On the role of bonding, emotional leadership, and partner choice in games of cooperation and conflict
Loerakker, B.A.

Citation for published version (APA):
Loerakker, B. A. (2018). On the role of bonding, emotional leadership, and partner choice in games of cooperation and conflict

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: http://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.
Human behavior is difficult to capture in simple and elegant models. In similarly looking environments different people show opposite behaviors and small changes in these environments can radically change the behavior of some, while huge changes might be completely ignored by others. The explanations for these behaviors have traditionally been approached from two different angles. In economics and (behavioral) game theory an individual is reduced to a set of beliefs, preferences and possible actions, whereby the, not always rational, mind rules over the body. On the other hand, in psychology and neurosciences, hormones, emotions and other direct physical impulses control humans with rationality playing a much smaller part. This thesis tries to harmonize both views to a certain extent, as it tries to show that both rational and emotional elements can be modeled together, studies destructive behavior, investigates the role of emotions in leadership and analyses the effect of partner choice, based on only physical characteristics, on the level of cooperation.

The author holds a BA in history and a MSc in economy from the University of Amsterdam as well as a MPhil from the Tinbergen Institute. He currently pursues a career outside of academia and lives in Amsterdam.
On the Role of Bonding, Emotional Leadership, and Partner Choice in Games of Cooperation and Conflict
ISBN: 978 90 3610 509 5

Cover illustration: Lejla Verheus

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 707 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.
On the Role of Bonding, Emotional Leadership, and Partner Choice in Games of Cooperation and Conflict

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 woensdag 2 maart 2018, te 12:00 uur

door

Ben Aloysius Loerakker

geboren te Amsterdam
**Promotiecommissie:**

**Promotores:**
- Prof. dr. F.A.A.M. van Winden Universiteit van Amsterdam
- Prof. dr. J. Sonnemans Universiteit van Amsterdam

**Overige leden:**
- Prof. dr. J.J.M. Potters Universiteit van Tilburg
- Prof. dr. A.J. Dur Erasmus Universiteit Rotterdam
- Prof. dr. R. Sloof Universiteit van Amsterdam
- Prof. dr. C.M. van Veelen Universiteit van Amsterdam
- dr. A. Ule Universiteit van Amsterdam

**Faculteit:** Economie en Bedrijfskunde
Voor Henriette en Joost
Acknowledgements

When I finally got the confirmation of my successful enrolment into the the Mphil program of the Tinbergen Institute in the spring of 2010, I was filled with joy. I celebrated by having diner with Helena in our then favorite Japanese restaurant and everything seemed me to be straightforward from there on. Soon I discovered that, although TI and the subsequent PhD research at CREED was a great and very valuable experience, it was by no means straightforward.

The year before my interest in the TI program was aroused by two courses in the ‘regular’ economics master at the UvA. Game Theory by Theo Offerman was the first course that I considered to be fun and where I enjoyed thinking about the presented problems. The second course that shaped my thinking was microeconomics taught by Hessel Oosterbeek. He always challenged me and the other students, and we learned how to carefully read economic papers. Theo and Hessel were also kind enough to write reference letters for me. I want to hereby thank them for both the inspiration and the letters.

At TI I soon found out that the courses were much more challenging, and most of them also more interesting, than the courses I followed in the previous years. Thanks to my fellow TI’ers though, I kept enjoying even the courses that I found to be very hard and perhaps less interesting. A special thanks hereby goes out to Chris, with whom I did most of the coursework and helped me through some of the most ‘challenging’ macro courses.

After the course work Frans van Winden became my supervisor and together with Nadège Bault introduced me to the world of neuroeconomics. Nadège also helped me develop my programming. I want to thank her for all the energy see put into our research and for teaching
me that the brain is often much more subtle than even we behavioral economist assume it to be. Next I want to thank Amsterdam Brain and Cognition (ABC) for the financial support making my PhD research possible.

Most influential to my research and most indebted though I am to Frans, who I want to thank for teaching me how to write and for the enormous amount of hours he spent reading and commenting on my work. I also very much enjoyed our discussions on research, politics and all things related to human behavior.

I would like to thank here also my roommates at CREED: Aaron, Boris, Margarita, Matze, Max and Simin. You always made sure the atmosphere was good, combining a dedication to research with sharp discussions and plenty of fun. Max deserves some special attention as we worked on the first two chapters of this thesis, followed courses in psychology and set our first steps in the neuro field together.

