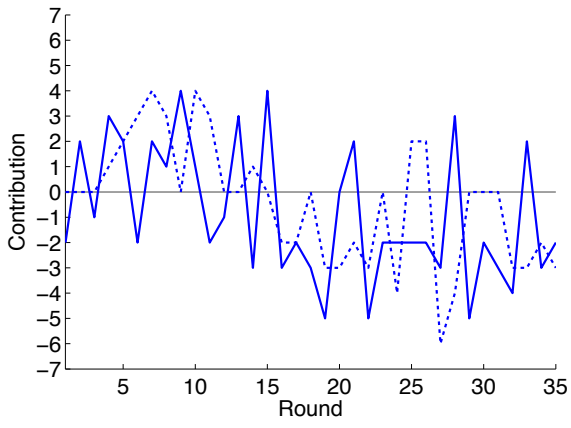
2489

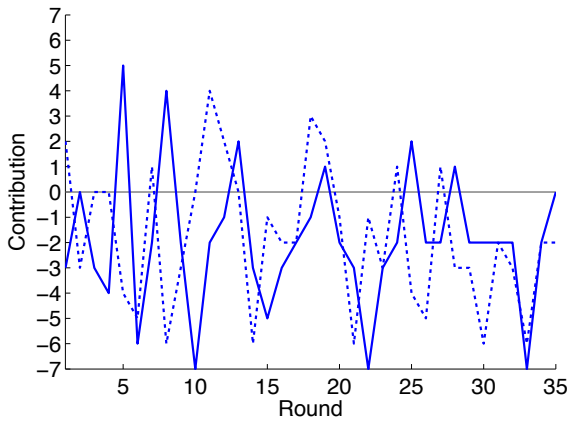


2490

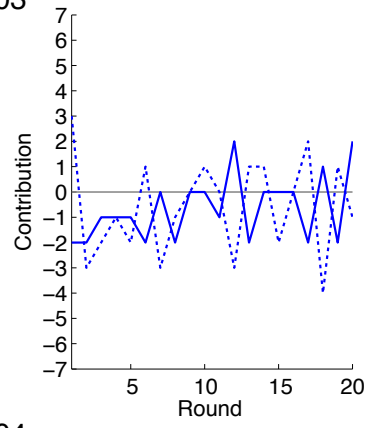




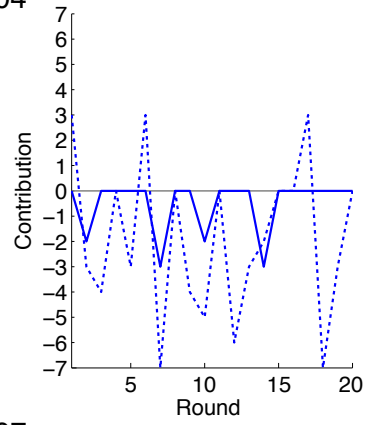
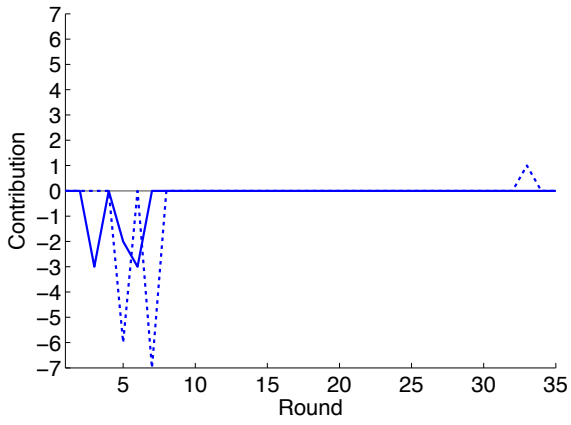




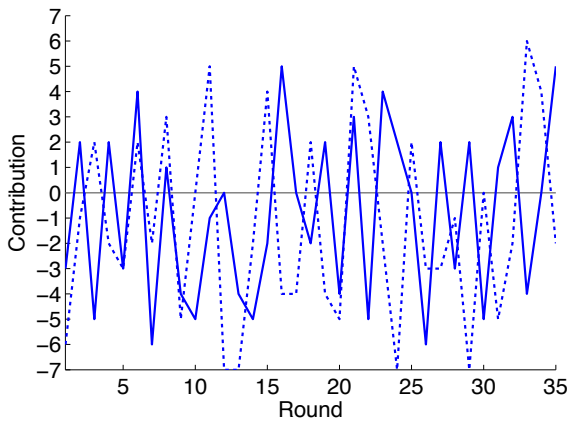
2603



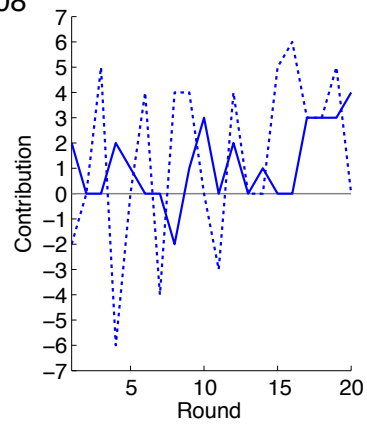
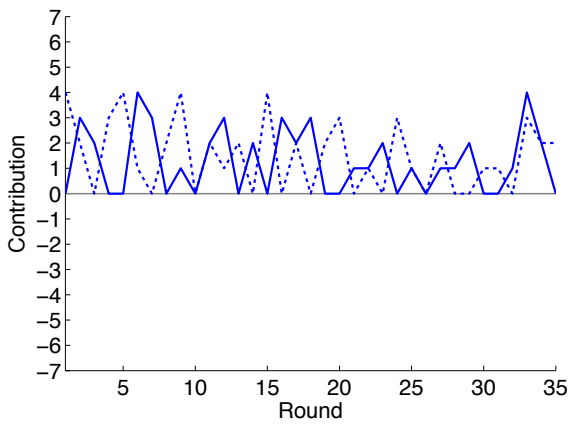
2604

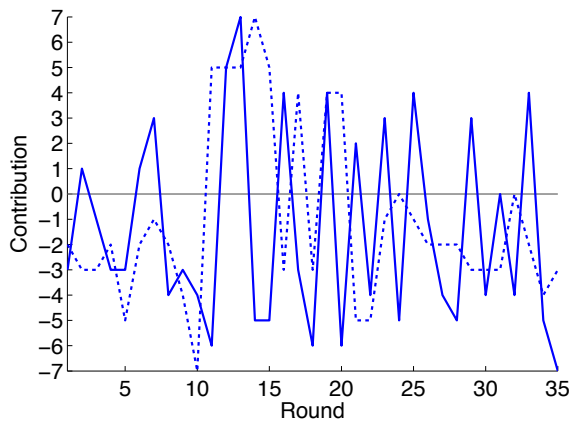


2607

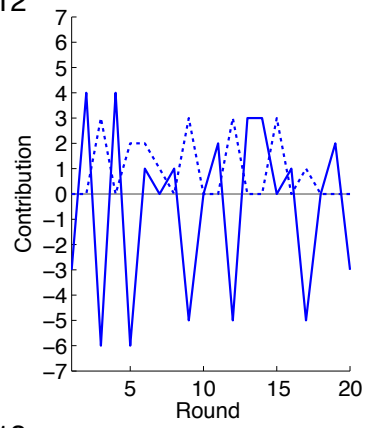


2608

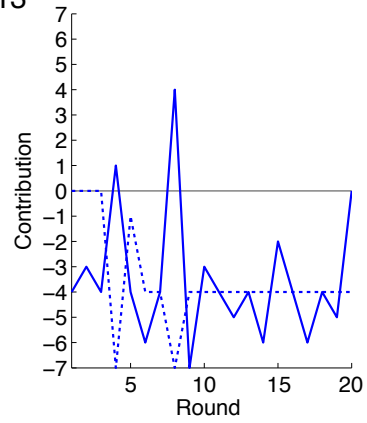
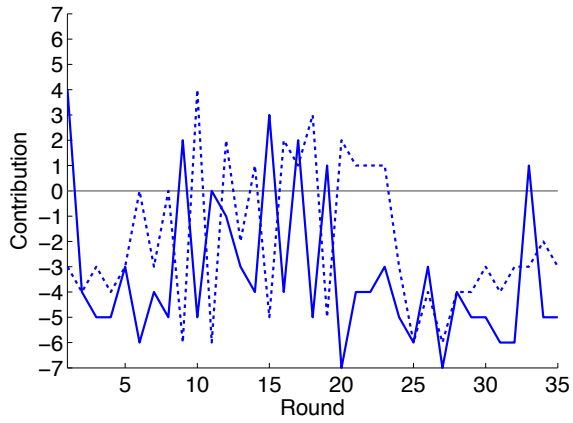




2612



2613



4.6.4 SVO and Contribution Correlations

| | ρ | p-value |
|-----------------|--------|---------|
| First Round | 0.0608 | 0.7729 |
| First 5 Rounds | 0.442 | 0.161 |
| Full First Game | 0.1457 | 0.4872 |

4.6.5 Additional fMRI data

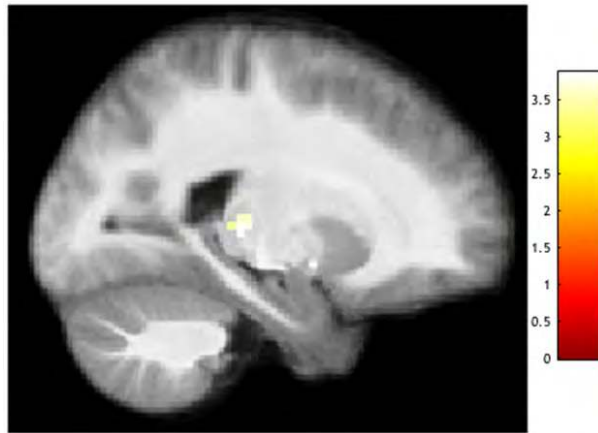


Figure 4.24: $X = -21$, uncorrected threshold $p = 0.001$

4.6.6 Questionnaires

Following is the questionnaire for scanned subjects. The questionnaires for non-scanned subjects merely differed in the numbering of the different parts of the experiment.

Questionnaire about the game

1. Did you have the feeling you were interacting with another person?

Not at all 1 2 3 4 5 6 7 very strongly

2. If you answered 4 or less (otherwise go to question 3), was it because:

Interacting through computers does not feel realistic to you

You had doubt you were seeing the actual choice of the other participant

Other: _____

3. How much do you like or dislike the participant you were paired with in part 2 and 3 of the experiment.

very unpleasant person 1 2 3 4 5 6 7 very nice person

4. How did you make your choice to take or contribute to the public account? What strategy did you use?

5. Assuming a given choice by the other participant, would the decision to take an additional token from the common account

Increase

Decrease

Not affect

your own payoff?

General questions

Birth Date:

Gender:

Place of birth:

Occupation/study background:

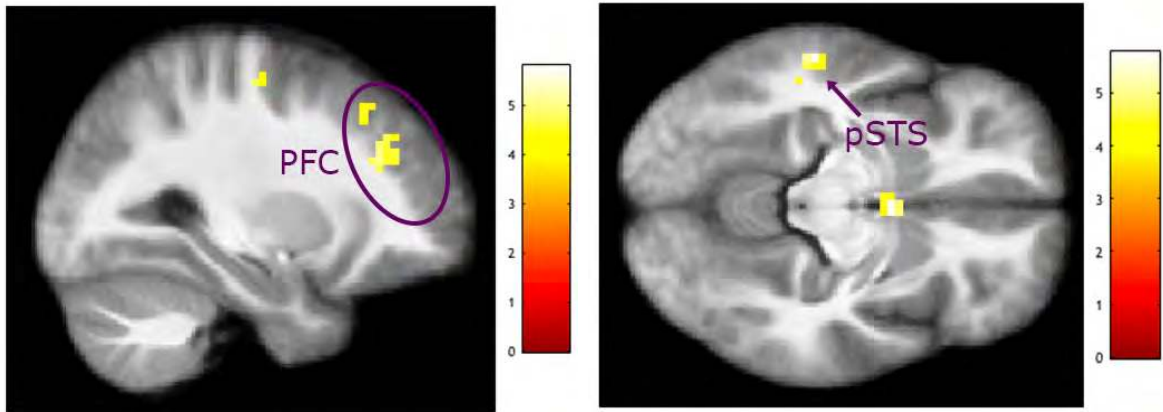
Level of study:

General Comments:

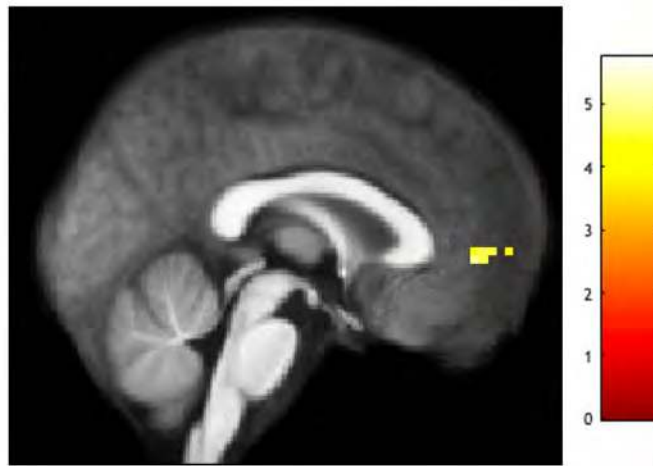
4.6.7 Subsample of subjects with high θ value

We selected a subset of subjects for which the estimate of the θ -parameter, the inverse of which indicates the degree of randomness in a subject's choices, was relatively high. Specifically, we chose the scanned participants for whom the estimated θ -value was larger or equal than the median value found¹⁵. Looking exclusively at this smaller sample, which contains 13 subjects, we get the following results: Contribution is still largely unresponsive, only lowering the voxel threshold to $p \leq 0.005$ uncorrected reveals a small cluster of positively correlated activity in the right insula (9 voxels, peak at 39, -4, 17), while some areas in the primary motor cortex appear to correlate negatively. SVO produces a similar negative contrast as before (peak at (-24, 26, 47), uncorrected p-value = 0.065, figure 4.25a). The major payoff cluster in the left pSTS stays intact (with two separately localized peaks at -57, -28, -10 and -63, -28, -4). In addition, the cluster that was previously identified as part of the ventral striatum now seems more specifically located at the hypothalamus (uncorrected p-value = 0.077, 20 voxels, peak at (3, 2, -10), figure 4.25b). Impulse correlates positively with a newly found cluster in the medial prefrontal cortex (uncorrected p-value = 0.047, 20 voxels, peak at (-2, 47, -1), figure 4.25c). The model parameters ($\alpha, \delta_1, \delta_{2P}, \delta_{2N}$), however, still do not correlate with activation in any areas of interest.

¹⁵This gave us the following list of subjects: sc2251, sc2260, sc2484, sc2485, sc2489, sc2490, sc2543, sc2544, sc2545, sc2598, sc2603, sc2604, sc2608.



(a) Negative parametric effect of SVO during (b) Contrast of receiving a higher/lower pay-decision. Sagittal view at $X = -28$. Uncorrected threshold $p = 0.001$.
 (b) Contrast of receiving a higher/lower pay-decision. Axial view at $Z = -10$. Uncorrected threshold $p = 0.001$.



(c) Parametric effect of impulse during display of other's decision. Sagittal view at $X = -1$. Uncorrected threshold $p = 0.001$

Figure 4.25: Selected subsample. T-map projected on averaged brain of all participants.

Chapter 5

Investors have feelings too¹

¹This chapter is based on work with Frans van Winden, see Hoyer and Van Winden (2016). For useful comments the authors thank participants of the 2015 ESA European and World Conferences, the 2015 FUR Conference, the 2015 CCC Meeting at the University of Norwich and seminars at the Tinbergen Institute and the University of Amsterdam. Financial support from the Research Priority Area Behavioral Economics of the University of Amsterdam is gratefully acknowledged.

5.1 Introduction

Relationships matter. This statement is true not only for everyday human interaction, but also when it comes to business. The experiment presented here was designed to shed further light on the role of relationships in one specific context: the interaction between an investor and a project manager, who can either be retained or replaced by a new project manager, following different experiences shared with that manager. In particular, we are focusing on the role that affect can have in this context. How affect is directed and how it develops are questions that have long been at the center of social psychology research, but have entered the field of experimental economics only fairly recently. We try to shed some more light on this issue in the specific context of an investment game, excluding the trust element that is often at the center of such games in an experimental context and focusing purely on social preferences.

Relationship banking is an important topic in microeconomics and finance (Boot, 2000) and has attracted attention in experimental economics (Brown and Zehnder, 2007; Cochard et al., 2004; Cornée et al., 2012). In both fields the focus has been on the strategic motives that come into play once an investor-borrower relationship extends through time. While it is of great relevance to answer questions such as the role of trust (Houser et al., 2010), regulation (Cornée et al., 2013; Brown and Serra-Garcia, 2014; Fehr and Zehnder, 2009; Lunawat, 2013), and reciprocity (Cochard et al., 2004), we believe that there is one more aspect that is part of such relationships. Any repeated interaction with another person may trigger emotional reactions in at least a subset of subjects. Inspired by experimental studies on affective relationships (van Dijk et al., 2002; Hoyer et al., 2014), our goal here is to contribute to the understanding of such non-strategic factors in the asymmetric context of the investor-borrower relationship with its inherent power imbalance. In further contrast to much of the research in the area of trust games, our focus is on the behavior of the investor, rather than the recipient of an investment. We are interested in the way that a personal relationship affects the decision making of the investor, and how it affects the way that the investor interprets information about the value of the investment op-

portunities as presented by the borrower, relative to the value of investing with another borrower. The individuals who make investment decisions within organizations are only human, hence it should be of relevance for any organization to understand what drives the behavior of those who make decisions in its name.

The specific setting studied in this experiment is a repeated investment decision, wherein subjects decide whether to proceed investing with project managers that they have been in contact with previously or to let them go and invest with a new project manager instead. There is anecdotal and empirical evidence that personal relationships play an instrumental role in banking. The following quote, taken from Uzzi (1999), who collected field data from lending officers, illustrates this notion: "After he [the entrepreneur] becomes a friend, you want to see your friend succeed and that goes along many lines. If I can be a part of helping them do that, it's a real good feeling and I'm providing a service not only to them but their employees.... So there's a lot of things that you kind of from a moral standpoint take into effect.... That is kind of a side effect of your relationship." However, there are also pitfalls in relationship banking, such as the hold-up problem and soft budget constraints, which can distort borrowers' incentives *ex post* if the lender finds herself forced to grant more credit just to preserve an earlier investment (Boot, 2000). We can point to more practical examples of situations in which an excessive focus on the relationship aspect of banking can have detrimental effects. A practical and dramatic example concerns the infamous Anglo Irish bank, the downfall of which contributed gravely to the struggles of the whole Irish economy during the financial crisis of the late 2000s, partly driven by the excessive interweaving of its fortunes with those of its lenders (Carswell, 2012). More to the point of this chapter, such relationships also have the potential to decrease the effectiveness of an investor to identify the most valuable projects to invest in. Relationship banking can hence have both positive and negative sides, which we will observe in our experiment.

When the early stage of a project provides negative signals about its value, investing more

money into a partner can become a question of loyalty². Stopping to provide financing to a borrower, firing an employee, and other self-interested acts that come at somebody else's cost are difficult, especially if we share a history with that person. It is human nature to feel responsibility for others and abandoning others goes against human nature. This is even more so as it is often difficult to evaluate which choice is going to be most profitable. In fact, when it comes to investment opportunities, especially in the context of venture capital, even repeated signals of bad performance do not necessarily imply that a company is not going to be successful. This is exemplified in the phenomenon of "pivoting" businesses, which change their business model before finding success³. At the same time, taking money away from an investment to put it into a completely new project means transferring it somewhere where there is even less available information. Because of the important role that uncertainty plays in the way that people make economic decisions, especially in a social context (Bosman and van Winden, 2010; Brock et al., 2013; Cappelen et al., 2013), the risk structure outlined in this paragraph is mimicked in the design of our experiment.

We contribute to the analysis of the nature of investment relationships by isolating the role of shared experiences with a project manager from the predictive power that such experiences might have for the future profitability of a project. Specifically, we investigate two different mechanisms: one in which project managers are given the opportunity to send or withhold a monetary transfer at the beginning of an interaction, and another in which investors merely experience success or failure of an independent additional project that is chosen by project managers at the beginning of the interaction. In both cases investors share an experience with a manager when they have to decide whether to stay with this manager or not: one that is positively charged (transfer or successful previous project) and one that is negatively charged (no transfer or failed previous project).

In the experiment the best response of an investor is not affected by the type of experience.

²It should be noted that our design restricts the action space of the investor to exiting the relationship, as opposed to voicing dissent with the quality of the project when loyal, using Hirschmann's terminology (Hirschman, 1970).

³Examples are Twitter (Carlson, 2011), Paypal (Penenberg, 2011), GroupOn (Penenberg, 2011) and BuzzFeed (Kafka, 2015)

What we are investigating is the question if they react to the different histories nonetheless, and if so, why. In the treatment that uses transfers, reciprocity provides a motivation for deviations from the pure best response. Moreover, this treatment shows some similarity to an experiment of Malmendier and Schmidt (2012), which finds such effects⁴. The impact of a shared history is not as obvious. However, there is a number of concepts that drive our hypothesis of a potential effect of such experiences. For example, investors could be driven to more positive reactions towards managers with whom they have shared positive experiences in the past on the basis of simply attaching a positive emotion to that interaction. Negative experiences could trigger the opposite reaction. Evidence suggests that even simple subliminal stimuli can cause liking or disliking, as demonstrated by mere exposure experiments (Zajonc, 2001)⁵. As we will see in the Design section, our managers' decisions essentially lead to random results, but investors might nevertheless attribute the success of the project to the manager's capability of selecting profitable projects. There is also evidence in experimental economics of an effect of unjust blame. Gurdal et al. (2013) show that principals routinely punish managers for events they had no influence on. Further arguments in support of our hypothesis of a "mere experience" effect will be discussed in section 5.2.

We find a strong difference in the investors' decisions after either having received a transfer or not and this reaction is significantly different from their behavior in a non-social control treatment without a project manager. We are not able to detect an increase in reaction to an experience shared with a manager relative to the non-social control treatment, despite the fact that post-experiment questionnaires indicate that a subset of investors reacted to the experience emotionally. Furthermore, decision times are similarly and significantly affected by the presence of a manager.

We begin with a literature survey in section 2. Section 3 presents the experimental design, together with an analysis of the investors' best responses and our hypotheses in section

⁴There are some notable differences between their experiment and ours; see section 2.

⁵There is also evidence that neurological processes related to preference ordering are activated when cues are not consciously recognizable (Pessiglione et al., 2008), and that subjects may unconsciously learn how to perform a task (Lebreton et al., 2009).

3. Results are presented in section 4, followed by a brief discussion in section 5. Section 6 concludes.

5.2 Literature

The relevance of relationships in the context of lending and borrowing has been recognized for a long time. In line with that, the term "relationship banking" has become a staple of the literature (Boot, 2000). Seeing how the act of providing credit to somebody else implies some expression of trust, this is hardly surprising: relationships can help facilitate trust-based interactions on a multitude of levels.

One element of relationship banking has only recently become actively researched: the creditors' preferences about whom they actually want to grant credit to, everything else being equal. Research on social distance (Goette et al., 2012) suggests that, if given a choice, people are much more cooperative towards people that they share social ties with. We are building on this idea.

Theoretically there is a strong connection between this chapter and social distance theory (Tajfel and Turner, 1979), insofar as one could look at the initial partner allocation in our experiment as related to the minimal group paradigm⁶. In a laboratory setting such minimal groups can lead to significantly more cooperative behavior in different environments, including investment situations. Examples within experimental economics can be found in Charness et al. (2007) and Chen and Li (2009). Akerlof and Kranton (2000, 2005) provide important arguments as to the role of identity in situations such as the one analyzed here, but differ to some extent in that the implicit focus lies on relatively low level members of an organization, such as employees, or, in our case, project managers. We explicitly focus on the role that identity plays for subjects that are better positioned in the hierarchy compared to those they interact with, as expressed through framing, their decision power, and outside options. We therefore look at a positive history as a source of a group identity that makes our investors look at their relationship with the relatively

⁶As we will see later, our investors actively choose partners, but do so in complete ignorance of who it is that they are choosing.

powerless managers as a team relationship.

There are a number of field studies that present evidence for the role of social relationships in the context of lending, especially using data from developing countries. Khwaja and Mian (2005) for example find that in a Pakistani sample firms with good political connections receive 45% bigger loans, even though they show 50% higher default rates. In a study that uses data from a highly developed country, Haselmann et al. (2014) find significantly positive effects for the influence of social proximity on lending in a German dataset of local banks. One of the measures used for social proximity is the shared service club membership of local bank board members and firm CEOs⁷.

The German example demonstrates that also in countries in which outright corruption is thought to be limited, social connections matter. We are not implying that such effects are necessarily indicative of corruption, or even merely inefficient: it is possible that the social connections that are being studied give banks better access to information that is vital in making an informed decision about which type of credit to grant. Even rather simple forms of relationships between a bank and a borrower can help reduce the default risk (Puri et al., 2013), implying that there is an objective value in these relationships⁸. Separating valuable information from favoritism is difficult to do in the field, opening up the potential for additional insight to be gained from a laboratory experiment.

Since we are interested in the relationship between a financier and a borrower, there is a connection to trust or investment games, which are widely used in the experimental economics literature, typically in some variation of the design of Berg et al. (1995). In these games an investor transfers a certain amount of money to a borrower, who has access to technology that can potentially increase the value of the investment (see Johnson and Mislin (2011) for an overview). This similarity is mostly superficial: the typical trust game largely focuses on the question of how the behavior of the borrower can be controlled or

⁷The relevance of social ties is by no means limited to the traditional banking sector. Duchin and Sosyura (2013) for example show that relationships are also highly relevant in the internal capital markets of companies. (Kuhnen, 2009) finds evidence of favoritism in manager choice in the mutual fund industry

⁸On the borrower side we see that even simple text message reminders can improve repayment if they include the name of the responsible loan officer, further showing that personal relationships can influence behavior on that end of the interaction (Karlan et al., 2015).

predicted by the lender. Investment games therefore have a second stage, in which the borrower can decide to keep or return the proceeds of the investment. Our focus in this project is on the affective reactions that drive an investor's decision making. In a repeated trust game it is difficult to disentangle the investor's desire to benefit a borrower with whom they have made positive experiences from their largely self-interested desire to invest in a manager who is more likely to repay the returns of a project. Our design in this experiment therefore shares much of the framing with trust games, but not the actual game design.

An example of an experiment that provides some separation between informative and emotional aspects of the investor/borrower relationship is presented by Brown and Zehnder (2007). The authors design a credit rating mechanism that provides information about a borrower's repayment history that is equivalent to information acquired in a treatment in which investors have previously been in a business relationship with the borrower. The main focus of their experiment is, however, on how that affects the borrowers' repayment discipline, rather than on potential effects on lenders, with the exception of hold-up strategies⁹. Their market mechanism is too complex to compare the investors' reaction to the relationship between treatments with and without the credit reporting facility. Similarly, Cornée et al. (2012) have treatments without repeated interaction, with repeated interaction, and with repeated interaction and additional information about the precise behavior of the borrower. However, also their design focuses on honest and dishonest behavior of the borrower, rather than the disentanglement of different motives of lenders. It can therefore not identify what role the relationship plays for them.

Clearly, the isolation of the emotional influence of bonding over shared success or failure is difficult if the behavior of the borrower has predictive power for the future income of the lender. For this reason we simplify the situation by creating a design that still has identifiable features of a repeated investment setting, but which eliminates this confounding factor. As we will see in section 5.3, our borrowers still make meaningful choices, but

⁹A lender can extract rents from a borrower by asking relatively high prices for the renewal of short term debt after a relationship has been established and provided the lender with exclusive positive information, because other lenders are not privy to the same information.

they are not predictive of future returns. Project managers select a project, which can either be more or less valuable, but they do not actually know which project it is. The project allocation is therefore essentially random, while still being a direct result of their decision.

Our experiment is framed as a investment game, but since it completely lacks an element of trust, it's design differs notably from most investment games. A comparison with a blame game (Gurdal et al., 2013) is more appropriate. In this game somebody is blamed for the outcome of a decision, even though there is no meaningful way in which it could have been known what the result of the decision was going to be. The results show that such behavior is a robust phenomenon not only in psychology, but also in a situation more in line with the methodology of experimental economics. Gurdal et al. (2013) mention two different potential drivers of such an effect: outcome bias (Baron and Hershey, 1988) and salient perturbations (Myerson, 1997). The former describes the effect that people rate the quality of someone else's decision making differently based on an uncertain outcome, even if the decision maker took all relevant information into account. Despite the fact that the result of a uncertain draw says nothing about the competency of the decision maker, it is rated higher after a positive outcome. In our so called History treatment project managers decide which project to implement without being able to distinguish more or less valuable projects, but the investors' evaluation of their decision quality might be affected by the outcome nonetheless. Salient perturbations, on the other hand, should not be of importance in this context. The concept describes the idea that agents interpret an unfamiliar situation in a manner that is more familiar to them. In our context this could imply that they assume the possibility of additional insight into the project quality as a function of some type of effort to be exuded by the manager. We think that our way of presentation prevents such mis-attributions, although it should be noted that this effect does not necessarily rely on a conscious misunderstanding of the situation, but merely on the situation being difficult to analyze and similar to a more familiar situation.

Yet another comparison can be drawn to experiments where different groups are constructed in the lab in order to analyze how the subjects' decisions in social games react

to that. Cason et al. (2015), for example, find noticeably higher rates of cooperation in an inter-group prisoner's dilemma game if the two groups played a successful minimum effort game together before engaging in the prisoners' dilemma game, as compared to a treatment where the minimum effort game was not present. This effect is reinforced by inter-group communication possibilities during the minimum effort game. Morita and Servátka (2013) find higher investment rates and lower rejection rates in a holdup problem if the first and second movers are from the same group than if they are from different groups. In their experiment subjects are assigned to groups using different shirts and group members perform a trivia task together prior to the holdup (trust) game.

We are further guided by a number of experiments on "affective ties", which are based on a model that treats agents' social preferences vis-à-vis other agents, with whom they interact repeatedly, as endogenous (van Dijk and van Winden, 1997). The utility specification in our predictions allows for this feature, albeit in a simpler way, since we merely have to track an investor's reaction to a binary state rather than a more complex space that is repeated over multiple periods. Nonetheless, results such as van Dijk et al. (2002) provide a foundation for our predictions, which are further supported by more recent experimental studies (Hoyer et al., 2014).

One distinction that should further be made is that our concept of a relationship based on a shared history clearly differs from the concept of intention based social preferences as it in Rabin (1993). Our project managers do not have any insight into the effect that their choice is going to have. Therefore, kind or unkind intentions cannot play any role in the investors' perception of the situation. This changes in the treatment in which project managers make an active choice to send a transfer or not. In this case the investor's response becomes much more similar to a response in a reciprocity setting (Fehr and Gächter, 2000). This treatment shares some features with an experiment by Malmendier and Schmidt (2012), who in their "gift" treatment give a manager ("producer" in their terminology) the option to send a gift prior to the investor ("decision maker") having to choose between investing with that manager or another manager. In this decision the projects associated with both managers are lotteries with different returns in case of a

success or failure¹⁰. One way of looking at our experiment is to see it as a combination of ideas that can be found in Malmendier and Schmidt (2012) and in Gurdal et al. (2013). In both papers the authors use a mechanism in which the choice that is disadvantageous to one partner benefits a previously unknown third party. In the first experiment the authors focus on favoritism and reciprocity as drivers for such behavior, while in the second the attribution of blame is investigated. Our design attempts to compare these mechanism¹¹.

5.3 Design and Hypotheses

The experiment consisted of three different treatments. Our main motivation was to isolate the role of different social experiences on the investors' project and manager choice in a stochastic environment. The three treatments are: a **History** treatment, in which the investor and the manager have experienced a success or failure together in a previous project; a **Transfer** treatment, in which a project's manager either sends a monetary transfer to the investor or not; and a **Control** treatment, which is similar to the History treatment, but does not include a manager, eliminating the social aspect completely.

5.3.1 Treatments

History

The History treatment has twice as many managers as investors. At the beginning of each round an investor chooses a manager from a pool of managers, who are presented in

¹⁰Major differences between our experiment and theirs are as follows. In their experiment investors ("decision makers") are assigned two potential managers ("producers"), whereas in our design managers are chosen endogenously from a pool. Furthermore, managers in our experiment are involved in two projects that affect the investor, and the projects are more complex to analyze due to the need to apply Bayesian updating to precisely calculate a best response. This difference in design also leads to a noticeably longer amount of time to pass between a transfer and the decision being made in our experiment. Finally, our investors decide between staying with a manager they interacted with before and switching to a new one, as opposed to deciding between two equally unknown managers. To explain their result, Malmendier and Schmidt model their decision makers behavior using a dynamic social preferences approach similar to what we do in this chapter.

¹¹Notable differences between our experiment and Gurdal et al. (2013) are the fact that their experiment clearly juxtapositions the result of an agent's choice with the hypothetical result had she chosen otherwise and again the more complex risk structure in our experiment.

the form of identical icons on a screen. The position of the icons is randomized in each round, so that the identities of the managers can not be tracked across rounds. The order in which investors make this choice is randomized anew for each round. Investors who have not yet made a choice and managers who have not yet been chosen see the screen with all icons until they have made a choice or have been chosen, respectively. The icons that represent managers who have already been chosen by an investor disappear from the screen one after another. Managers are also informed which icon they are represented by. Managers who are not chosen by any investor are redirected to a waiting screen¹².

Managers who have been chosen by an investor choose one out of eight potential projects. Each project either has a success probability of $\frac{1}{4}$ or $\frac{3}{4}$. Both types of projects are equally likely and neither investors nor managers can identify the projects at the time of choosing (i.e. their positions on the screen are randomized anew for every decision). The decisions are made in the same order as the choices of the investors, that is a manager who was chosen third is also the third to choose a project. Since all managers chose from the same set of projects, a manager who has been chosen by the final of 8 investors has only one project to choose from. The project choice screen works in the same way as the investor screen: randomly positioned projects disappear one after another once they are chosen and are no longer available to other managers. After a project has been chosen a manager is asked to "implement" it by clicking on a box that symbolizes the project. Both investor and manager see a 5 second long animation similar to the "processing" animation typically found on computers¹³, after which the success or failure of the project is announced.

After investor and manager have observed the result, the manager chooses a *second* project, which has no relation to the first project in any way. This implies that the success or failure of the first project provides no information at all about the success probability of any later project. The understanding of the last point was tested before the beginning of the experiment.

¹²To ensure attention inactive managers were given the possibility to watch a neutral video while they were inactive. We did not test possible behavioral effects of the video, but they would be irrelevant, since we do not analyze the managers' behavior.

¹³see the online demo (section 5.8.1) for an example.

After observing the outcome of this second project, investors are now given the choice to either stay with this project and project manager or to choose an alternative project manager and project. If the investor chooses the first option the manager is redirected to the implementation screen once more. After the implementation of the second project both parties are informed about the success or failure of the second implementation of the project. If s/he chooses to change managers the investor first has to wait until all investors have made their decision. Once that is the case all investors who opted to replace their managers are assigned a new random order and choose a new manager from the pool of managers who were not chosen to be managers at the beginning of the round. Newly chosen managers then choose a new project with the same blind procedure as before, which they subsequently implement. After all results have been observed the round ends. There is a total of eight rounds, which only differ in the payoffs of the alternative projects, as explained later in this section.

Transfer

The Transfer treatment follows the same general structure as the History treatment, with one difference. Whereas every round of the History treatment starts with a project that is completely unrelated to future projects, this part is now replaced. Instead managers who have been chosen by an investor are now given the option to transfer money to the investor or not. They are endowed with an extra 10 experimental currency (ECU) for this transfer. If a manager decides to make that transfer these 10 units are doubled and investor's earnings grow by 20 units¹⁴.

After deciding whether to transfer money or not, the manager chooses a project from a pool of 8 different projects using the same procedure as in the History treatment. It is then implemented in exactly the same way. After this project has been chosen and implemented, investors face the same decision as in the History treatment: to stay with the same project manager and project as before or to choose a new manager, who then

¹⁴The size of the transfer was chosen based on the observation that a transfer that was similarly sized relative to a project's expected earnings lead to reasonably evenly distributed decisions to transfer and not to transfer in Malmendier and Schmidt (2012).

chooses a new project.