Another aspect that made life at CREED good were the colleagues in general and specifically the table football. Whether it took up too much or too little time I am unsure of, but it was equally serious and hilarious. In this light, I want to especially thank Jeroen, Theo and Joël for the fearsome competition and the many laughs.

Obviously I couldn’t have made it through the entire PhD saga without my friends, ‘The Boys’. The countless hours I have been (and will be) with you in the bars of Amsterdam I consider to be time very well spent.

Furthermore I would like to thank my mother Henriette, Joost, my brothers Jan and Lex, and my sister Lejla for their continuous warmth. You are always there for me when I need your help. I am very lucky with such a family.

Finally, and most of all, I would like to thank my girlfriend Helena. She has been with me through this entire chapter of my life and supported me throughout it. She always stayed positive and kept motivating me, although I can be very difficult, grumpy, and unreasonable, a remarkable feat for which I will be forever thankful.
Contents

1 Introduction ................................................................. 1

2 Destructive Behavior in a Fragile Public Good Game* ........... 9

   2.1 Introduction ................................................................ 9
   2.2 Experiment ............................................................. 11
       2.2.1 Symmetric Treatment (SYM) ............................... 11
       2.2.2 Framing Treatment (FRAME) ............................... 11
       2.2.3 Asymmetric Treatment (ASYM) ........................... 12
   2.3 Results ..................................................................... 12
   2.4 Conclusion ............................................................. 18

Appendix 2.A Payoff matrix .................................................. 19
### 3 Asymmetry in the Development of Cooperative and Antagonistic Relationships. A Model-Based Analysis of a Fragile Public Good Game Experiment.*

<table>
<thead>
<tr>
<th>Section</th>
<th>Subsection</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>21</td>
</tr>
<tr>
<td>3.2</td>
<td>Theory</td>
<td></td>
</tr>
<tr>
<td>3.2.1</td>
<td>Fragile Public Good Game</td>
<td>23</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Model</td>
<td>25</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Model analysis and hypotheses</td>
<td>27</td>
</tr>
<tr>
<td>3.3</td>
<td>Methods</td>
<td>30</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Experiment</td>
<td>30</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Estimation</td>
<td>32</td>
</tr>
<tr>
<td>3.4</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>3.4.1</td>
<td>Descriptive statistics</td>
<td>33</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Group level results</td>
<td>35</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Individual level results</td>
<td>38</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Predictive performance out-of-sample and model comparison</td>
<td>39</td>
</tr>
<tr>
<td>3.5</td>
<td>Applying the Ties Model to the Repeated Prisoner’s Dilemma</td>
<td>43</td>
</tr>
<tr>
<td>3.6</td>
<td>Conclusion</td>
<td>51</td>
</tr>
</tbody>
</table>

#### Appendix 3.A Proof of Propositions

3.A.1 Proposition 1: Contributions outside of $-5 \leq C_{it} \leq 5$ can never be part of any equilibrium if $-1 \leq \alpha \leq 1$. 52

3.A.2 Proposition 2: All symmetric equilibria with $-5 \leq C_{it} \leq 5$ are possible. 54

3.A.3 Proposition 3: Asymmetric equilibria exist if $C_i C_j > 0$ and $|C_i| \leq 5$ and $|C_j| \leq 5$ 56

3.A.4 Proposition 4: For the socially optimal choices to be a stable equilibrium under the myopic model, both players satisfying $0.2 < \frac{\delta_2}{1 - \delta_1} < 0.25$ is a necessary, but not a sufficient condition. 57

#### Appendix 3.B Reference Point

iv
5 Does partner choice foster cooperation?

5.1 Introduction
5.2 Literature
5.3 Design and hypotheses
5.3.1 Experimental procedures and treatments
5.3.2 Hypotheses
5.4 Results
5.4.1 Treatment results
5.4.2 Results per role
5.4.3 Estimation results at treatment level
5.4.4 Estimation results at role level
5.4.5 Behavior of Choosers
5.5 Conclusion

Appendix 5.A Instructions

Bibliography

Samenvatting (Summary in Dutch)
Chapter 1

Introduction

Human behavior is difficult to capture in simple and elegant models. In similarly looking environments different people show opposite behaviors and small changes in these environments can radically change the behavior of some, while huge changes might be completely ignored by others. The explanations for these behaviors have traditionally been approached from two different angles. In economics and (behavioral) game theory an individual is reduced to a set of beliefs, preferences and possible actions, whereby the, not always rational, mind rules over the body. On the other hand, in psychology and neurosciences, hormones, emotions and other direct physical impulses control humans with rationality playing a much smaller part (for an informative discussion regarding this divide see: Gul and Pesendorfer (2008) and Camerer et al. (2005)). This thesis tries to harmonize both views to a certain extent, as it tries to show that both rational and emotional elements can be modeled together, investigates the role of emotions in leadership and studies the effect of partner choice, based on only physical characteristics, on the level of cooperation.