Control

The Control treatment eliminates the social element that is present in the two other treatments. Investors now choose and implement their own projects instead of choosing a manager who then chooses and implements a project. Managers are not part of this treatment. Apart from that difference this treatment is exactly identical to the History treatment. Projects are chosen by the investors from a pool of eight projects in the same manner as in the other treatments.

Figure 5.1 illustrates the design of a single round in all three treatments.

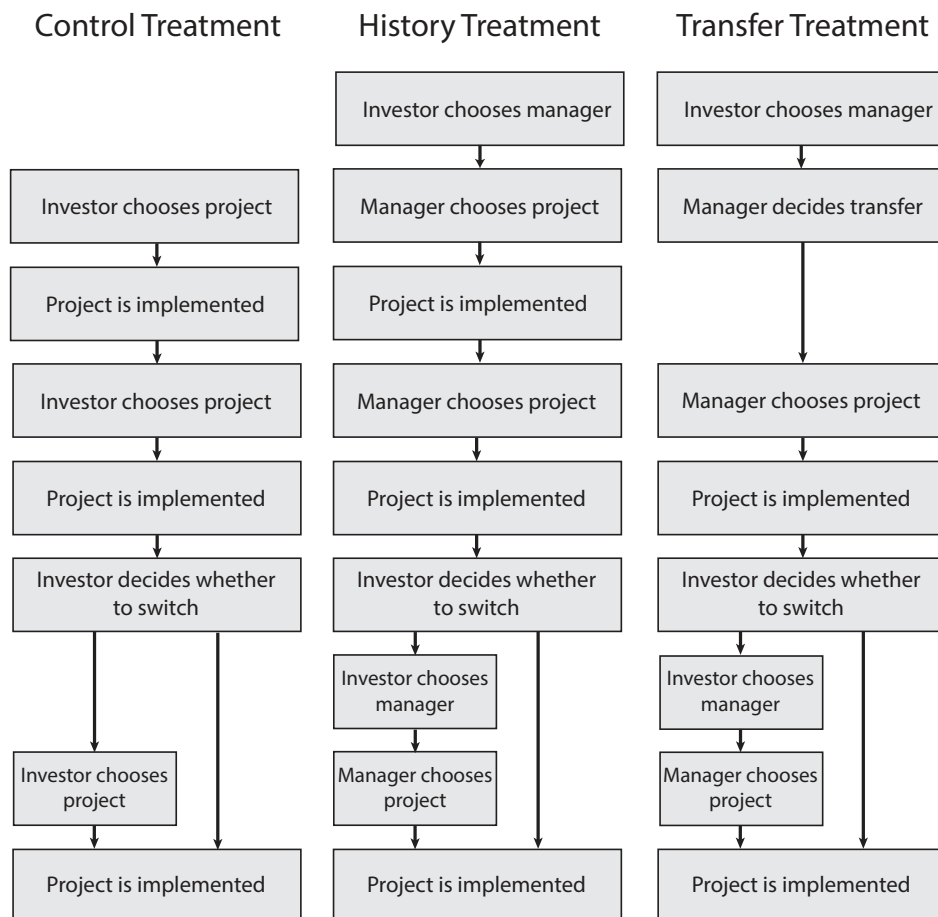


Figure 5.1: Design of Treatments

5.3.2 Projects

The following explains the earnings of investors and managers and the investor's best response.

A manager who is actively managing a project at a given time receives 200 experimental currency units irrespective of the project's success or failure. Managers who are inactive during the first project (Transfer treatment) or the first and second project (History treatment) also receive the same 200 units¹⁵. During the final project inactive managers receive nothing.

Ignoring all social aspects of this experiment for the moment, a profit maximizing investor must use past observations as a signal for the underlying success probability of the project in order to determine the best response.

In every round an investor can only choose one project. All projects either have a high ($p = \frac{3}{4}$) or a low ($p = \frac{1}{4}$) success probability. The ex-ante probability of both types of projects is 50%. With the exception of the alternative project that an investor can switch to at the end of a round, all projects generate earnings of 300 in case of a success and 100 in case of a failure. In order to precisely calculate the expected value of a project with unknown success probability and the investor's best response we therefore have to calculate the expected value of both types of projects and then combine them to get to the overall expected value:

$$E(\pi_H) = \frac{3}{4}300 + \frac{1}{4}100 = 250 \quad (5.1a)$$

$$E(\pi_L) = \frac{1}{4}300 + \frac{3}{4}100 = 150 \quad (5.1b)$$

where we use π_H and π_L for projects with known high or low success probabilities, respectively.

If the project in question is a completely new project (π) this implies an expected value of

$$E(\pi) = \frac{1}{2}E(\pi_H) + \frac{1}{2}E(\pi_L) = 200 \quad (5.2)$$

¹⁵We made this choice to eliminate inequity aversion as much as possible from the experiment.

The probability of observing the good outcome with payoff 300 is therefore $\frac{1}{2}$.

If however a project has been implemented in the previous period its success or failure provides information about its underlying success probability. Using Bayesian updating we can calculate the probability of the project being of the good type after having observed a successful draw:

$$P(\pi = \pi_H | success) = \frac{P(success | \pi_H)P(\pi_H)}{P(success)} = \frac{\frac{3}{4} \cdot \frac{1}{2}}{\frac{1}{2}} = \frac{3}{4} \quad (5.3)$$

Using the same procedure we get $P(\pi = \pi_L | success) = \frac{1}{4}$, $P(\pi = \pi_H | failure) = \frac{1}{4}$, and $P(\pi = \pi_L | failure) = \frac{3}{4}$. Combining equations (5.3) and (5.1) we can calculate the expected value of a project that was observed to succeed:

$$\begin{aligned} E(\pi | success) &= P(\pi = \pi_H | success)E(\pi_H) + P(\pi = \pi_L | success)E(\pi_L) \\ &= \frac{3}{4} \left(\frac{3}{4} \cdot 300 + \frac{1}{4} \cdot 100 \right) + \frac{1}{4} \left(\frac{1}{4} \cdot 300 + \frac{3}{4} \cdot 100 \right) \\ &= \frac{5}{8} \cdot 300 + \frac{3}{8} \cdot 100 = 225 \end{aligned} \quad (5.4)$$

Similarly we can calculate the expected value after observing a project to fail to be

$$E(\pi | failure) = \frac{3}{8} \cdot 300 + \frac{5}{8} \cdot 100 = 175 \quad (5.5)$$

Facing the decision whether to implement an old project again or choose a new one, a risk neutral selfish investor would therefore stay with a project that has been successful before (to earn $E(\pi | success) = 225$ in expectation) and choose a new manager with an unknown project if the first project implementation was a failure (to earn $E(\pi) = 200$ in expectation).

However, investors face a more complex situation. During the first (Transfer treatment) or first and second project (History and Control treatments) they earn 300 units in case of a success and 100 units in case of a failure. The alternative project has different returns, of which they are informed when they have to decide whether to stay with the original

manager and project or have a new manager choose a new project. For this reason the most convenient way of expressing the expected value of an original project is the more general

$$E(\pi^O|h) = P(\pi = \pi_H|h)E(\pi_H^O) + P(\pi = \pi_L|h)E(\pi_L^O) \quad (5.6a)$$

$$E(\pi^A) = \frac{1}{2}E(\pi_H^A) + \frac{1}{2}E(\pi_L^A) \quad (5.6b)$$

where π_H^O and π_L^O stand for the high and low success probability type of the original project and π_H^A and π_L^A for the high and low success probability type of the alternative project. h is a particular history of experiences.

The original and alternative project's returns are chosen such that they are either equal in their variance¹⁶ or their expected earnings or both. As illustrated in table 5.3 in the appendix, we offer three combinations of returns in which the alternative project has higher expected earnings, one with lower expected earnings, two with a lower variance and one with a higher variance. In five of the cases the alternative project either has a higher expected value or a lower standard deviation, therefore we take the alternative projects as the benchmark both in the appendix and the results section when describing differences in expected value and standard deviation. In order to get the most efficient experimental design possible we condition the alternative project returns that investors are offered on the success or failure of the previous project.

Calculating the optimal decision in the way outlined above is a task that is challenging and we do not expect participants to be very good at this part of the task¹⁷. In fact, there are reasons to think of it as even beneficial. One is the greater degree of realism that subjects face if they are not able to perfectly determine the value of the different options they are facing. Another reason is that situations which present a subject with a high cognitive load are understood to be more likely to trigger impulsive behavior from subjects (Duffy and Smith, 2014), in particular in situations that call for other-regarding

¹⁶That is, up to a negligible difference.

¹⁷In the instructions to the Control and History treatments subjects were told a second time that information from earlier draws could be used to estimate the success probability of a project, on top of merely outlining the design of the experiment. This was not the case in the Transfer treatment.

behavior (Cornelissen et al., 2011; Schulz et al., 2014).

Every investor faced each combination of returns exactly once and the order of the different combinations was randomized so as to ensure that the distribution of experienced orders was as flat as possible.

5.3.3 Presentation and Organization

Much of the experimental design was driven by the aim to provide an engaging experience for subjects, as the blind matching and project choice procedures are fairly impersonal. This was the main reason to implement the experiment with the computerized equivalent of a choice method in which subjects blindly choose cards that indicate their assigned managers and projects in turn. The act of choosing a partner should trigger a stronger engagement than if a partner had been assigned in purely random manner. A similar logic applies to the active project choice by the manager. We reinforced these effects by showing subjects that the pools of available managers and projects were constantly depleting and by using a design language that promotes the notion that projects are actually implemented, similar to the animations used in computer games to illustrate the execution of projects or tasks. The mechanic of choosing whether to stay with the project (and manager) or to choose anew was designed using a deliberately slow animation to reinforce the notion that this decision, which is our main outcome variable, is of relevance¹⁸.

The original instructions as they were presented to participants and an interactive example round of each treatment of the experiment can be accessed on http://www.mhoyer.com/inv_feelings.

The participants' understanding was checked using a quiz that covered the most important features of the experiment, including the concept that Bayesian updating can be performed in this setting. After the experiment subjects answered a short questionnaire covering demographic variables and some short questions about their emotional state during different situations in the experiment (see appendix 5.8.2).

The experiment was run in 12 sessions at the CREED laboratory of the University of

¹⁸The animation in question took 3 seconds.

Amsterdam in March and April of 2015. A total of 222 participants participated. Both the Transfer and the History treatment had 87 participants, of which a third (29) were investors. 48 participants were in the Control treatment, all of which were investors. At the end of each session a random round was chosen for payout. In the History and Control treatments no show-up fee was paid. The replacement of the first project with a relatively low-value transfer in the Transfer treatment required us to pay a show-up fee of 7 euros in the Transfer treatment to ensure satisfactory minimum earnings for participants. Sessions took approximately 70 minutes on average including instructions and payout and participants earned an average of 16.55 euros.

5.3.4 Hypotheses

For the reasons outlined in sections 5.1 and 5.2 we hypothesize that an investor's preference for the earnings of a project manager is stronger a) if the first project was a success relative to the situation in which it was a failure (History treatment), and b) if the manager sent a transfer relative to the situation in which it was withheld (Transfer treatment). From now on we will speak of a positive experience or a negative experience whenever we summarize the two different cases across treatments.

A simple formalization of this idea is to incorporate the manager's earnings into the investor's utility function and multiply it with a weight α which adjusts based on experience. We can then compare the expected utility of switching to the alternative project project (π^A , see (5.6a)) with the expected utility of staying with the original (π^O , see (5.6b)). Using a simple linear function and reformulating (5.6a) to the situation in the experiment, we get the following expected utility from choosing either of the two possible options:

$$E(U(\pi^O|h)) = E(\pi^O|h) + \alpha 200 \quad (5.7a)$$

$$E(U(\pi^A)) = E(\pi^A) \quad (5.7b)$$

Under the assumption that α is larger in case of a positive history compared to the opposite situation there are therefore more combinations of project payoffs for which $E(U(\pi^O|h))$

is bigger than $E(U(\pi^A))$ after a positive than after a negative experience. Therefore we expect a higher proportion of investors to stay with their original project and manager after positive than after negative experiences.

It should be noted that the second project in the History treatment - which finds it's equivalent in the first project of the Transfer treatment - can be expected to have similar effects as those we ascribe to the first project of the History treatment. They are however much more difficult to analyze due to being confounded with the calculation of the expected value of proceeding with the original project. Moreover it does not lend itself well to inter-treatment comparison in our design since we do not know how precisely the effect from the potential transfer interacts with an additional experience effect of a different type.

In this simple analysis we have ignored both models of inequality aversion and considerations concerning the income earned in the stages of the the experiment that precede the investor's decision. Canonical models of inequity aversion do not play a role in the investor's behavior, since at the time the decision is made all project managers have earned exactly the same amount. Whatever decision the investor makes, the number of managers who earned more or less than she and the size of the difference in earnings will not be affected¹⁹. Similarly, considerations such of efficiency do not affect our argument, since they coincide with the self-interested best response. Note that we do not model general social preferences regarding anybody but the first project manager (i.e. they have an α of zero). We think that this is natural due to the relatively big group of alternative managers. It would however not make a meaningful difference if we would take their income into account, because prior to the investor's decision all managers have had exactly identical earnings.

The direction of the assumed mechanism is the same in the History and in the Transfer treatment, the only difference is in the motivation. Whereas in History we assume that the investor is more concerned about the earnings of the original manager if they experienced

¹⁹More details and an extension covering the role of inequity aversion in a design without fixed manager earnings can be found in appendix 5.9.

success in the very first project, in Transfer the trigger is whether the manager chose to send the transfer or not, analogous Malmendier and Schmidt (2012)²⁰.

Hypothesis 5. *The probability of switching to the alternative project is lower in case of a positive experience than after a negative experience.*

As outlined in section 5.3, there are two potential problems in our design because subjects could be confused by the fact that one project - the first project in the control and history treatments - is not predictive of the success probability of future projects, whereas another project in fact is predictive. In addition, a positive experience could generally affect the subjects' emotional state regarding any familiar project, making them feel more positive about the original project, as opposed to the person who chose it. Behavior born from more general types of misunderstanding probabilities, such as the gambler's fallacy, add further potential problems. Without a method to control for these effects we would not be able to attribute the supposed result in hypothesis 1 to the assumed social effect of sharing a positive or negative history (or receiving a transfer or not). Therefore, in addition to the first hypothesis, we also require that the effect size of the different experiences is larger in History and Transfer than in Control.

Additionally, due to existing literature (such as Malmendier and Schmidt (2012)) we have a stronger prior for the existence of such an effect after a transfer than after a positive history. Based on the idea that a direct transfer, which may trigger reciprocity on the side of the investor, is a stronger intervention than merely sharing a common history we expect the overall effect to be strongest in the Transfer treatment.

Hypothesis 6. *The switching probability effect of different experiences follows the order Control < History < Transfer.*

²⁰It should be noted that inactive managers were not compensated for the transfer stage. Therefore active managers that had not sent the transfer had a slightly higher income level than inactive managers.

5.4 Results

Table 5.1 presents demographic data about the participants in the experiment and specifies the histories that the investors in the different treatments experienced prior to making their decision about staying with the same project (and manager) or not. We define a "positive" experience to describe either the experience of a successful project or a transfer, whereas a "negative" experience describes a failed project or not receiving a transfer. The distribution of positive and negative experiences in the Control treatment is perfectly balanced at 192 each by design, while in the History treatment the balance is not perfect because some sessions were run with only 18 or 21 instead of 24 participants due to low show-up, leading to a success rate of 48.7%. Experienced histories in the Transfer treatment are a function of the participants' decision making: Managers sent the voluntary transfer in 166 out 232 possible cases, a grand total of 71.6%. This is close enough to our optimal distribution of 50% to allow us to make statements about the reaction of investors to either receiving the transfer or not²¹.

| | N | Age | Female | Economics Students | Experienced Histories | | | |
|--------------------------|-----|-------|--------------|--------------------|-----------------------|--------------|--------------|--------------|
| | | | | | -/- | -/+ | +/- | +/+ |
| Control Treatment | 48 | 22.65 | 24 (50%) | 31 (64.6%) | 83 (21.6%) | 109 (28.4%) | 109 (28.4%) | 83 (21.6%) |
| History Treatment | 87 | 22.07 | 60 (69%) | 57 (65.5%) | | | | |
| History, Investors only | 29 | 22.1 | 21 (72.4%) | 14 (48.3%) | 61 (26.3%) | 58 (25%) | 56 (24.1%) | 57 (24.6%) |
| Transfer Treatment | 87 | 22.26 | 48 (55.2%) | 70 (80.5%) | | | | |
| Transfer, Investors only | 29 | 22.76 | 16 (55.2%) | 24 (82.8%) | 34 (14.7%) | 32 (13.8%) | 80 (34.5%) | 86 (37.1%) |
| Total | 222 | 22.27 | 132 (59.46%) | 158 (71.17%) | 178 (20.99%) | 199 (23.47%) | 245 (26.65%) | 226 (26.65%) |

A "+" indicates either a successful project or a transfer, a "-" indicates a failed project or the absence of a transfer. One manager's age was ignored due to obvious misreporting.

Table 5.1: Demographic Data and Experienced Histories

There is no notable change in the investors' decision across the 8 rounds of the experiment (figure 5.2)²². In the Transfer treatment there appears to be a slight increase in the

²¹The hypothesis of equal transfer ratios in all rounds is rejected with 0.04% significance due to one outlier in round 3, where 90% of all transfer were sent. Excluding that round the hypothesis cannot be rejected (Chi-square). Regressing the transfer decision on a trend in a random effects model produces a significantly negative coefficient at the 5%-level (see figure 5.7 in the appendix).

²²The null hypothesis of equal project switching rates in different rounds cannot be rejected (p=0.66) and there is no discernible trend.

second half of the experiment²³.

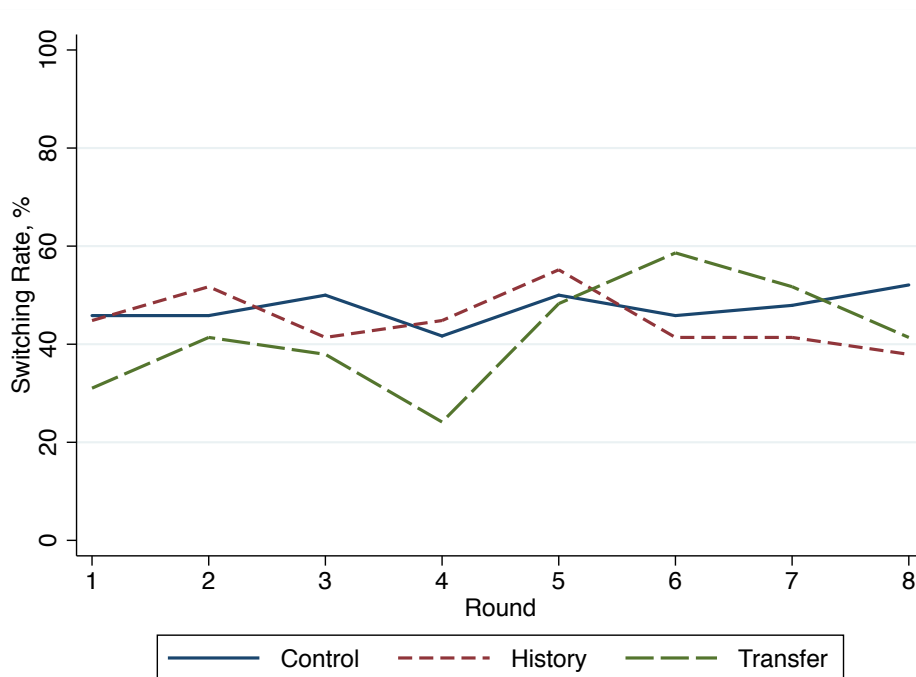


Figure 5.2: Switching Rate across Rounds

We begin our investigation into the investor behavior with a simple question: Does the experience made at the very beginning of a round matter? Figure 5.3a shows the proportions of investors that decided to choose a new project (and manager) after a positive experience in the different treatments. Recall that a perfectly selfish investor with a perfect ability to perform Bayesian updating would switch in 62.5% to 75% of all cases, irrespective of the experience or treatment. We observe a switching rate of 53.6% in case of a negative experience and 38.4% in case of a positive experience and the difference is highly significant with a chi-square test statistic of 16.1²⁴. This first result confirms hypothesis 1:

Result 1. *A positive experience leads to a significant drop in switching rates relative to a negative experience.*

²³ $p=0.069$ in a regression of only the trend and a constant in a random effects model.

²⁴We use a clustered chi-square procedure (Stata package `clchi2`). Here and later we cluster at the subject level. While the subjects interact indirectly, we argue that there is no possible channel for behavioral spillover within a group of investors, allowing us to treat different investors as independent. We also ran a test on only the first round as a robustness check, but results are only reported if they differ qualitatively using common significance criteria.

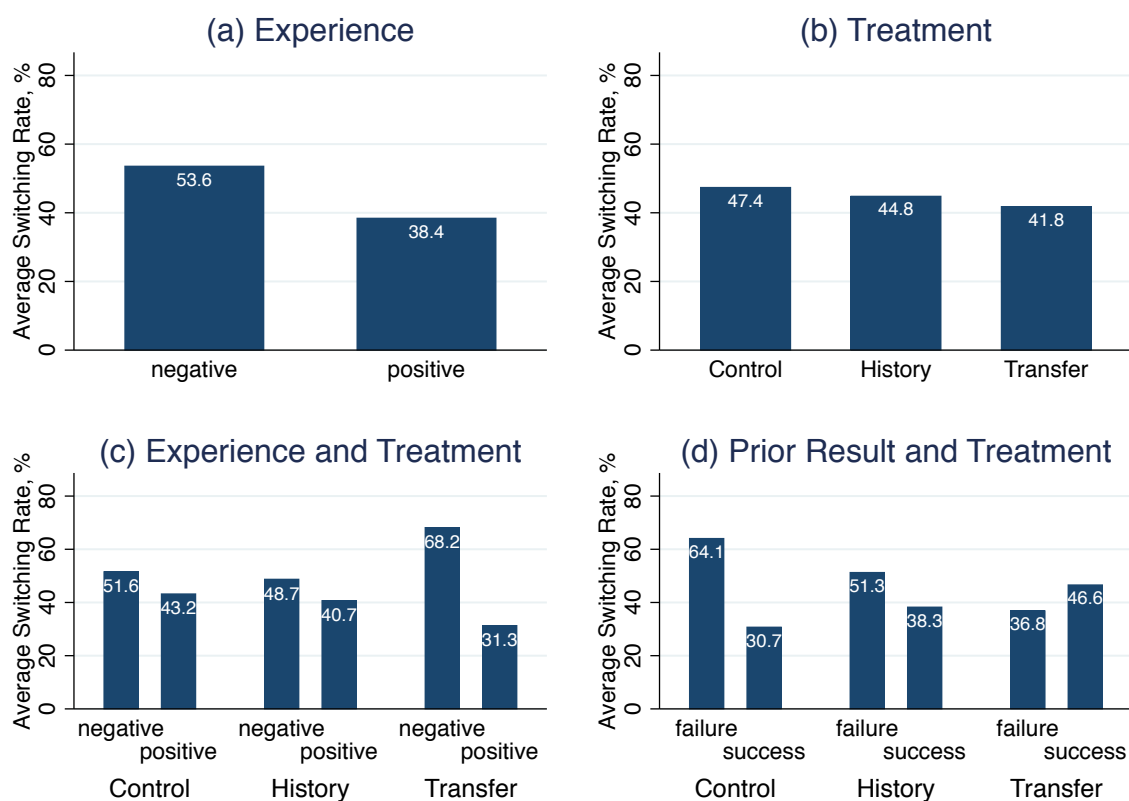


Figure 5.3: Switching Rates

Next we look at differences in the switching rate between the different treatments. Figure 5.3b shows overall switching rates in the three treatments. We observe a constant decrease going from the Control to the History and the Transfer treatment. The differences are not significant, however²⁵.

The natural next step is to focus on the difference in switching ratios relative to the different types of experience in the separate treatments, see figure 5.3c. While the difference in switching rates is substantial in the Transfer treatment (36.9%), the difference in the History treatment (8%) is not only nigh-identical to the Control treatment (8.3%), but even slightly smaller. The only treatment in which the investors' behavior differs significantly between experiences is found in the Transfer treatment.

Another dimension along which we can separate the investors' decisions is the result of the project that was implemented just prior to the decision, which we refer to as the "prior

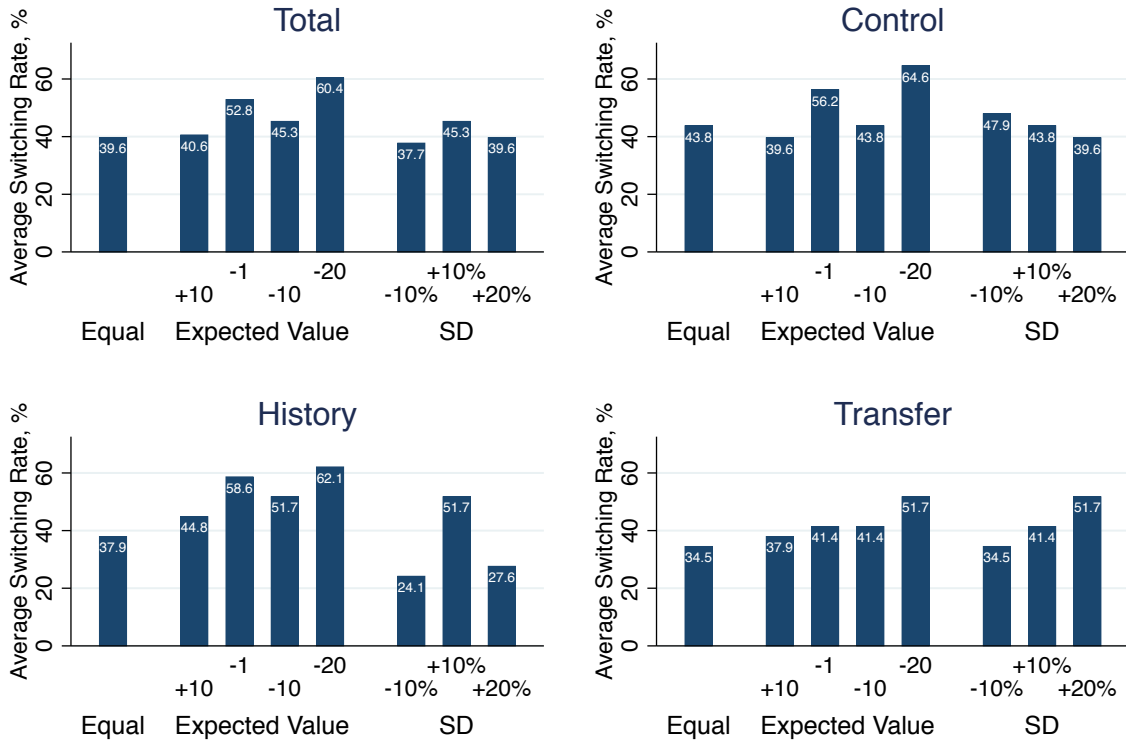
²⁵The lowest p-value is found comparing the Control and Transfer treatments at $p = 0.207$.

project” from now on. While the value of the alternative project was adjusted to the expected value of the original project as it could be calculated using Bayesian updating, we might still expect a positive experience effect relative to the result of this project. This is however not what we find, as the difference decreases between Control and History and even reverses in Transfer (Fig 5.3d).

The ability of the subjects to correctly perform Bayesian updating is not at the core of this analysis and not necessary for the interpretation of the other results. However, investors have a monetary incentive to switch projects more often if it is relatively beneficial to do so. Figure 5.4 distinguishes the different alternatives that investors faced in the experiment. Generally speaking, there seems to be a discernible effect when comparing the most extreme cases of positive or negative differences in expected value (19.8%, $p < 0.01$)²⁶. However, we do not see the monotonic increase in switching rates with increasing differences in expected value that one would expect. The same is true for the projects with different variances, where we would expect increasing switching rates the lower the variance of the alternative project becomes.

²⁶In this case the data were insufficient to run a meaningful test using only the first round.

Project Switching by Dilemma Type



Original and alternative projects either have the same expected value and variance or differ in one of the two dimensions. The labels used here refer to the situation in which the respective values for the original project differ from the benchmark alternative project as indicated in the label, the other dimension is always identical between projects. For example, in the case of expected value -20 the original project has an expected value that is 20 units lower than the alternative project, implying that switching is the best response for a purely self-interested investor. Expected value differences are in absolute values and differences in standard deviation are in relative values, rounded to full percentage points.

Figure 5.4: Project Switching by Dilemma Type

So far we have only compared the investors' behavior relative to their different experiences within the three treatments. In order to answer hypothesis 2 we need to go one step further. We hypothesized that the *difference* between the investors' switching rates after a positive experience and after a negative experience should be smallest in the Control treatment, and largest in the Transfer treatment. A visual inspection of figure 5.3c suggests that the Transfer treatment indeed has the greatest difference, but it seems unlikely that the first part of the hypothesis, which concerns the difference between Control and History, holds. To come to a more conclusive statement we construct a number of different panel regressions, in which we interact the treatments with a dummy variable

for the experience (Table 5.2). Irrespective of the specification, the results fall in line with the first impression from figure 5.3c. The coefficient of the interaction term between the Transfer treatment and the experience dummy is always significant at the 1%-level, while the coefficient of the interaction term between the History treatment and the experience dummy is positive but not significant. This impression is confirmed by running chi-square tests over the differences in ratios predicted by the logit coefficients in the different treatments²⁷. We also predicted the switching probability using only positive experiences, to prevent potential overlap with inequality aversion motives in the Transfer treatment. The difference in predicted switching rates between Control and Transfer was negative and marginally significant (-9.4% , $p = 0.08$, specification (4)). Running a regression similar to specification (4), but using History as the baseline confirms that there is no difference in the differences in Control and History, but the Transfer dummy shows a significant interaction with the experience dummy. In conclusion, we can only partly confirm hypothesis 2:

Result 2. *Switching rates after different experiences are not significantly different between the Control and History treatments. In the Transfer treatment the difference is significantly larger than in the Control treatment and the History treatment.*

Using the whole sample we also see that the result of the prior project affects switching rates negatively (a positive outcome leads to a lower switching rate). The difference between the expected values of the original and the alternative projects affects switching in the expected direction (if the original project has a relatively high expected value we predict less switching), whereas standard deviation differences do not show a significant effect, although also pointing in the expected direction (original projects with a relatively high future variance lead to more switching)²⁸.

²⁷Predicted between treatment change in the difference of switching ratios relative to experience, keeping all other variables at their mean and using specification (4) from table 5.2: Control vs History 3.9%, $p = 0.64$; Control vs Transfer: 23.5%, $p < 0.01$.

²⁸See table 5.4 in the appendix for the same regression using a probit model. Results are qualitatively comparable.

| | Investor switches project | | | |
|---------------------------------------|---------------------------|----------------------|----------------------|----------------------|
| | (1) | (2) | (3) | (4) |
| History | -0.119 (-0.48) | | -0.190 (-0.74) | -0.263 (-1.00) |
| Transfer | 0.699* (2.22) | | 0.630 (1.95) | 0.623 (1.91) |
| Positive Experience | -0.348 (-1.66) | -0.707*** (-4.73) | -0.497* (-4.62) | -0.499* (-2.30) |
| History \times Positive Experience | 0.0160 (0.05) | | 0.139 (0.40) | 0.163 (0.46) |
| Transfer \times Positive Experience | -1.214** (-3.18) | | -1.075** (-2.75) | -1.047** (-2.67) |
| Prior Result Positive | | -0.708*** (-4.81) | -0.699*** (-4.71) | -0.723*** (-4.83) |
| Expected Value Difference | | -0.0282** (-3.18) | -0.0269** (-3.01) | -0.0258** (-2.87) |
| SD Difference | | 0.00437 (0.30) | 0.00329 (0.37) | 0.00294 (0.33) |
| Round | | | | 0.0196 (0.61) |
| Female | | | | 0.0662 (0.40) |
| Age | | | | 0.00825 (0.29) |
| Economics Student | | | | -0.206 (-1.16) |
| Choice number | | | | -0.0526 (-1.56) |
| Constant | 0.0664 (0.43) | 0.452** (3.21) | 0.408* (2.26) | 0.486 (0.65) |
| Individuals | 106 | 106 | 106 | 106 |
| N | 848 | 848 | 848 | 848 |

Random effects model with z -statistics in parentheses, using robust standard errors

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.2: Investor decision regressions, logit

Next we investigate the behavior of the investors as expressed in their decision times. We observe that decision times in the two social treatments are significantly longer than in the Control treatment (Figure 5.5)²⁹. At the same time the difference in decision times

²⁹Using clustered t-test, both Control vs History and Control vs Transfer have p-values below 0.001.

between History and Transfer is negligible at 0.5 seconds. This result seem to be driven by a smaller proportion of investors who made their decision very quickly (see density estimate in figure 5.8 in the appendix) and is stable throughout all rounds (see figure 5.9 in the appendix). Again we repeat this analysis by using a panel regression, which confirms our findings (Table 5.5 in the appendix). Interestingly, we find no indication that decision speeds are significantly correlated with either the decision made by the investor or the absolute difference in expected value or variance between the two projects³⁰.

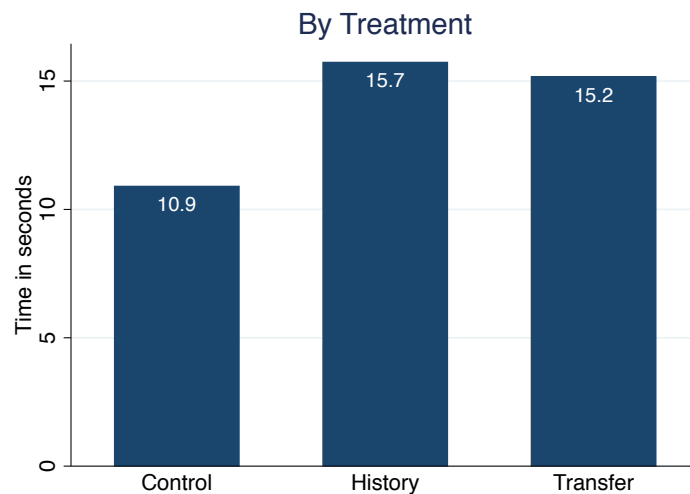


Figure 5.5: Investor decision times, in seconds

As part of our post-experiment questionnaire we asked investors a set of questions regarding their reaction to either having received a transfer or the first project having been a success, dependent on the treatment³¹. Figure 5.6 shows the distribution of their an-

As was done with investor decisions, decision time comparisons were also ran on only the first round as a robustness check (here using a simple logit instead of a random effects panel model), but results are only reported if they differ qualitatively using common significance criteria. Note that all decision times include the 3 seconds that a investor has to wait as part of the confirmation screen, plus additional waiting time if they decided to change their decision before confirming. In History and Transfer subjects saw an additional reminder of the effect a decision has on the managers, but that was identical in all 8 rounds and hence unlikely to be relevant for this comparison.

³⁰In the most complete specification (4) the coefficients for the absolute difference in expected value and variance both point in the direction that implies that decisions in situations in which the difference is larger are made quicker, but fail to achieve significance. Note, though, that in both social treatments with a positive experience decisions to stay with the current manager are made slower than decisions to switch, a relationship that completely reverses with a negative experience, but only in the social treatments (Control treatment: -0.2 to -0.2 seconds; History treatment: from -2.6 to +2.1 seconds; Transfer treatment: from -2 to +4.8 seconds, see figure 5.10). The differences are weakly significant in a regression using an experience/investor decision dummy when we pool the social treatments, but not when analyzed in any treatment in isolation.