One of the behaviors that is hard to explain by the more traditional economic line of reasoning is repeated costly destructive behavior hurting others. Although this behavior is bad for everyone involved we find it in the spread of computer viruses, vandalism and ongoing vendettas. Destructive behavior, however, has been understudied for a long time in (experimental) economics. One example of destructive behavior is the rejection of strictly positive monetary
payoffs by responders in the ultimatum game, hereby often fairness considerations are cited as the main reason for rejection (for a comprehensive overview see: Güth and Kocher (2014)). Recently, though, there has been more attention for dynamic situations where agents seem to develop preferences that make them gain utility by hurting others, not just out of fairness considerations, whereby an agent gives up a higher payoff in order to diminish the payoff of another player. Examples of studies are Zizzo (2003), Nikiforakis and Engelmann (2011), and Bolle et al. (2014). Often, though, games have been used where the least (or the most, for instance, in the ultimatum game) cooperative decision an agent could make was equal to the standard (rational, selfish) Nash equilibrium action. A new element in our studies is the introduction of the Fragile Public Good (FPG) game. The game is a two-player non-linear public good game with a constant increase of the marginal cost of contribution. Next to contributing to a common account, both players can also take from the common account. This has the same, detrimental, effect on the taking player as contributing does, but instead of helping the other player is hurts the other player. The game thus has the interesting feature that it is symmetric in the sense that it is a non-linear public good game that gives as much room for cooperative as for destructive behavior.

In Chapter 2 we investigate the dynamics of negative bilateral relationships and the effects of framing on such relationships in an experimental setting. In a two-player FPG game in a partner setting, we find that individuals make destructive decisions. More striking, though, is that very subtle changes in the environment could lead to much more destructive behavior. In a game where the Nash equilibrium contribution is negative, but not destructive, twice as many destructive decisions (around 20% of the total decisions) are made and the length of the feuds, situations where both players make repeated destructive decisions, increases accordingly. Furthermore, we find support for the affective ties model of van Dijk and van Winden (1997). In this model the weight a player, \(i\), attaches to the payoff given to player \(j\), \(U_{ij}\), is expressed as a tie, \(\alpha_{ij}\). The tie is not per definition symmetric and is impacted by impulses, \(I_{ijt}\). These impulses are caused by the behavior of \(j\) vis-à-vis a reference point. In short: If \(j\) behaves
nicer (to i) than I’s reference contribution, i will take the payoff of j into account more and if j behaves worse he or she will take the payoff of j into account less. The model also allows for negative ties, in such instances players are willing to reduce their own payoff in order to reduce the payoff of the other. In our experiment, dyads that make repeatedly destructive decisions spiral into a sour relationship and find it increasingly difficult to get out of them. This is in line with negative emotional ties being created and then strengthened by the mutual destructive decisions.

In chapter 3 the possibilities and explanatory power of the aforementioned ties model are tested and forward-looking agents are introduced to the model. We investigate a linearized version of the van Dijk and van Winden model. Chapter 3 is based on the same series of experiments as those researched in Chapter 2. In this chapter, however, we focus on the predictive power and estimation of the tie model. Furthermore, we build on the work done on this model by Bault et al. (2017). We find that the ties model predicts behavior of subjects out-of-sample significantly better than a standard fixed interdependent utility model as well as an established learning model.

The introduction of forward-looking agents taught us more about the modeling of behavior and its consequences. I realized that the impulse, albeit usually interpreted as an emotional mechanism, could also be interpreted as a signal of willingness to cooperate in the classical behavioral sense. This leaves a simple model that is able to capture both emotional as well as forward looking elements in its lone two parameters. We find that the forward-looking parameter is insignificant on the group level in all model specifications should therefore not be interpreted as a lack of strategic thinking by our subjects. For future research it would be interesting to determine more precisely what size the affective part of the impulse is and what the beliefs oriented part is. A potential way of disentangling the two parts would be to change, in a partner setting of a multi round FPG game, the continuation probability instead of keeping it constant to either 1 or zero (the last round). The intuition being that we would expect the affective part to remain, while the belief part would update according to the continuation
probability.