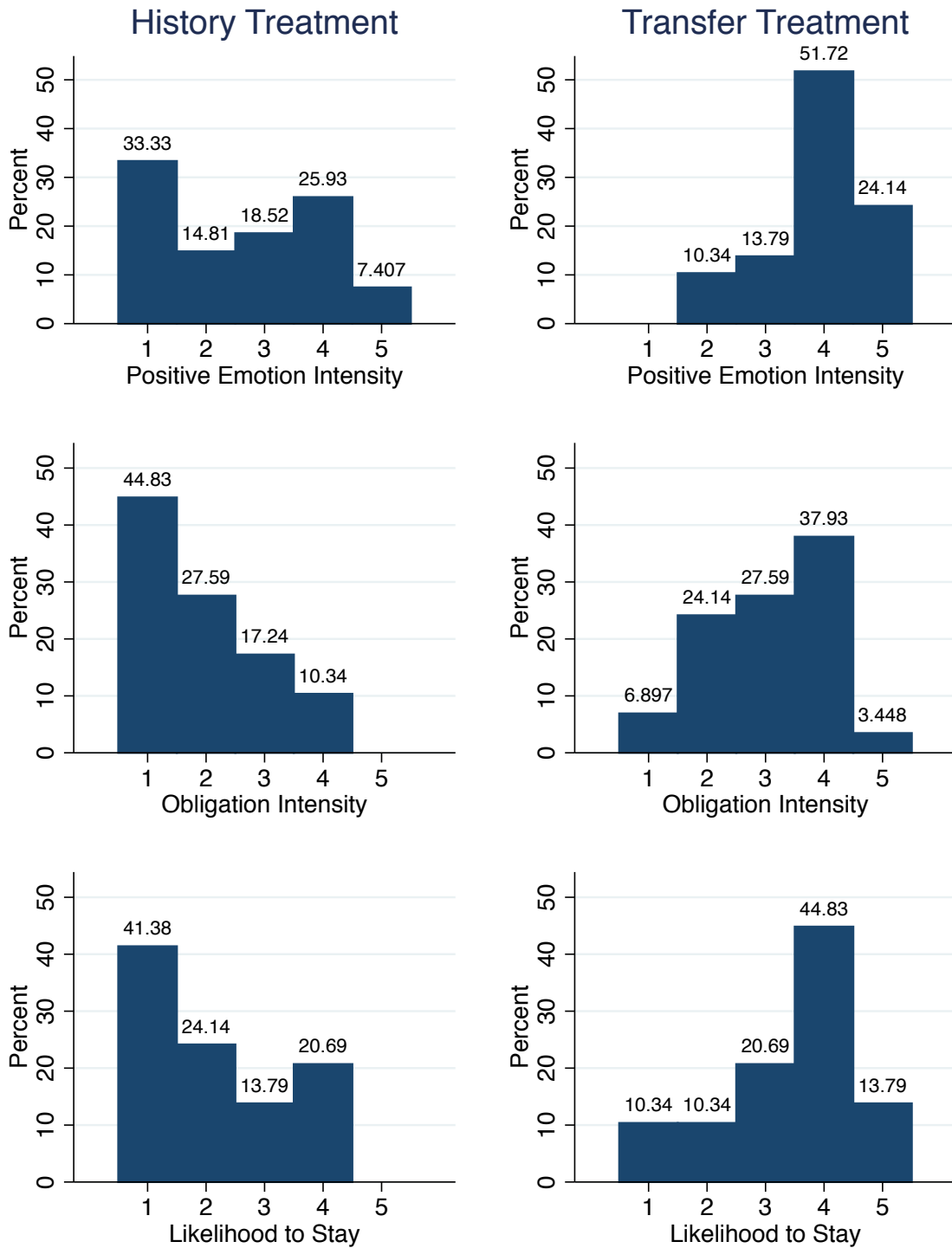
³¹See section 5.8.2 in the appendix for the exact questions.

swers in the two different treatments. Asked if they felt a positive emotion towards a manager with whom they had shared a positive experience, the distribution of answers in the History treatment is bimodal with 33% choosing 1, the lowest possible choice on our 5-point scale. In the Transfer treatment 0% chose 1 and 75% chose a value of 4 or higher. This picture is confirmed by the questions asking for a sense of obligation towards that manager and a direct question concerning their likelihood to stay with such a manager. There is always a lot more mass on the upper part of distribution in the Transfer treatment compared to the History treatment³². The investors' answers to the emotion and "likely to stay" questions are predictive of their reaction to their experiences, with the obligation question also being marginally significant (see table 5.7 in the appendix³³). Correlations between the questions regarding the feeling of positive emotion and feeling of obligation are between 0.35 and 0.58 and always at least weakly significant, but among the other possible combinations only the Transfer treatment emotion and "likelihood to stay" correlation is significant without pooling the treatments³⁴.

³²In all questions average answers are significantly higher in the Transfer treatment with p-values below 0.001.

³³The interaction for the emotion and the likelihood to stay variables with experience are significant at the 5%-level, while the obligation variable just misses that mark. Running the same regressions separately in the History and Transfer treatment mostly results in results too weak to make any conclusive statements about the directionality of the effect, with the exception of the "likelihood to stay" variable, the coefficient of which is negative and significant with a p-value of 2.9% when interacted with experience in the Transfer treatment.

³⁴see table 5.6 in the appendix.



Two subjects answered they had not experienced a success in the first project when asked for their emotion rating, leaving 27 observations. In all other cases we have answers from all 29 investors in both treatments.

Figure 5.6: Questionnaire: investor scores on emotion, obligation, and likelihood to stay with a project manager after a positive experience

5.5 Discussion

Based on the results presented in the previous section, one result is clearly established: We observe a much lower rate of switching to alternative projects and managers after a transfer has been sent than after the absence of a transfer. The predicted size of the effect reaches almost 37%, despite the fact that investors only face the decision whether they want to stay with the same manager or not after an intermediate project. Not only does this project come with the cognitively strenuous task of having to evaluate the relative value of the two options, it also introduces a non-negligible amount of time that passes between the transfer and the decision. This makes our study a much more demanding test of the impact of direct transfers than the one that was implemented in Malmendier and Schmidt (2012). Moreover, our design further differs from that experiment by having an investor decide between staying with a manager or switch to a new one, as opposed to choosing between two completely unknown managers ("decision makers"). In our case the decision is one about an ongoing relationship, rather than a simple binary choice between two otherwise identical partners. The comparison with a non-social, yet otherwise comparable treatment makes for another difference.

Moving to the History treatment we observe more mixed results. A majority of investors (14/26) reports to have felt a relatively positive emotion towards a manager after a successful first project³⁵. Reported scores for a sense of obligation and a likelihood to stay with the manager were lower, but still 8 and 10 investors reported relatively high scores, respectively. Moreover, the decision times in this treatment are almost identical to the Transfer treatment, while decision times in the Control treatment are significantly shorter. Despite the self-reported emotional reaction to the experience shared with the original manager and the apparently more time demanding complexity of the decision compared to the non-social setting we cannot detect a noticeable effect on the investors' incentivized decision. We are therefore forced to conclude that the hypothesized effect of being more willing to forgo future earnings in exchange for the ability to benefit the earnings of a

³⁵Defined as reporting three or higher on the emotion intensity scale.

manager after a positive history than after a negative history was not observable in this experiment. There are a number of potential reasons for this.

First, it is possible that the situation experienced in the laboratory was too artificial to trigger the kind of social reaction that we were interested in. This explanation can go in two different directions: Either investors genuinely did not care about the fate of their managers, or they cared only to the extent that it did not hurt themselves. We feel rather confident that the first explanation can be rejected on the basis of the significantly higher decision times observed in the History treatment as compared to the Control treatment. The fact that decision times were nigh-identical in the History and Transfer treatments would seem to imply that the decision presented investors with a similar degree of complexity. Combined with the observation that the Transfer treatment did show a significantly larger effect of the transfer decision, this leads us to believe that subjects may have perceived some kind of conflict between their relationship induced preferences and their own self-interest.

A second explanation could be that investors were not fully aware of the fact that both active and inactive managers earned the same in the first two rounds of the History treatment, and also forgot that the managers' earnings are independent of the projects' success or failure. If some investors thought that inactive managers earned nothing or less than active managers that could have given them an additional motive to switch projects after a shared success, an effect that would be weaker after a shared failure. A standard inequity aversion utility function (Fehr and Schmidt, 1999), where we define all participants rather than only the matched manager as the relevant group, predicts investors would switch away from losing projects more often in order to equalize earnings (see appendix 5.9.2). This effect might cancel out some of the hypothesized decrease in switching away after a shared success. However, for this to be relevant a subject would have to misunderstand or forget two carefully explained features of the experiment, one of which - the fact that manager earnings are independent of success or failure - was tested in a post-instruction quiz. Note, furthermore, that if the subjects had only forgotten that inactive managers also receive income during the first two projects that would have

affected switching rates after both histories equally.

In fact, the fixed earnings of managers might also have counteracted the detectability of the hypothesized effect. It is conceivable that the investors' perception of a bond or group identity with the manager that they interacted with might have been stronger if the managers had also been exposed to payoff uncertainty, as opposed to merely choosing a project and observing the result. There is suggestive evidence that the experience of a common threat can drive attachment (Carter, 1998), whereas our design explicitly exposes only the investor to the threat of losing.

To put the impact of different types of experiences into perspective we note that their effect on the investor's earnings was much less dramatic in the Transfer treatment, compared to the other two treatments. A success of the first game implied a profit of 300 ECU, 200 ECU more than a failure of that same project. In comparison to this the size of the transfer, 20 ECU, was negligible. This was a deliberate design choice, intended to make the transfer a primarily symbolic act. Nonetheless, it makes the clear result found for the Transfer treatment all the more notable.

The subjects' understanding of the relative values of the different projects presented to them was at best tenuous, as is demonstrated in figure 5.4. At least in the Control treatment one would expect a dramatic difference in switching rates between the situation in which switching is advantageous compared to when it is disadvantageous. While we had expected a somewhat better performance on the aggregate level, this is not in and of itself a problem for the investigation of our group level results, as the dilemmas that subjects are presented with are identical across treatments. In fact, as indicated in the design section, we intentionally chose a rather difficult stochastic environment under the assumption that that would give us a better chance of finding the hypothesized effects. There is little reason to believe that the intensity of the investor's loyalty towards a manager is weakened by the complexity of the situation that she is facing. If that is true it should be of relatively greater relevance in the decision making process that leads to the ultimate behavior if the monetary value of the different alternatives is relatively difficult to determine. This argument is rooted in the idea of cognitive efficiency: The analysis of

a relatively complex situation is cognitively demanding, therefore other decision factors can become more important with greater complexity.

5.6 Conclusion

We have shown conclusively that a project manager's decision to send a voluntary transfer to an investor at the beginning of a relationship changes the investor's decision whether to stay with that manager or switch to another manager at a later stage. This is true despite the fact that our design introduces additional steps in between the transfer and the investor's decision. However, we do not detect a similar effect for a merely shared positive history related to the outcome of a project. The relative difference in the likelihood to stay with or switch is not noticeably different from a control treatment in which the investor also takes on the role of the manager. Our hypothesis that positive experiences would facilitate bonding in a stochastic setting is therefore not substantiated without qualification by our participants' behavior.

Other measures such as decision time and the post-experiment questionnaires suggest that subjects did take the fate of the project manager into account in their decision making even when transfers were not possible. This factor is apparently not sufficiently relevant to sway their actual behavior significantly. It seems safe to assume that if such an effect exists its impact may be moderated by the saliency of the experience shared between investor and manager. We cannot exclude, therefore, that our design just failed to generate a sufficiently strong bond or group identity. The experience of a random draw from a relatively small lottery is clearly not as involving an experience as the type of long term business relationships we were inspired by. In any case, the effect appears to be a lot smaller than the effect that is triggered by the decision whether to send a transfer or not.

5.7 Appendix A

5.7.1 Additional Figures and Tables

| Original Project | | | | Alternative Project | | | | Comparison Alternative Project and Original Project | | | |
|------------------|------------------|------------|------------|---------------------|---------------------|----------------|--------------------|---|-----------------|--|--|
| EV ex ante | EV after failure | SD ex ante | SD ex post | Earnings if success | Earnings if failure | Expected value | Standard Deviation | Difference EV* | Difference SD** | | |
| 200 | 175 | 100 | 96.82 | 272 | 78 | 175 | 97 | 0 | 0% | | |
| 200 | 175 | 100 | 96.82 | 262 | 68 | 165 | 97 | 10 | 0% | | |
| 200 | 175 | 100 | 96.82 | 273 | 79 | 176 | 97 | -1 | 0% | | |
| 200 | 175 | 100 | 96.82 | 282 | 88 | 185 | 97 | -10 | 0% | | |
| 200 | 175 | 100 | 96.82 | 292 | 98 | 195 | 97 | -20 | 0% | | |
| 200 | 175 | 100 | 96.82 | 283 | 67 | 175 | 108 | 0 | -10% | | |
| 200 | 175 | 100 | 96.82 | 263 | 87 | 175 | 88 | 0 | 10% | | |
| 200 | 175 | 100 | 96.82 | 256 | 94 | 175 | 81 | 0 | 20% | | |

* The difference in expected value is expressed as the absolute difference in ECU by which the original project differs from the benchmark alternative project.
 ** The difference in Standard deviation is the relative difference in variance of the original project compared to the alternative project, rounded to full percentage points.

| Original Project | | | | Alternative Project | | | | Comparison Alternative Project and Original Project | | | |
|------------------|------------------|------------|------------|---------------------|---------------------|----------------|--------------------|---|-----------------|--|--|
| EV ex ante | EV after success | SD ex ante | SD ex post | Earnings if success | Earnings if failure | Expected value | Standard Deviation | Difference EV* | Difference SD** | | |
| 200 | 225 | 100 | 96.82 | 322 | 128 | 225 | 97 | 0 | 0% | | |
| 200 | 225 | 100 | 96.82 | 312 | 118 | 215 | 97 | 10 | 0% | | |
| 200 | 225 | 100 | 96.82 | 323 | 129 | 226 | 97 | -1 | 0% | | |
| 200 | 225 | 100 | 96.82 | 332 | 138 | 235 | 97 | -10 | 0% | | |
| 200 | 225 | 100 | 96.82 | 342 | 148 | 245 | 97 | -20 | 0% | | |
| 200 | 225 | 100 | 96.82 | 333 | 117 | 225 | 108 | 0 | -10% | | |
| 200 | 225 | 100 | 96.82 | 313 | 137 | 225 | 88 | 0 | 10% | | |
| 200 | 225 | 100 | 96.82 | 306 | 144 | 225 | 81 | 0 | 20% | | |

* The difference in expected value is expressed as the absolute difference in ECU by which the original project differs from the benchmark alternative project.
 ** The difference in Standard deviation is the relative difference in variance of the original project compared to the alternative project, rounded to full percentage points.

(a) Possible situations after an experienced failure

(b) Possible situations after an experienced success

Table 5.3: Project Combinations

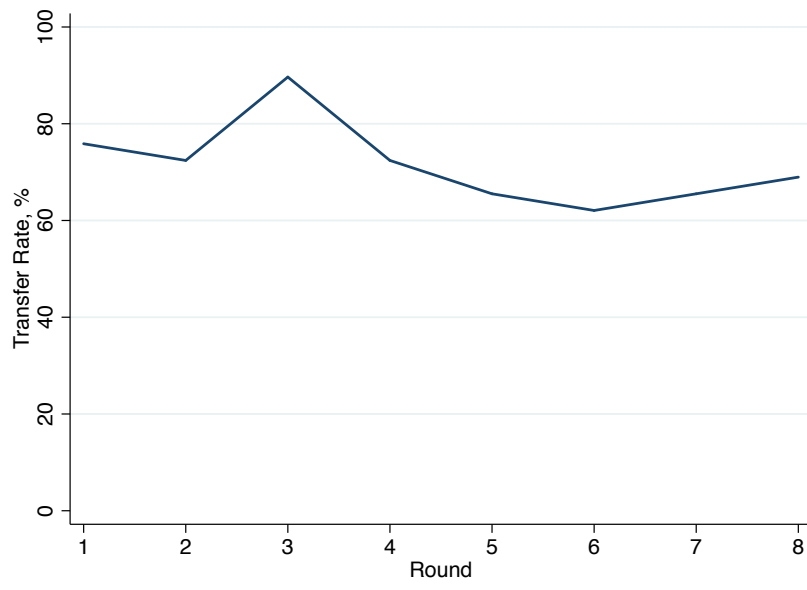


Figure 5.7: Transfer Decisions over Different Rounds

| | Investor switches project | | | |
|--------------------------------|---------------------------|----------------------|----------------------|----------------------|
| | (1) | (2) | (3) | (4) |
| History | -0.119 (-0.48) | | -0.190 (-0.74) | -0.263 (-1.00) |
| Transfer | 0.699* (2.22) | | 0.630 (1.95) | 0.623 (1.91) |
| Positive Experience | -0.348 (-1.66) | -0.707*** (-4.73) | -0.497* (-2.29) | -0.499* (-2.30) |
| History × Positive Experience | 0.0160 (0.05) | | 0.139 (0.40) | 0.163 (0.46) |
| Transfer × Positive Experience | -1.214** (-3.18) | | -1.075** (-2.75) | -1.047** (-2.67) |
| Prior Result Positive | | -0.708*** (-4.80) | -0.699*** (-4.71) | -0.723*** (-4.83) |
| Expected Value Difference | | -0.0282** (-3.18) | -0.0269** (-3.01) | -0.0258** (-2.87) |
| SD difference | | 0.00437 (0.50) | 0.00329 (0.37) | 0.00294 (0.33) |
| Round | | | | 0.0196 (0.61) |
| Female | | | | 0.0662 (0.40) |
| Age | | | | 0.00825 (0.29) |
| Economics Student | | | | -0.206 (-1.16) |
| Choice number | | | | -0.0526 (-1.56) |
| Constant | 0.0664 (0.43) | 0.452** (3.21) | 0.408* (2.26) | 0.486 (0.65) |
| Individuals | 106 | 106 | 106 | 106 |
| N | 848 | 848 | 848 | 848 |

Random effects model with z -statistics in parentheses, using robust standard errors

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.4: Investor decision regressions, probit

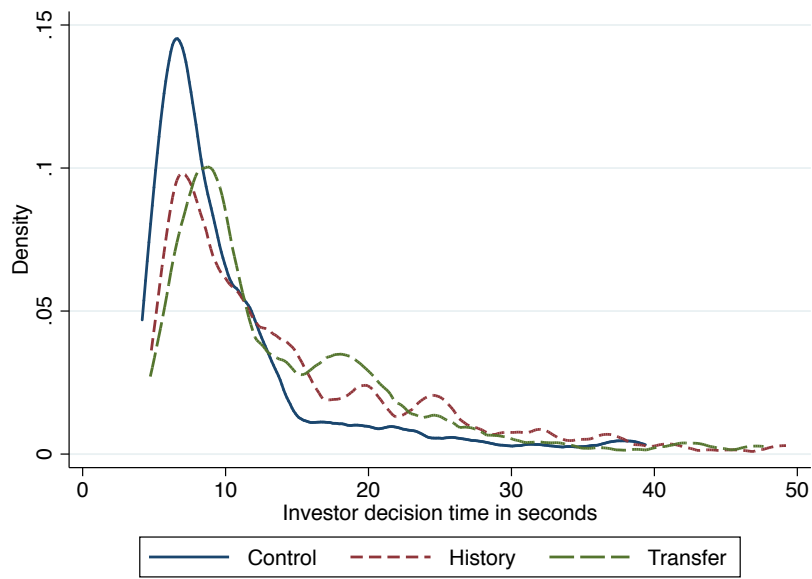


Figure 5.8: Decision time density estimate, ignoring outliers above 50 seconds (Epanechnikov kernel with bandwidth of 1 second)