Another addition to the van Dijk and van Winden model was allowing for a dichotomy between the strength of positive and negative impulses. At the start it was not clear whether positive or negative emotional impulses would have a stronger effect on behavior. However, there was no reason to believe that these impulses should be equally strong. On one hand, there is evidence by Baumeister et al. (2001) and Baumeister and Leary (1995) that negative emotions affect behavior stronger. On the other hand there are studies by Rand et al. (2009) and King-Casas et al. (2005) who find that positive signals are better picked up than negative signals. In our study we find that the positive impulses have a stronger effect on behavior than negative impulses. Finally, the third chapter also shows that the, seemingly ad-hoc, strategy shifts in a repeated prisoners dilemma, as described by Dal Bó and Fréchette (2011) and Fudenberg et al. (2012), can be largely explained by a simple ties model.

Chapter 4 investigates the effect of emotion and leadership in a group conflict game. The motivation for this chapter came from the pillarization of the Dutch society in the early stages of the twentieth century. A period in Dutch history in which strongly opposed factions -Socialists, Catholics, Protestants and Liberals- lived almost completely at cross-purposes. Although seemingly hating each other, there were few altercations and both country and government were to a large extent stable. This period culminated in the pacification of 1917 in which three major laws regarding voting, schooling, and labor conditions were passed. Although, the demarcation in Dutch society was sharp, the leadership of the different factions or pillars stayed in good contact, was on very friendly personal terms and was able to deal with issues in a very pragmatic way. In front of their own audiences (or pillars), however, leaders were very negative about the other pillars, often blaming them for the failures in society. This drew our attention to the effect of leaders on outcomes of conflict. Could leaders broker peace in a difficult and hostile environment or would they stir up the masses in order to enhance conflict. In our current study we miss quite some aspects that were present in the Dutch pillarized society, though, most notably the ability of leaders to communicate among each other is missing. This research, thus, needs
to be seen as a first step in studying the effects that leaders have on conflict (games). Further studies could have leaders communicating with each other and allow leaders more means of communication with their followers.

More generally speaking, politics and leadership are often about the communication of emotions. Not seldom directed towards a (perceived) enemy. A group conflict game is a good instrument to study the effects of emotional leadership as there is a clear other group, that represents another faction in society or a different society altogether. In our game, two groups of interacted for multiple rounds, always in the same composition. In every round all players were endowed with tokens they could either keep or invest in the conflict. The members of the winning group got the tokens they held for themselves and got one-fourth of the tokens kept privately by the other group. The outcome of the conflict was determined by a lottery, whereby the odds where directly proportional to the amount of tokens invested in the conflict by each group. Every group had a leader. The selection of the leader was decided by an auction. In all treatments the leader contributed to the game before the others and this contribution was visible only for his group members.

In the game, however, contribution has not such clear consequences as in for instance a public good game. In a public good game contributing is (at least above a certain threshold) bad for the individual and good for the group. In the conflict game we used contribution is (again above a threshold) bad for an individual, good for his or her own group and bad for the other group. This in combination with the probabilistic nature of our game makes for more uncertainty, something that reflects reality well, but makes it hard for the leaders in our game to communicate clearly. This is amplified by the fact that our leaders had only a minimal instrument, a video evoking a basic emotion, to influence emotions. In the Baseline treatment every 3 rounds a neutral video clip was shown, in the Random Emotions (RE) treatment emotion evocation clips were shown randomly and in the Strategic Emotions (SE) treatment leaders could every 3 rounds select the emotion to be invoked in his or her followers.

We study the effect of leading-by-example as well as emotional leadership. In order to give
the leader the possibility to effectuate both forms of leadership, he or she was not only allowed to contribute to the conflict first but could in one treatment also choose a basic emotion that was evoked in his or her followers by specially selected and tested movies. The study shows that followers tend to follow the contribution of the leaders rather closely and that although the selected emotions do have an effect on followers, leaders do not appear to make use of this fact strategically. In general, leaders do not make use of their advantage of moving first in such a setting by free-riding on the contributions of the followers, unlike in the standard Nash equilibrium in which they contribute zero. Instead, they contribute more than their followers. Other behavioral models, like the ties model, can predict this behavior and are overall better in predicting the contribution level than the standard model.

Chapter 5 deals with the laboratory environments used in experimental economics. One of the crucial differences between the lab and the outside world - the environment we also as experimenters ultimately want to study - is that outside the lab people know, and often choose with whom they interact. If in economic experiments participants have the opportunity to choose their partner this choice is typically based upon prior behavior of the candidate partners, and the subject of study is typically reputation or belief formation. What I wanted to study, though, was the direct effect of choice, where the characteristics of the potential candidates were as limited as their physical appearance. This follows a recent stream of literature that looks into the effects on beliefs as well as behaviors of observing certain physical characteristics in others. Examples are: Tognetti et al. (2012) on cooperation, Bonnefon et al. (2013) on trust, and Van Leeuwen et al. (2017) on predictable rejections in an ultimatum game. In my final chapter I investigate how this possibility to choose your own partner based on pictures of the potential partners affects behavior.

When allowing participants to see with whom they (might) interact there are some effects that in a normal experimental environment it is hard if not impossible to correct for. Participants have the possibility for reciprocity outside the lab and take their prior relationships with them in the lab. Especially, if they are allowed to choose their partner the latter problem can be
severe. To deal with this problem I organized an experiment in two geographically separated locations, without the need for organizing it sequentially. The geographical distance limited the possibilities for reciprocity outside the lab and severely lowered the probabilities of existing relationships across the two subject pools. In the experiment of our study participants play a one-shot linear public good game with one partner as their main task. In one of the two treatments half of the subjects in one laboratory (so one fourth of the total subjects) was allowed to choose a partner from two potential partners. This created four types in the Choice treatment: Choosers, Chosen, Passives (those who are not able to choose, nor able to be chosen) and Leftovers. The behavioral differences between those types, except for the Choosers, are by and large insignificant when compared to each other or when compared to the Baseline treatment, where all participants were randomly matched. Before the main task, all participants were asked to do a circle test in order to elicit interpersonal distributional preferences (as in Sonnemans et al. (2006)). Moreover, they were asked to state their beliefs about the expected contribution of eight potential partners (both parts being incentivized), including their two potential partners they could choose from if they were Choosers. Participants were informed this was the case, but did not know which two would end up being their potential partners as this was determined by a draw only after the circle test and the belief elicitation. After the main task the circle test and belief elicitation were again performed, but now only with their actual partner of the main task. Before the final questionnaire, participants were also asked to rate the attractiveness of their partner.

I find that both affective as well as rational channels play an important role in the decision to contribute to the public good game. The effect of the ability to choose a partner, however, is not clear. Although participants in the treatment where they are allowed to choose their partner contribute around 10% more to the public good than participants in baseline, this difference is not significant at the conventional level. In the decision for a partner, not only the initial interpersonal preference for the potential partner (or initial tie, in terms of the ties model) and the expected contribution of the other (or belief) play an important role, but also attractiveness.
Attractiveness, however does not seem to effect the eventual contribution decision, indicating that beliefs about the contribution of the potential partners might mediate attractiveness.
Chapter 2

Destructive Behavior in a Fragile Public Good Game*

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 each other 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

*This chapter is based on joint work with Nadège Bault, Max Hoyer, and Frans van Winden, published in Economics Letters, see Hoyer et al. (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.
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 paper 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., 2014). 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. (2014) 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\(^1\). 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.

### 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 before\(^2\). Because only contributions can be made, this is a case of positive framing.

---

\(^1\)Formally, 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.

\(^2\)Payoff function: \(V_A(c_A, c_B) = 10(c_A + c_B) + 10(14 - c_A) - (|c_A - 7| + (c_A - 7)^2)\)
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 the equilibrium prescribes to take one token out of the common account, while both subjects contributing 5 tokens is the social optimum\(^3\).

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 tables\(^4\).

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 paper. 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.

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.

\[ V_A(c_A, c_B) = 10(14 + c_A + c_B) + 11(7 - c_A) + |c_A| - c_A^2 \]

\(^3\)Payoff function:

\(^4\)Instructions are available upon request.
Table 2.1: Average contributions

<table>
<thead>
<tr>
<th>Average contribution</th>
<th>SYM</th>
<th>FRAME</th>
<th>ASYM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>2.28 (2.01)</td>
<td>2.07 (1.69)</td>
<td>1.83 (3.3)</td>
</tr>
<tr>
<td>First round</td>
<td>1.26 (2.55)</td>
<td>-0.02 (2.55)</td>
<td>0.92 (3.06)</td>
</tr>
<tr>
<td>Rounds 26-34</td>
<td>2.44 (2.35)</td>
<td>2.49 (3.85)</td>
<td>2.13 (3.85)</td>
</tr>
<tr>
<td>Last round (35)</td>
<td>0.68 (2.43)</td>
<td>0.98 (1.81)</td>
<td>0.31 (3.81)</td>
</tr>
</tbody>
</table>

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 ($p = 0.001$). 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 (2013) 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.