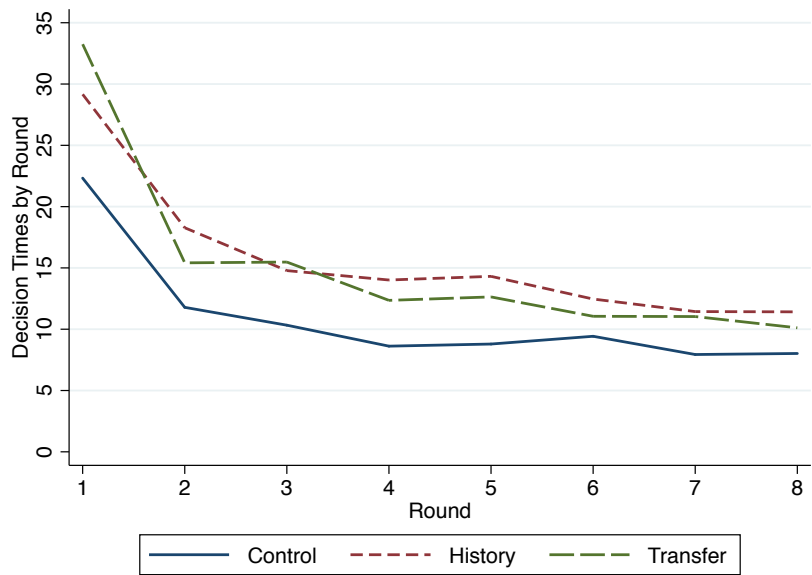


Figure 5.9: Decision time over different rounds

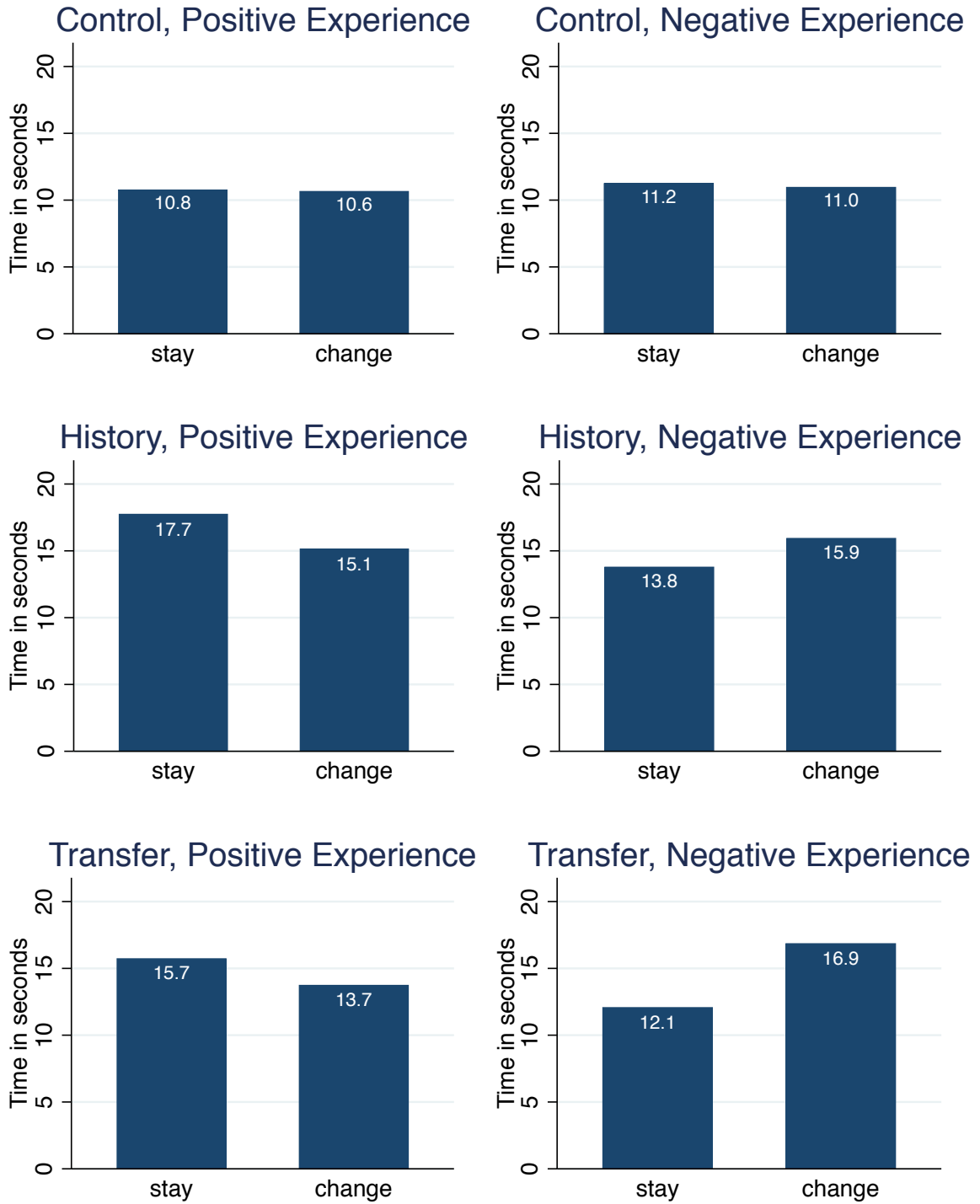


Figure 5.10: Decision time by Experience

| | Decision Time | | | |
|------------------------------------|---------------|----------|----------|-----------|
| | (1) | (2) | (3) | (4) |
| History | 3.595* | | 3.440* | 3.686** |
| | (2.54) | | (2.43) | (2.62) |
| Transfer | 4.267* | | 4.180* | 4.734** |
| | (2.55) | | (2.49) | (2.94) |
| Positive Experience | -0.389 | 0.485 | -0.622 | -0.606 |
| | (-0.35) | (0.63) | (-0.56) | (-0.60) |
| History × Positive Experience | 2.523 | | 2.801 | 2.619 |
| | (1.39) | | (1.55) | (1.58) |
| Transfer × Positive Experience | 0.110 | | 0.292 | -0.831 |
| | (0.06) | | (0.15) | (-0.47) |
| Investor Switches Project | | -0.360 | -0.376 | -0.106 |
| | | (-0.46) | (-0.48) | (-0.15) |
| Prior Result Positive | | -1.490* | -1.587* | -1.599* |
| | | (-1.97) | (-2.10) | (-2.32) |
| Absolute Expected Value Difference | | 0.00198 | 0.000921 | -0.00277 |
| | | (0.03) | (0.02) | (-0.05) |
| Absolute SD Difference | | -0.0586 | -0.0586 | -0.0614 |
| | | (-0.95) | (-0.95) | (-1.10) |
| Round | | | | -1.843*** |
| | | | | (-12.80) |
| Female | | | | -0.0618 |
| | | | | (-0.07) |
| Age | | | | -0.208 |
| | | | | (-1.29) |
| Economics Student | | | | 1.558 |
| | | | | (1.53) |
| Choice Number | | | | -0.0132 |
| | | | | (-0.08) |
| Constant | 11.10*** | 14.32*** | 12.48*** | 24.48*** |
| | (12.72) | (13.66) | (10.42) | (5.81) |
| Individuals | 106 | 106 | 106 | 106 |
| N | 848 | 848 | 848 | 848 |
| R2 | 0.0434 | 0.0071 | 0.0499 | 0.2009 |

Random effects model with z -statistics in parentheses, using robust standard errors

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.5: Investor decision time regressions

| | History | Transfer | Total |
|-----------------------------------|------------------|-----------------|-----------------|
| Emotion and Obligation | 0.54 (0.0038)*** | 0.35 (0.0598)* | 0.58(0.0000)*** |
| Obligation and Likelihood to Stay | 0.15 (0.4260) | 0.06 (0.7430) | 0.32 (0.0149)** |
| Emotion and Likelihood to Stay | 0.11 (0.5742) | 0.41 (0.0266)** | 0.42(0.0012)*** |

Table 5.6: Correlations between questionnaire answers

| | Investor Switches Project | | |
|---|---------------------------|---------------------|---------------------|
| | (1) | (2) | (3) |
| Transfer | -0.0152 (-0.05) | -0.203 (-0.73) | 0.396 (1.47) |
| Positive Experience | 0.134 (0.25) | -0.0514 (-0.10) | 0.167 (0.35) |
| Result previous project | -0.00932 (-0.05) | -0.0217 (-0.11) | -0.0346 (-0.17) |
| Expected Value Difference | -0.0296* (-2.39) | -0.0288* (-2.36) | -0.0276* (-2.26) |
| SD difference | 0.0112 (0.91) | 0.0137 (1.13) | 0.0135 (1.12) |
| Emotion | 0.210 (1.55) | | |
| Positive Experience \times Emotion | -0.333* (-2.11) | | |
| Obligation | | 0.424** (2.71) | |
| Positive Experience \times Obligation | | -0.337 (-1.88) | |
| Likelihood to Stay | | | 0.0304 (0.22) |
| Positive Experience \times Likelihood to Stay | | | -0.409* (-2.53) |
| Constant | -0.438 (-0.97) | -0.800* (-1.99) | -0.0670 (-0.17) |
| Individuals | 56 | 58 | 58 |
| N | 448 | 464 | 464 |

Random effects model with z -statistics in parentheses, using robust standard errors

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5.7: Investor decision regressions

5.8 Appendix B: Supplementary Material

5.8.1 Instructions

An interactive version of the instructions is available on http://www.mhoyer.com/inv_feelings, together with an interactive demo of one round of the experiment. A copy of the files hosted there is available from the library of the University of Amsterdam together with the digital version of this thesis.

5.8.2 Questionnaires

Transfer, manager

1. What is your age (in numbers)?
2. What is your gender?
 - female
 - male
3. What is your (primary) study program (if not a student please choose that)?
4. How would you describe your decision making process when deciding whether to send a transfer or not?
5. Did you send any transfers to your investor after being chosen?
 - Yes
 - No
6. If yes, which was the most important reason to do so?
 - Transferring doubled the income for the group as a whole
 - I hoped making the transfer would make the investor stay with my project
 - I just tried to be nice to the investor
7. Were you disappointed by an investor who switched to another project?
 - Yes, every time.
 - Yes, but only if my project was better than the alternative.
 - Yes, but only if I had sent the transfer.
 - No, the investor can choose what they want.
 - Not applicable, every investor I met stayed with my project.

History, Manager

1. What is your age (in numbers)?
2. What is your gender?

- female
- male

3. What is your (primary) study program (if not a student please choose that)?

Transfer, investor

1. What is your age (in numbers)?

2. What is your gender?

- female
- male

3. What is your (primary) study program (if not a student please choose that)?

4. How would you describe your decision making process when choosing whether to stay with a project manager or not in general?

5. Did you calculate the success probability of a project?

- Yes
- No
- I tried to, but failed

6. Did you try to calculate the expected value of the different projects? (expected value is probability times earnings)

- Yes
- No

7. Did you feel a positive emotion towards a manager who sent you a transfer?

- 1 - Not at all
- 2
- 3
- 4
- 5 - Very strongly
- The first project never succeeded.

Next page:

8. Did you feel a sense of obligation towards a manager who sent you a transfer?

- 1 - Not at all
- 2
- 3
- 4
- 5 - Very strongly
- Never received a transfer

9. Were you more likely to stay with a project and manager if the manager sent you a transfer earlier?

- 1 - Not at all
- 2
- 3
- 4
- 5 - A lot
- Never received a transfer

History, investor

1. What is your age (in numbers)?

2. What is your gender?

- female
- male

3. What is your (primary) study program (if not a student please choose that)?

4. How would you describe your decision making process when choosing whether to stay with a project manager or not in general?

5. Did you calculate the success probability of a project?

- Yes
- No
- I tried to, but failed

6. Did you try to calculate the expected value of the different projects? (expected value is probability times earnings)

- Yes
- No

7. Did you feel a positive emotion towards a manager if the first project succeeded?

- 1 - Not at all
- 2
- 3
- 4
- 5 - Very strongly
- The first project never succeeded

Next page:

8. Did you feel a sense of obligation towards a manager who's first project was a success?

- 1 - Not at all

- 2
- 3
- 4
- 5 - Very strongly
- Never received a transfer

9. Were you more likely to stay with a project and manager if the manager's first project was a success and, if so, how much more?

- 1 - Not at all
- 2
- 3
- 4
- 5 - A lot
- Never received a transfer

Control, investor

1. What is your age (in numbers)?

2. What is your gender?

- female
- male

3. What is your (primary) study program (if not a student please choose that)?

4. How would you describe your decision making process when choosing whether to stay with a project or not in general?

5. Did you calculate the success probability of a project?

- Yes
- No
- I tried to, but failed

6. Did you try to calculate the expected value of the different projects? (expected value is probability times earnings)

- Yes
- No

5.9 Appendix C: Alternative Theories

5.9.1 The Role of Other Social Preference Models in our Design

Looking at the History treatment, in the design that is implemented in this experiment many social decision theories popular within economics do not play any role whatsoever. This is a direct result of two design features:

First, project managers do not know what project they are choosing and do not have any influence on the income of the investor beyond the act of choosing a project that is more or less likely to be successful. This precludes any influence of reciprocity of any kind.

Secondly, the payment scheme chosen for the project manager effectively nullifies many concerns that investors might have about the effect that their choice might have on other subjects. Since by the time the investors' decisions are made every project manager has earned the exact same amount - 200 ECU for each of the first two projects - motives such as inequity aversion or envy are meaningless. Since investors cannot affect the type of distribution of the others' earnings in any way, this is true irrespective of the exact theory applied, such as for example Fehr and Schmidt (1999) or Bolton and Ockenfels (2000). In fact, social welfare concerns are also irrelevant, therefore excluding approaches such as simple max-min preferences or the model in Charness and Rabin (2002).

In the Transfer treatment the situation is slightly different. Since the transfer decision is intentional, intention-based theories predict an effect. Still, inequality-oriented motives play a negligible role: They are only relevant if the project manager withholds the transfer. In that case she earns 10 ECU more from the exchange than managers who did not send the transfer or, more importantly, were inactive. This however is only a one-twentieth of the fixed income that managers earn if they are recruited for the final project, giving us confidence that we can ignore it as a confounding factor.

5.9.2 The Case of Flexible Project Manager Payment

In section 5.5 we mentioned that one potential explanation for not being able to pick up a result is that subjects might have thought of the design in terms of a outcome dependent payment systems for managers. We also mentioned that such a system could be argued to potentially have lead to stronger results, creating a stronger emotional connection between investors and managers. To shed further light on how these two claims can coexist we will now investigate the theoretical implications of a flexible manager payment system in our design.

Game Design and Best Response

We start by changing the payment of the project manager from a fixed fee (200 ECU in the experiment) to a share of the return of the project, assigning s to the investor and $1-s$ to the manager ($0 < s < 1$). In a world in which earnings are outcome dependent it makes sense to normalize the income of inactive managers at 0. We assume that the income from a failed project is positive and the income from a successful project is only restricted by the condition that it be strictly larger than the income from a failed project. This is a more general formulation than the fixed earnings of 300 and 100 used in equation (5.1), but still ensures that nobody can make losses from a project, which will be convenient later on. Together with the assumption that both project types, whether they have a high or low success probability, generate each outcome with a probability that is strictly larger than zero, implies that even a project with a low success probability has positive expected value. The only further restriction on projects is that, while different projects can have different returns, we impose that returns after successes and failures are such that for any projects of the high or low success probability type (π_H, π_L) or an unknown history (π) the ordering $0 < E(\pi_L) < E(\pi) < E(\pi_H)$ holds. Compared to the situation in section 5.3, equations (5.1) through (5.6) stay fundamentally the same, but are scaled

by the factor s and generalized, leading to the following:

$$E(\pi^O|h) = s \left(P(\pi = \pi_H|h)E(\pi_{HO}) + P(\pi = \pi_L|h)E(\pi_{LO}) \right) \quad (5.8a)$$

$$E(\pi^A) = s \left(\frac{1}{2}E(\pi_{HA}) + \frac{1}{2}E(\pi_{LA}) \right) \quad (5.8b)$$

The situation differs more when we start constructing the equivalent of equation (5.7). In particular we are looking at an investor with a positive or negative regard for the manager with whom she has some kind of previous history, henceforth manager M_1 . We weigh the influence of this manager's earnings by factor α_{M_1} . With a utility function that is linear in the investor's and the manager's payoffs the expected utility of staying with the current project is:

$$\begin{aligned} E\left(U(\pi_o|h)\right) &= s \left(P(\pi = \pi_H|h)E(\pi_H) + P(\pi = \pi_L|h)E(\pi_L) \right) + \\ &\quad \alpha_{M_1}(1-s) \left(P(\pi = \pi_H|h)E(\pi_H) + P(\pi = \pi_L|h)E(\pi_L) \right) \\ &= \left(s + (1-s)\alpha_{M_1} \right) \left(P(\pi = \pi_H|h)E(\pi_H) + P(\pi = \pi_L|h)E(\pi_L) \right) \end{aligned} \quad (5.9)$$

where we ignore income from the first project (History treatment) or the transfer stage (Transfer treatment) for the moment.

The expected utility of choosing a new project is unaffected by any preferential other-regarding factor and free of any informative history about the project success. As a result it is identical to the expected value of a project of unknown type that is managed by a manager who has not been encountered previously.

$$E\left(U(\pi)\right) = s \left(\frac{1}{2}E(\pi_H) + \frac{1}{2}E(\pi_L) \right) \quad (5.10)$$

This implies that, even if the expected value of an alternative project is higher than the expected value of a previously implemented project, an investor might decide to stay with the current project and project manager, if the specific other-regarding preference factor α_M is big enough. Vice versa, a negative α_M could make the investor switch projects,

even if the expected value of the alternative project is lower than that of the original project.