5We use the Mann-Whitney U-test with dyad averages as observations unless otherwise mentioned.
All treatments show an increase in contributions over time until the (usually observed) sharp decline at the end; see figure 2.1. 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

<table>
<thead>
<tr>
<th>Percentage of destructive decisions</th>
<th>SYM</th>
<th>FRAME</th>
<th>ASYM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>11.25%</td>
<td>6.93%</td>
<td>21.14%</td>
</tr>
<tr>
<td>Last round</td>
<td>13.9%</td>
<td>5.36%</td>
<td>25%</td>
</tr>
</tbody>
</table>

Most relevant to this paper is the occurrence and development of destructive behavior hurting both partners. Relative to all decisions, destructive decisions count 11% in SYM, 7% in

---

6Contributions in the last round are lower than in the ten rounds before ($p < 0.01$ in all treatments.)
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 ASYM.\(^7\) Summing the destructive decisions of each dyad we find a difference only between SYM and ASYM (p = 0.094)\(^8\). 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.

\[ \text{Figure 2.2: Between-subject variance across treatments/rounds} \]

Background measures seem to play a role in explaining destructive behavior: In ASYM and

---

\(^7\)Means 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.

\(^8\)In line with a result of Nikiforakis et al. (2012), who find higher rates of counter-punishment in a treatment with increased normative conflict.
SYM contributions are lower and destruction rates higher in female dyads than in male dyads at 5% significance levels and SVO correlates with individual decisions\(^9\). It appears that sour relationships do indeed develop in the FPG game, as illustrated in two examples in figure 2.3.

![Graphs showing contribution over rounds for Subject A and Subject B in SYM](image)

**Figure 2.3:** Examples of sour relationships in SYM.

![Bar chart showing distribution of feuds](image)

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

\(^9\)In 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.
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 onwards\textsuperscript{10}.

\textsuperscript{10}Coefficients 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.
Table 2.4: Number of distinct dyads with at least one feud of a certain length

<table>
<thead>
<tr>
<th>Feud length</th>
<th>SYM</th>
<th>FRAME</th>
<th>ASYM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feuds of any length</td>
<td>16 (52)</td>
<td>4 (6)</td>
<td>18 (57)</td>
<td>38 (115)</td>
</tr>
<tr>
<td>Length = 1</td>
<td>14 (29)</td>
<td>3 (5)</td>
<td>16 (30)</td>
<td>33 (64)</td>
</tr>
<tr>
<td>Length = 2</td>
<td>9 (13)</td>
<td>0 (0)</td>
<td>7 (12)</td>
<td>16 (25)</td>
</tr>
</tbody>
</table>

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.
## Appendix 2.A Payoff matrix

<table>
<thead>
<tr>
<th>Choice Player 1</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take</td>
<td>84</td>
<td>94</td>
<td>104</td>
<td>114</td>
<td>124</td>
<td>134</td>
<td>144</td>
<td>154</td>
<td>164</td>
<td>174</td>
<td>184</td>
<td>194</td>
<td>204</td>
<td>214</td>
<td>224</td>
</tr>
<tr>
<td>Contribute</td>
<td>98</td>
<td>108</td>
<td>118</td>
<td>128</td>
<td>138</td>
<td>148</td>
<td>158</td>
<td>168</td>
<td>178</td>
<td>188</td>
<td>198</td>
<td>208</td>
<td>218</td>
<td>228</td>
<td>238</td>
</tr>
<tr>
<td>Player 2 Take</td>
<td>110</td>
<td>120</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
</tr>
<tr>
<td>Contribute</td>
<td>120</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
<td>260</td>
</tr>
<tr>
<td>Player 2 Take</td>
<td>128</td>
<td>138</td>
<td>148</td>
<td>158</td>
<td>168</td>
<td>178</td>
<td>188</td>
<td>198</td>
<td>208</td>
<td>218</td>
<td>228</td>
<td>238</td>
<td>248</td>
<td>258</td>
<td>268</td>
</tr>
<tr>
<td>Contribute</td>
<td>134</td>
<td>144</td>
<td>154</td>
<td>164</td>
<td>174</td>
<td>184</td>
<td>194</td>
<td>204</td>
<td>214</td>
<td>224</td>
<td>234</td>
<td>244</td>
<td>254</td>
<td>264</td>
<td>274</td>
</tr>
<tr>
<td>Player 2 Take</td>
<td>138</td>
<td>148</td>
<td>158</td>
<td>168</td>
<td>178</td>
<td>188</td>
<td>198</td>
<td>208</td>
<td>218</td>
<td>228</td>
<td>238</td>
<td>248</td>
<td>258</td>
<td>268</td>
<td>278</td>
</tr>
<tr>
<td>Contribute</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
<td>260</td>
<td>270</td>
<td>280</td>
</tr>
<tr>
<td>Player 2 Take</td>
<td>138</td>
<td>148</td>
<td>158</td>
<td>168</td>
<td>178</td>
<td>188</td>
<td>198</td>
<td>208</td>
<td>218</td>
<td>228</td>
<td>238</td>
<td>248</td>
<td>258</td>
<td>268</td>
<td>278</td>
</tr>
<tr>
<td>Contribute</td>
<td>134</td>
<td>144</td>
<td>154</td>
<td>164</td>
<td>174</td>
<td>184</td>
<td>194</td>
<td>204</td>
<td>214</td>
<td>224</td>
<td>234</td>
<td>244</td>
<td>254</td>
<td>264</td>
<td>274</td>
</tr>
<tr>
<td>Player 2 Take</td>
<td>128</td>
<td>138</td>
<td>148</td>
<td>158</td>
<td>168</td>
<td>178</td>
<td>188</td>
<td>198</td>
<td>208</td>
<td>218</td>
<td>228</td>
<td>238</td>
<td>248</td>
<td>258</td>
<td>268</td>
</tr>
<tr>
<td>Contribute</td>
<td>120</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
<td>260</td>
</tr>
<tr>
<td>Player 2 Take</td>
<td>110</td>
<td>120</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
<td>250</td>
</tr>
<tr>
<td>Contribute</td>
<td>98</td>
<td>108</td>
<td>118</td>
<td>128</td>
<td>138</td>
<td>148</td>
<td>158</td>
<td>168</td>
<td>178</td>
<td>188</td>
<td>198</td>
<td>208</td>
<td>218</td>
<td>228</td>
<td>238</td>
</tr>
</tbody>
</table>

Figure 5: Earnings of player 1, SYM
Chapter 3

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

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.¹ 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

*This chapter is based on a joint work with Nadège Bault, Maximilian Hoyer and Frans van Winden, see Loerakker et al. (2017). 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.

¹See Chaudhuri (2011) and Plott and Smith (2008) for an overview.
direct behavioral (van Dijk et al., 2002, Bault et al., 2016) as well as neurobiological (Bault et al., 2017) 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.\(^2\) 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 not been investigated in a horse race with other models using out-of-sample predictions regarding the same game.\(^3\) 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?\(^4\)

In this paper we will address each of these four issues. To address the first two, a novel game

\(^2\) Abbink 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 destroying their own earnings. More recently Bolle et al. (2014) 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.

\(^3\) See Bault et al. (2016)

\(^4\) See for instance Dal Bó and Fréchette (2011) and Fudenberg et al. (2012)
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.

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 in the opposite direction.

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, etcetera,
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. (2016). In this model $\alpha_{ijt}$ 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:

$$V_{it} = U_{it} + \alpha_{ijt}U_{jt}$$

(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...
This leads us to the following simple extension of eq. (3.1):

\[ V_{it} = U_{it} + \alpha_{ijt} U_{jt} + (U_{it+1} + \alpha_{ijt} U_{jt+1}) \]  

(3.2)

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

\[ U^e_{it} = 210 - C^2_{it} - |C_{it}| + 10C^e_{jt} \]

With

\[-7 \leq C_{it} \leq 7\]

(3.3)

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):

\[ C^e_{jt+1} = \gamma_i C_{it} + (1 - \gamma_i)C^e_{jt} \]

With \( 0 \leq \gamma \leq 1 \)

(3.4)

From these equations it follows that the optimal contribution for player \( i \) depends on both the parameter \( \gamma_i \), indicating how strong agent \( i \) believes he can influence agent \( j \), and \( \alpha_{ijt} \), which expresses the weight \( i \) assigns to the payoff for \( j \). We specify the latter as:

\[ \alpha_{ijt} = \delta_1 \alpha_{ijt-1} + \delta_2 I_{ijt-1} \]