The arguments for why we should observe the effect hypothesized about in the main experiment is therefore fairly similar to what we described in section 5.3. What differs however are the implications of other theories of social preferences, which we investigate in the rest of the appendix.

General inequity aversion

Since we cannot directly observe if the investor develops a specific regard for the original manager's earnings, we have to rely on their decision to indirectly analyze their behavior for the presence of the presumed effect. We will now look at other potential motives for staying with the original manager. In particular, could different types of inequity aversion explain this behavior?

A complicating factor would be if investors do not only develop a specific factor α_M for the well-being of their immediate peer, but also exhibit general inequity aversion that incorporates the whole pool of potential managers – in this case adding all possible managers in a experiment session. We argue that, since the investor can shift income to at most one additional manager, we can restrict ourselves to the effect that such a decision would have on the manager that is ultimately chosen, even though the particular manager has not yet been selected when the investor decides whether to change projects or not. To analyze this situation formally we need to describe all histories that can arise out of the up to two (History treatment) different projects that can be experienced prior to the decision whether to stay with the current project and project manager or not. From this point onwards we speak only of the History treatment. Equivalent arguments apply to the Transfer treatment.

Assume one investor I and two managers M_1 and M_2 and different histories of the implementations of the first two projects, h_{ff} (all project draws were failures) and h_{fs}, h_{sf}, h_{ss} (the first and/or the second project were successful). We summarize these three cases by writing \bar{h} . Assuming a completely linear payment scheme, the relative wealth position

of the investor and the first manager is only determined by factor s . In all histories the incomes of the investor and M_1 are at least as high as the income of M_2 . After any history \bar{h} we know for certain that their income prior to the decision being made is larger than that of M_2 . Staying with the original partner is equivalent to choosing to repeat the previous project and choosing a new partner is equivalent to letting them choose a new, random project, therefore we will express the investor's choices in terms of choosing projects rather than partners. We denote cumulative earnings of agent x at time t by Π_t^x , the utility of the investor at time t as a function of history h and project choice c as $U_t(h, c)$, and use E_2 to describe an expectation that is formulated at time $t = 2$, just after the second (History treatment) or first (Transfer treatment) project has been implemented and before the decision to proceed or switch has been made.

We start by assuming max-min preferences for the investor. If that is the true model for her preferences we can express her expectation of the utility at $t = 3$ that results from choosing the alternative project at $t = 2$ as

$$\begin{aligned}
E_2\left(U_3(h, \pi_A)\right) &= \min\left(E_2(\Pi_3^I(h, \pi_A)), E_2(\Pi_3^{M_1}(h, \pi_A)), E_2(\Pi_3^{M_2}(h, \pi_A))\right) \\
&= \min\left(\Pi_2^I(h) + s\left(\frac{1}{2}E(\pi_H) + \frac{1}{2}E(\pi_L)\right), \Pi_2^{M_1}(h), \right. \\
&\quad \left. \Pi_2^{M_2}(h) + (1-s)\left(\frac{1}{2}E(\pi_H) + \frac{1}{2}E(\pi_L)\right)\right)
\end{aligned} \tag{5.11}$$

The expected utility from choosing a new draw of the original project is

$$\begin{aligned}
E_2\left(U_3(h, \pi_O)\right) &= \min\left(\Pi_2^I(h) + s\left(P(\pi_H|h)E(\pi_H) + P(\pi_L|h)E(\pi_L)\right), \right. \\
&\quad \left. \Pi_2^{M_1}(h) + (1-s)\left(P(\pi_H|h)E(\pi_H) + P(\pi_L|h)E(\pi_L)\right), \right. \\
&\quad \left. \Pi_2^{M_2}(h)\right)
\end{aligned} \tag{5.12}$$

We first look at the three different histories subsumed under \bar{h} : Since $\Pi_2^{M_2}(\bar{h}) \leq \Pi_2^{M_1}(\bar{h})$ (strict inequality unless the return of a failed project is zero) the expected utility of switching is greater or equal than the alternative that would result from sticking with the current manager, irrespective of the expected values of π_L , π , or π_H . Following the fourth history, h_{ff} , the result is fundamentally the same, with the sole exception

being the situation in which the income from a failed project is exactly zero. In that case everybody's earned income is zero and max-min preferences lead to indifference between the two options, because any choice gives at least one manager an income of zero. Thus we can state conclusively that the max-min preference can only ever drive an investor to be more prone to switching to an alternative project and manager, but not the opposite. What we are really interested in however is how much more or less of a factor this becomes after different histories. The active manager's income scales with the amount earned in the projects directly leads to the answer: The effect is smallest after h_{ff} , equally big after h_{fs} and h_{sf} , and biggest after h_{ss} . This is a direct result of the fact that $(1-s)\Pi^{M_1}(h_{ff}) < (1-s)\Pi^{M_1}(h_{fs}) = (1-s)\Pi^{M_1}(h_{sf}) < (1-s)\Pi^{M_1}(h_{ss})$.

Note that the situation is different if inactive managers earn the same as employed managers during the first and second period of the game, as was the case in section 5.3. In that case $\Pi_2^{M_1}(h) = \Pi_2^{M_2}(h)$ and the only difference between the expected utility following the two choices is in the relative size of the factors $\frac{1}{2}E(\pi_H) + \frac{1}{2}E(\pi_L)$ and $P(\pi_H|h)E(\pi_H) + P(\pi_L|h)E(\pi_L)$. Max-min preferences can therefore only ever play a role if the investor has earned less than the managers and in that case their effect points in the opposite direction as our main hypothesis implies.

Inequity preferences of the Fehr and Schmidt (1999) type have slightly different implications. We use the standard formulation $U_i(x) = x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max(x_j - x_i, 0) - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max(x_i - x_j, 0)$ with $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$. Assuming $s \geq \frac{1}{2}$, the only possible difference is advantageous to the investor, i.e. only the investor's β is of relevance. Since dis-utility from earning differences only affects utility linearly in the model, it does not matter if the welfare difference between the investor and the original manager or the welfare difference between the investor and the alternative manager becomes grows more extreme. The investor only cares about the sum of inequality, but not its distribution among the two managers. If $s < \frac{1}{2}$, the active manager earns more from a project than the investor. This triggers the "envy" parameter α of the utility function, which is multiplied with the difference between their respective earnings. Since the inactive manager does not have any accrued income and the model assumes $\beta_i \leq \alpha_i$, the envy aspect of the utility

function increases the investor's motive to stay with the current manager. Similarly to the mechanism applied in the max-min case, this motive is stronger the higher previous earnings are, making it more relevant after histories h_{fs} and h_{sf} , and biggest after h_{ss} . ERC (Bolton and Ockenfels, 2000) preferences do not consider the total distribution of incomes, but merely the agent's own income and the relative income of the agent compared to the average income of the group. As such, even without using any of ERC's other assumptions, we know that the distribution of income between the two managers is not of relevance to the investor's utility. Since ERC also imposes strictly increasing utility in own income it therefore implies that the project with the highest expected value is preferred.

As in the design used in the experiment, the equivalent argument in the Transfer treatment is simultaneously simpler - we only have to consider the outcomes of one project as opposed to two - and more complicated - managers who withhold the transfer have earned more than those who send it or are inactive³⁶. The former difference does not change the direction of any of the described effects. The latter difference can be shown to motivate an investor to switch away from an original manager who withheld a transfer more often than from one who sent it, using any of the theories presented here.

Concluding Remarks on Flexible Manager Payment

In conclusion, we have shown that in the History treatment the only effect that the presented theories of inequity aversion could explain would be a propensity to switch more often to a different project and manager after successful than after failed projects. Therefore they are unequivocally pointing in the opposite direction of the hypothesis that forms the basis of this experiment, namely that investors place a positive weight on the future income of managers with whom they experienced a positive history. Using a result-dependent payment system for the managers would therefore be a valid approach from the perspective of hypothesis testing. As we have seen, it would however create many situations in which at least some conventionally found forms of social preferences (max-min,

³⁶In case of a loss a withholding manager now also has higher earnings than the investor.

ERC) describe effects that work in the opposite direction as the suspected effect, making it very hard to detect. This is why we decided to choose a fixed fee for the managers instead.

Summary

The way people make decisions in social situations is in many ways a product of how they feel about those affected by their decisions. Both positive and negative emotions are important factors in decision-making. This thesis presents results from a series of laboratory experiments designed to contribute to our understanding of emotions in social economic games. In particular, we focus on the contrast between negative and positive emotional reactions towards other individuals.

Chapters 2 to 4 cover a series of experiments in which groups of two players interacted repeatedly in public good games that allow the expression of both destructive and cooperative behaviour. Chapter 2 is based on the publication "Destructive behavior in a fragile public good game" (Hoyer et al., 2014). In that paper we investigate socially destructive behavior in a repeated game in which the costs of acting increasingly destructively scale in exactly the same way as the costs of acting increasingly cooperatively. This gives us the unique opportunity to compare these two types of behavior directly. We find substantial evidence of destructive decisions, sometimes leading to sour relationships characterized by persistent hurtful behavior. Changing the framing of the game to positive framing induces fewer destructive decisions, while shifting the selfish Nash equilibrium towards minimal taking doubles the share of destructive decisions to more than 20%.

The third chapter uses data from the same experiment, but moves from group level analysis to the individual level. We estimate a modified version of a social ties model (van Dijk and van Winden, 1997) in order to directly model the emotional impact of different types of behavior by others onto our subjects decision making. Contrary to our hypothesis we find that cooperative behavior by others has a greater impact on the level of a tie with a partner than destructive behavior. This stands in some contrast to

results from other areas, where negative events typically lead to stronger reactions than positive ones, such as the loss aversion effect.

Chapter 4 uses a similar game, but takes it to a fMRI-based experiment. We investigate whether the processing of destructive behavior is similar or distinct from the processing of constructive behavior. In addition, we investigate the role that our subjects social value orientation plays on their processing of the game. We observe a high degree of destructive behavior. In line with Bault et al. (2014) higher tie-values correlate with activation in a particular region of the pSTS, both for constructive and destructive decisions. Payoff differences between self and other also correlate with pSTS activity.

In chapter 5 we stay true to the topic of investigating the way in which social preference can adapt based on previous experiences, but move to a different situation: we investigate the relationship between an investor and a project manager. Other than in the earlier chapters, we also investigate different types of potential sources for such relationships, namely direct transfers and the experience of having shared a successful or unsuccessful experience in the past. If investors want to abandon a project with low success probability, they also have to change project managers. An additional joint project or a voluntary transfer precedes their interaction. We hypothesize that investors favor projects that are managed by project managers with whom they have shared positive experiences in the past, which should lead to a lower rate of project change than in comparable situations with shared negative experiences. The role of this social element is isolated using a control treatment in which the role of the project manager does not exist. Interaction through voluntary transfers plays a clear and significant role in the investors decision making, whereas the influence of merely sharing a positive or negative experience proves more complex than anticipated.

In conclusion, all chapters of this thesis investigate different aspects of the interaction between two different people. We began by focusing on the difference between positive and negative ties that result from cooperative or destructive behavior by others. While the prevalence of destructive behavior was generally lower and destructive feuds were a comparatively rare sight, we did not find that destructive behavior by others caused

stronger reactions than cooperative behavior. Neither the model-based analysis, which only took behavior into account, nor the analysis using fMRI showed clear signs pointing in that direction. The final chapter, which focused on the distinction between different kinds of positive and negative experiences instead of merely direction, found clear differences between intentional and circumstantial shared experiences, with the latter proving much less effective at prompting a behavioral reaction. This implies that a more rigorous investigation of circumstantial experiences in the style of the early chapters would most likely not be fruitful. At the same time, the topic of what differentiates the processing of positive and negative experiences in a social setting and their impact on affective states still provides a multitude of open questions, as we still do not know what drives the differences in prevalence and persistence of negative and positive ties.

Samenvatting

De beslissingen die mensen nemen in sociale dilemma's worden op vele manieren beïnvloed door de gevoelens die mensen hebben voor diegenen die geraakt worden door deze beslissingen. Zowel positieve als negatieve emoties spelen een belangrijke rol in deze besluitvorming. Deze scriptie bespreekt de resultaten van een serie experimenten die zijn ontworpen om het inzicht in het effect van emoties in sociale dilemma's te vergroten. In het bijzonder richt ik me hierbij op het contrast tussen de (ontwikkeling van) positieve en negatieve emoties ten opzichten van anderen.

Hoofdstukken 2 tot en met 4 behandelen een serie experimenten waarin twee spelers herhaaldelijk een publiek goed spel spelen waarbij spelers zowel constructief als destructief gedrag konden tonen. Hoofdstuk 2 is gebaseerd op het artikel "Destructive behavior in a fragile public good game" (Hoyer et al., 2014). In dit artikel onderzoeken we destructief gedrag in een herhaald spel waarbij de kosten van destructief gedrag dezelfde zijn als die voor cooperatief gedrag. Dit geeft ons de unieke mogelijkheid deze twee type gedrag direct te vergelijken. We vinden hierbij substantieel bewijs voor destructief gedrag, wat soms leidt tot een verstoorde relatie waarin beide partijen elkaar voortdurend bewust schaden. Het veranderen van de aankleding, niet de inhoud, van het experiment leidde tot verschillen in de aanwezigheid van destructief gedrag. Een positieve setting leidde tot minder destructieve beslissingen, terwijl het verplaatsen van het baatzuchtige Nash-evenwicht naar het negatieve gedeelte van de keuzeruimte het aantal destructieve keuzes verdubbelde tot wel 20% van de totaal gemaakte keuzes.

Het derde hoofdstuk analyseert data van hetzelfde experiment, maar richt zich op het individuele niveau. We schatten een aangepaste versie van het 'social-ties' model (Van

Dijk and Van Winden, 1997) om het emotionele effect op de proefpersonen, veroorzaakt door verschillende gedragingen van anderen, te modelleren. Tegengesteld aan onze hypothese zien we dat cooperatief gedrag van anderen een groter absoluut effect heeft op de 'social-tie' dan destructief gedrag. Dit lijkt te conflicteren met de bevindingen uit andere onderzoeksvelden, waar negatieve gebeurtenissen over het algemeen tot sterkere reacties leiden dan positieve gebeurtenissen. Een bekend voorbeeld hiervan is het 'loss aversion effect'.

Hoofdstuk 4 gebruikt een soortgelijk spel als de voorgaande hoofdstukken, nu in een fMRI experiment. We onderzoeken of hersenen destructief gedrag anders verwerken dan constructief gedrag. Daarbij bekijken we de rol die de sociale waarde oriëntatie van onze proefpersonen speelt bij de verwerking van het spel. We zien in deze setting relatief veel destructief gedrag. Overeenkomstig met Bault et al. (2015) correleren ook in deze studie hoge 'tie-values' met activiteit in een bepaalde regio in de pSTS, zowel voor constructieve als voor destructieve beslissingen. Ook verschillen verschillen in beloning tussen gepaarde spelers correleren met activiteit in de pSTS.

Ook hoofdstuk 5 behandelt het onderwerp van de verandering in sociale preferenties ten gevolge van eerdere ervaringen, maar onderzoekt dit nu in een andere omgeving: de relatie tussen een investeerder en een project manager. Anders dan in de eerdere hoofdstukken onderzoeken we in dit hoofdstuk verschillende ontstaanwijzes van dergelijke relaties, te weten directe betalingen en gedeelde succesvolle en onsuccesvolle ervaringen. Als investeerders in dit experiment uit een project met een lage kans van slagen willen stappen moeten ze van project manager veranderen. Deze interactie wordt vooraf gegaan door ofwel een eerdere interactie, ofwel een vrijwillige betaling. Onze hypothese is dat investeerders een voorkeur hebben projecten die worden beheerd door managers met wie ze positieve ervaringen delen, wat zou moeten leiden tot minder veranderingen van project door managers na een positieve ervaring dan na een negatieve ervaring. De rol van de sociale interactie wordt gesoleerd door het gebruik van een controle experiment waarin er geen project manager is. We vinden dat vrijwillige betalingen een duidelijke en significante rol spelen in in het beslissingsproces van de investeerder en dat er bij het delen van

positieve en negatieve ervaringen geen duidelijk patroon te herkennen is.

Alle hoofdstukken van deze thesis onderzoeken verschillende aspecten van de interactie tussen twee mensen. Als eerste richten wij ons op het verschil tussen negatieve en positieve onderlinge banden die resulteerde in cooperatief dan wel destructief gedrag. Hierbij vinden we dat er minder destructief dan cooperatief gedrag werd vertoond, dat destructieve vetes zeldzaam waren en dat destructief gedrag geen sterkere emotionele impact had dan vergelijkbaar cooperatief gedrag. Nog de gedragsanalyse, nog de fMRI-resultaten wijzen op een verschil in effectgrote tussen de positieve en negatieve impulsen. Het laatste hoofdstuk, waarin verschillende soorten positieve en negatieve ervaringen worden geanalyseerd, vindt een duidelijk verschil tussen intentionele en door de omgeving gegenereerde ervaringen, waarbij op de laatste veel minder gereageerd wordt. Dit impliceert dat een zelfde soort onderzoek naar deze door de omgeving gecreerde ervaringen, zoals naar intentionele ervaringen gedaan is in de eerdere hoofdstukken, waarschijnlijk weinig op zal leveren. Tegelijkertijd blijven er nog vele open vragen bestaan over de verschillen tussen de verwerking van positieve en negatieve ervaringen in een sociale omgeving. Zo weten we nog steeds niet wat het verschil in de aanwezigheid van positieve en negatieve sociale banden precies veroorzaakt.

Bibliography

- Abbink, K. and Herrmann, B. (2009). Pointless vendettas. *Norwich: University of East Anglia CBESS Working paper*, 2.
- Abbink, K. and Sadrieh, A. (2009). The pleasure of being nasty. *Economics Letters*, 105(3):306–308.
- Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *Quarterly journal of Economics*, pages 715–753.
- Akerlof, G. A. and Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic perspectives*, pages 9–32.
- Andreoni, J. (1995). Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics*, 110(1):1–21.
- Andreoni, J. and Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, pages 737–753.
- Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., and Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of neurophysiology*, 94(1):327–337.
- Baron, J. and Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of personality and social psychology*, 54(4):569.
- Bartels, A. and Zeki, S. (2000). The neural basis of romantic love. *Neuroreport*, 11(17):3829–3834.
- Bartra, O., McGuire, J. T., and Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of bold fmri experiments examining neural correlates of subjective value. *Neuroimage*, 76:412–427.
- Bault, N., Fahrenfort, J. J., Pelloux, B., Ridderinkhof, K. R., and van Winden, F. (2017). An affective social tie mechanism: Theory, evidence, and implications. *Journal of Economic Psychology*, 61:152–175.
- Bault, N., Pelloux, B., Fahrenfort, J. J., Ridderinkhof, K. R., and van Winden, F. (2014). Neural dynamics of social tie formation in economic decision-making. *Social cognitive and affective neuroscience*, 10(6):877–884.

- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., and Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology*, 5(4):323.
- Baumeister, R. F. and Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin*, 117(3):497.
- Behrens, T. E., Hunt, L. T., Woolrich, M. W., and Rushworth, M. F. (2008). Associative learning of social value. *Nature*, 456(7219):245.
- Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, 10(1):122–142.
- Bhanji, J. P. and Delgado, M. R. (2014). The social brain and reward: social information processing in the human striatum. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1):61–73.
- Bolle, F., Tan, J., and Zizzo, D. (2013). Vendettas. *American Economic Journal: Microeconomics (forthcoming)*.
- Bolton, G. E. and Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American economic review*, pages 166–193.
- Bone, J., Hey, J., and Suckling, J. (2004). A simple risk-sharing experiment. *Journal of Risk and Uncertainty*, 28(1):23–38.
- Bone, J. D., Hey, J. D., and Suckling, J. R. (2003). Do people plan ahead? *Applied Economics Letters*, 10(5):277–280.
- Boot, A. W. (2000). Relationship banking: What do we know? *Journal of financial intermediation*, 9(1):7–25.
- Bosman, R. and van Winden, F. (2002). Emotional hazard in a power-to-take experiment. *The Economic Journal*, 112(476):147–169.
- Bosman, R. and van Winden, F. (2010). Global risk, investment and emotions. *Economica*, 77(307):451–471.
- Breitmoser, Y. (2015). Cooperation, but no reciprocity: Individual strategies in the repeated prisoner’s dilemma. *American Economic Review*, 105(9):2882–2910.
- Brewer, M. B. and Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology*, 50(3):543–549.
- Brock, J. M., Lange, A., and Ozbay, E. Y. (2013). Dictating the risk: Experimental evidence on giving in risky environments. *The American Economic Review*, 103(1):415–437.
- Brown, M. and Serra-Garcia, M. (2014). The threat of exclusion and relational contracting. *Working Paper*.
- Brown, M. and Zehnder, C. (2007). Credit reporting, relationship banking, and loan repayment. *Journal of Money, Credit and Banking*, 39(8):1883–1918.

- Camerer, C. and Ho, T. H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4):827–874.
- Cappelen, A. W., Konow, J., Sørensen, E. Ø., and Tungodden, B. (2013). Just luck: An experimental study of risk-taking and fairness. *The American Economic Review*, 103(4):1398–1413.
- Carlson, N. (2011). The real history of Twitter. *Business Insider*. Online; accessed 22-September-2015.
- Carswell, S. (2012). *Anglo Republic: Inside the Bank that Broke Ireland*. Penguin.
- Carter, C. S. (1998). Neuroendocrine perspectives on social attachment and love. *Psychoneuroendocrinology*, 23(8):779–818.
- Cason, T. N., Lau, S.-H. P., and Mui, V.-L. (2015). Prior interaction, identity, and cooperation in the inter-group prisoners dilemma. *Working Paper*.
- Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly journal of Economics*, 117(3):817–869.
- Charness, G., Rigotti, L., and Rustichini, A. (2007). Individual behavior and group membership. *The American Economic Review*, 97(4):1340.
- Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14(1):47–83.
- Chen, Y. and Li, S. X. (2009). Group identity and social preferences. *The American Economic Review*, pages 431–457.
- Cochard, F., Van, P. N., and Willinger, M. (2004). Trusting behavior in a repeated investment game. *Journal of Economic Behavior & Organization*, 55(1):31–44.
- Cookson, R. (2000). Framing effects in public goods experiments. *Experimental Economics*, 3(1):55–79.
- Cornée, S., Masclet, D., et al. (2013). Long-term relationships, group lending and peer sanctioning in microfinance: New experimental evidence. *Working Paper*.
- Cornée, S., Masclet, D., and Thenet, G. (2012). Credit relationships: Evidence from experiments with real bankers. *Journal of Money, Credit and Banking*, 44(5):957–980.
- Cornelissen, G., Dewitte, S., and Warlop, L. (2011). Are social value orientations expressed automatically? decision making in the dictator game. *Personality and Social Psychology Bulletin*, page 0146167211405996.
- Dal Bó, P. and Fréchet, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *The American Economic Review*, 101(1):411–429.
- D’Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Baetens, E., Luxen, A., Maquet, P., and Salmon, E. (2007). Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. *Cognitive Neuroscience, Journal of*, 19(6):935–944.

- Decety, J. and Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, 41(2):127–138.
- Declerck, C. H., Boone, C., and Emonds, G. (2013). When do people cooperate? the neuroeconomics of prosocial decision making. *Brain and cognition*, 81(1):95–117.
- van Dijk, F., Sonnemans, J., and van Winden, F. (2002). Social ties in a public good experiment. *Journal of Public Economics*, 85(2):275–299.
- van Dijk, F. and van Winden, F. (1997). Dynamics of social ties and local public good provision. *Journal of Public Economics*, 64(3):323–341.
- Duchin, R. and Sosyura, D. (2013). Divisional managers and internal capital markets. *The Journal of Finance*, 68(2):387–429.
- Duffy, S. and Smith, J. (2014). Cognitive load in the multi-player prisoner’s dilemma game: Are there brains in games? *Journal of Behavioral and Experimental Economics*, 51:47–56.
- Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, pages 848–881.
- Fahrenfort, J. J., Van Winden, F., Pelloux, B., Stallen, M., and Ridderinkhof, K. R. (2012). Neural correlates of dynamically evolving interpersonal ties predict prosocial behavior. *Frontiers in neuroscience*, 6.
- Falk, A. and Fischbacher, U. (2006). A theory of reciprocity. *Games and economic behavior*, 54(2):293–315.
- Fareri, D. S., Chang, L. J., and Delgado, M. R. (2012a). Effects of direct social experience on trust decisions and neural reward circuitry. *Frontiers in neuroscience*, 6.
- Fareri, D. S., Niznikiewicz, M. A., Lee, V. K., and Delgado, M. R. (2012b). Social network modulation of reward-related signals. *The Journal of neuroscience*, 32(26):9045–9052.
- Fehl, K., Sommerfeld, R. D., Semmann, D., Krambeck, H.-J., and Milinski, M. (2012). I dare you to punish me - vendettas in games of cooperation. *PloS ONE*, 7(9):e45093.
- Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960):785–791.
- Fehr, E. and Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The journal of economic perspectives*, pages 159–181.
- Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868.
- Fehr, E. and Zehnder, C. (2009). Reputation and credit market formation: How relational incentives and legal contract enforcement interact. *Working Paper*.

- Fermin, A. S., Sakagami, M., Kiyonari, T., Li, Y., Matsumoto, Y., and Yamagishi, T. (2016). Representation of economic preferences in the structure and function of the amygdala and prefrontal cortex. *Scientific reports*, 6.
- Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, 71(3):397–404.
- Fisher, H., Aron, A., and Brown, L. L. (2005). Romantic love: an fmri study of a neural mechanism for mate choice. *Journal of Comparative Neurology*, 493(1):58–62.
- Fließbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., and Falk, A. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, 318(5854):1305–1308.
- Fudenberg, D., Rand, D. G., and Dreber, A. (2012). Slow to anger and fast to forgive: cooperation in an uncertain world. *The American Economic Review*, 102(2):720–749.
- Garvert, M., Moutoussis, M., Kurth-Nelson, Z., Behrens, T., and Dolan, R. (2015). Learning-induced plasticity in medial prefrontal cortex predicts preference malleability. *Neuron*, 85(2):418 – 428.
- Glimcher, P. W. and Fehr, E. (2014). *Neuroeconomics: Decision making and the brain*. Academic Press.
- Goerg, S. J. and Walkowitz, G. (2010). On the prevalence of framing effects across subject-pools in a two-person cooperation game. *Journal of Economic Psychology*, 31(6):849–859.
- Goette, L., Huffman, D., and Meier, S. (2012). The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups. *American Economic Journal: Microeconomics*, 4(1):101–115.
- Gurdal, M. Y., Miller, J. B., and Rustichini, A. (2013). Why blame? *Journal of Political Economy*, 121(6):1205–1247.
- Hampton, A. N., Bossaerts, P., and O’Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences*, 105(18):6741–6746.
- Haselmann, R., Schoenherr, D., and Vig, V. (2014). Lending in social networks. *Working Paper*.
- Heatherton, T. F., Wyland, C. L., Macrae, C. N., Demos, K. E., Denny, B. T., and Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *Social Cognitive and Affective Neuroscience*, 1(1):18–25.
- Hill, C. A., Suzuki, S., Polania, R., Moisa, M., O’Doherty, J. P., and Ruff, C. C. (2017). A causal account of the brain network computations underlying strategic social behavior. *Nature Neuroscience*, 20(8):1142–1149.
- Hirschman, A. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Harvard University Press.

- Houser, D., Schunk, D., and Winter, J. (2010). Distinguishing trust from risk: an anatomy of the investment game. *Journal of economic behavior & organization*, 74(1):72–81.
- Hoyer, M. (2012). Social value orientation and spillovers of cooperation.
- Hoyer, M., Bault, N., Loerakker, B., and van Winden, F. (2014). Destructive behavior in a fragile public good game. *Economics Letters*, 123(3):295–299.
- Hoyer, M. and van Winden, F. (2016). Investors have feelings too. *Working Paper*.
- Izuma, K., Saito, D. N., and Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, 58(2):284–294.
- Johnson, N. D. and Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5):865–889.
- Kafka, P. (2015). Interview with Jonah Peretti. *Recode.net*. [Online; accessed 22-September-2015].
- Kagel, J. H., Roth, A. E., and Hey, J. D. (1995). *The handbook of experimental economics*. Princeton university press Princeton, NJ.
- Karlan, D. S., Morten, M., and Zinman, J. (2015). A personal touch: Text messaging for loan repayment. *Working Paper*.
- Keysar, B., Converse, B. A., Wang, J., and Epley, N. (2008). Reciprocity is not give and take: Asymmetric reciprocity to positive and negative acts. *Psychological Science*, 19(12):1280–1286.
- Khadjavi, M. and Lange, A. (2015). Doing good or doing harm: experimental evidence on giving and taking in public good games. *Experimental Economics*, 18(3):432–441.
- Khwaja, A. I. and Mian, A. R. (2005). Do lenders favor politically connected firms? rent provision in an emerging financial market. *Quarterly Journal of Economics*, 120(4).
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science*, 308(5718):78–83.
- Kohls, G., Perino, M. T., Taylor, J. M., Madva, E. N., Cayless, S. J., Troiani, V., Price, E., Faja, S., Herrington, J. D., and Schultz, R. T. (2013). The nucleus accumbens is involved in both the pursuit of social reward and the avoidance of social punishment. *Neuropsychologia*, 51(11):2062–2069.
- Krienen, F. M., Tu, P.-C., and Buckner, R. L. (2010). Clan mentality: evidence that the medial prefrontal cortex responds to close others. *The Journal of Neuroscience*, 30(41):13906–13915.
- Krueger, F., McCabe, K., Moll, J., Kriegeskorte, N., Zahn, R., Strenziok, M., Heinecke, A., and Grafman, J. (2007). Neural correlates of trust. *Proceedings of the National Academy of Sciences*, 104(50):20084–20089.

- Kuhnen, C. M. (2009). Business networks, corporate governance, and contracting in the mutual fund industry. *The Journal of Finance*, 64(5):2185–2220.
- Kuhnen, C. M. (2015). Asymmetric learning from financial information. *The Journal of Finance*, 9:2029–2061.
- Laury, S. K. and Holt, C. A. (2008). Voluntary provision of public goods: experimental results with interior nash equilibria. *Handbook of experimental economics results*, 1:792–801.
- Lebreton, M., Jorge, S., Michel, V., Thirion, B., and Pessiglione, M. (2009). An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron*, 64(3):431–439.
- Ledyard, J. (1995). Public goods: A survey of experimental research. In Kagel, J. and A.E., R., editors, *Handbook of Experimental Economics*. Princeton University Press, Princeton.
- Leider, S., Möbius, M. M., Rosenblat, T., and Do, Q.-A. (2009). Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics*, 124(4):1815–1851.
- Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of economic dynamics*, 1(3):593–622.
- Li, W., Mai, X., and Liu, C. (2014). The default mode network and social understanding of others: what do brain connectivity studies tell us. *Frontiers in human neuroscience*, 8.
- Liebrand, W. and McClintock, C. (1988a). The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *European Journal of Personality*, 2:217–230.
- Liebrand, W. B. and McClintock, C. G. (1988b). The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *European journal of personality*, 2(3):217–230.
- Loerakker, B., Bault, N., Hoyer, M., and van Winden, F. (2016). Asymmetry in the development of cooperative and antagonistic relationships. a model-based analysis of a fragile public good game experiment. *Working Paper*.
- Lunawat, R. (2013). An experimental investigation of reputation effects of disclosure in an investment/trust game. *Journal of Economic Behavior & Organization*, 94:130–144.
- Malmendier, U. M. and Schmidt, K. M. (2012). You owe me. *Working Paper*.
- McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, 98(20):11832–11835.
- Miller, B. L. and Cummings, J. L. (2007). *The human frontal lobes: Functions and disorders*. Guilford press.

- Morita, H. and Servátka, M. (2013). Group identity and relation-specific investment: An experimental investigation. *European Economic Review*, 58:95–109.
- Myerson, R. (1997). *Game Theory: Analysis of Conflict*. Harvard University Press.
- Nikiforakis, N. and Engelmann, D. (2011). Altruistic punishment and the threat of feuds. *Journal of Economic Behavior & Organization*, 78(3):319–332.
- Nikiforakis, N., Noussair, C. N., and Wilkening, T. (2012). Normative conflict and feuds: The limits of self-enforcement. *Journal of Public Economics*, 96(9):797–807.
- Offerman, T. (2002). Hurting hurts more than helping helps. *European Economic Review*, 46(8):1423–1437.
- Penenberg, A. L. (2011). Eric Lies is a lean startup machine. *Fast Company*. [Online; accessed 22-September-2015].
- Pessiglione, M. and Delgado, M. R. (2015). The good, the bad and the brain: neural correlates of appetitive and aversive values underlying decision making. *Current Opinion in Behavioral Sciences*, 5:78–84.
- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., and Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, 59(4):561–567.
- Plott, C. R. and Smith, V. L. (2008). *Handbook of experimental economics results*, volume 1. Elsevier.
- Poldrack, R. A., Mumford, J. A., and Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. Cambridge University Press.
- Puri, M., Rocholl, J., and Steffen, S. (2013). What kinds of bank-client relationships matter in reducing loan defaults and why?
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American economic review*, 83(5):1281–1302.
- Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D., and Nowak, M. A. (2009). Positive interactions promote public cooperation. *Science*, 325(5945):1272–1275.
- Rangel, A. and Clithero, J. (2014). The computation of stimulus values in simple choice. In Glimcher, P. and Fehr, E., editors, *Neuroeconomics*, chapter 8, pages 125–148. Academic Press.
- Reuben, E. and Van Winden, F. (2008). Social ties and coordination on negative reciprocity: The role of affect. *Journal of Public Economics*, 92(1):34–53.
- Rick, S. (2011). Losses, gains, and brains: Neuroeconomics can help to answer open questions about loss aversion. *Journal of Consumer Psychology*, 21(4):453–463.
- Rilling, J. K. and Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual review of psychology*, 62:23–48.

- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, 22(4):1694–1703.
- Roth, A. E. and Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, 8(1):164–212.
- Ruff, C. C. and Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, 15(8):549–562.
- Saijo, T. (2008). Chapter 85 spiteful behavior in voluntary contribution mechanism experiments. volume 1 of *Handbook of Experimental Economics Results*, pages 802 – 816. Elsevier.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626):1755–1758.
- Schulz, J. F., Fischbacher, U., Thöni, C., and Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology*, 41:77–87.
- Sharot, T. and Garret, N. (2016). Forming beliefs: Why valence matters. *Trends in Cognitive Sciences*, 20(9):25–33.
- Singer, T., Seymour, B., O’Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, 439(7075):466–469.
- Singer, T. and Tusche, A. (2014). Understanding others: Brain mechanisms of theory of mind and empathy. In Glimcher, P. and Fehr, E., editors, *Neuroeconomics*, chapter 27, pages 513–532. Academic Press.
- Sonnemans, J., van Dijk, F., and van van, F. (2006). On the dynamics of social ties structures in groups. *Journal of Economic Psychology*, 27:187–204.
- Sonnemans, J., Schram, A., and Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior & Organization*, 34(1):143–161.
- Strombach, T., Weber, B., Hangebrauk, Z., Kenning, P., Karipidis, I. I., Tobler, P. N., and Kalenscher, T. (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. *Proceedings of the National Academy of Sciences*, 112(5):1619–1624.
- Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations*, 33(47):74.
- Tricomi, E., Rangel, A., Camerer, C. F., and O’Doherty, J. P. (2010). Neural evidence for inequality-averse social preferences. *Nature*, 463(7284):1089–1091.

- Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics*, pages 1039–1061.
- Uzzi, B. (1999). Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *American sociological review*, pages 481–505.
- Wagner, D. D., Haxby, J. V., and Heatherton, T. F. (2012). The representation of self and person knowledge in the medial prefrontal cortex. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(4):451–470.
- Willinger, M. and Ziegelmeyer, A. (1999). Framing and cooperation in public good games: an experiment with an interior solution. *Economics Letters*, 65(3):323–328.
- Yang, Y., Onderstal, S., and Schram, A. (2016). Inequity aversion revisited. *Journal of Economic Psychology*, 54:1–16.
- Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current directions in psychological science*, 10(6):224–228.
- Zeki, S. (2007). The neurobiology of love. *FEBS letters*, 581(14):2575–2579.
- Zink, C. F., Tong, Y., Chen, Q., Bassett, D. S., Stein, J. L., and Meyer-Lindenberg, A. (2008). Know your place: neural processing of social hierarchy in humans. *Neuron*, 58(2):273–283.
- Zizzo, D. J. (2003). Money burning and rank egalitarianism with random dictators. *Economics Letters*, 81(2):263–266.