(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 paper, though, we will use the next, more
general specification, which differentiates between positive and negative impulses:

\[
\alpha_{ijt} = \begin{cases} 
\delta_1 \alpha_{ijt-1} + \delta_2 N_t I_{ijt-1} & I_{ijt-1} \leq 0 \\
\delta_1 \alpha_{ijt-1} + \delta_2 P_t I_{ijt-1} & I_{ijt-1} > 0
\end{cases}
\]

(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). The tie mechanism described above can also be interpreted as an information extraction mechanism, used to determine if the other player is a friend or a foe (Bault et al., 2017). A higher positive and negative impulse parameters can then be seen as a preference for respectively cooperative or destructive behavior.

### 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 zero,\(^5\) and omitting the subscripts of \( \alpha \) and \( \gamma \), 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 \( 10\alpha \). If players expect to be able to influence their counterpart the cost-benefit analysis becomes more complicated. There are benefits of \( 10\gamma \), 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

\(^5\)In Appendix A also addresses the case where \( C_{it} \) is smaller than zero.
about the fact that the other faces a cost of reciprocating or imitating.

The following propositions can be proven: 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 \(\alpha\) 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 \(\alpha\) 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 \(\alpha\) parameter.

**Proposition 3.1.** Contributions outside of the interval \([-5, 5]\) can never be part of any equilib-
rium if \(-1 \leq \alpha \leq 1\).

**Proof.** See Appendix A.1.

Next, focusing first on symmetric equilibria \((C_{it} = C_{jt})\), we arrive at the following proposi-
tion:

**Proposition 3.2.** Any contribution level where \(C_{it} = C_{jt} \in [-5, 5]\) can be part of an equilib-
rium.

**Proof.** See Appendix A.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 con-
straints 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 3.3.** Asymmetric equilibria exist if either \(-5 \leq C_{it}, C_{jt} < 0\) or \(0 \leq C_{it}, C_{jt} \leq 5\).

**Proof.** See Appendix A.3.

Our last proposition establishes a parameter restriction for efficient cooperation. For con-
venience, we restrict ourselves here to the myopic model as this models turns out to be most
relevant in our study. Similar restrictions including $\gamma$ could be derived for the model that allows for forward looking behavior (see Appendix A.2).

**Proposition 3.4.** *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: $0.2 \leq \frac{\delta_2}{1-\delta_1} \leq 0.25$, which is a necessary but not a sufficient condition.*

Proof. See Appendix A.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 $\alpha$ 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 3.5.** *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 3.6.** *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. (2016) 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 3.7.** 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 fullfill the restrictions outlined in this proposition:

**Hypothesis 3.8.** For the majority of the subjects \(0.2 \leq \frac{\delta_2}{1-\delta_{11}} \leq 0.25\) will not hold.

### 3.3 Methods

#### 3.3.1 Experiment

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.
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.

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 horizontal axis and the payoff to Other is on the vertical 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 $\alpha$, as they indicate that a person is willing to give something up in order to decrease the payoff for Other. The tangent that results from the test can be interpreted as an indication of the initial alpha ($\alpha_0$) and will be used as such later on.

After this SVO test the participants played 35 rounds of the Fragile Public Good (FPG) 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 feedback they would receive during the actual game. After every round they saw the

---

6Situations where individuals prefer negative payoffs over positive payoffs for themselves are not taken into consideration here and very seldom observed.
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 in this paper.

3.3.2 Estimation

For our estimation procedure we follow Bault et al. (2016). To close the model and to enable us to estimate it we introduce a random variable $\epsilon_{ik}/\theta_i$ as a noise term. Where $\theta_i$ represents the rationality or choice intensity of player $i$. If we now assume $\epsilon_{ik}$ to be i.i.d. and double-exponentially distributed, we arrive at a multinomial logit model. Now let $\pi_{ikt}$ be the probability that a player chooses contribution $k$ in period $t$, then if we multiply all these probabilities we obtain our likelihood measure:

$$\prod_t \pi_{ikt} = \prod_t \frac{e^{\theta_i V_{ikt}}}{\sum_h e^{\theta_i V_{ih} t}}$$  \hspace{1cm} 0 < \theta < \infty \hspace{1cm} (3.8)

Estimation requires a value for $\alpha_0$, 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
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 $\alpha$ 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 calculation of the average.

A summary of the behavior during the FPG games is given below in table 1. Looking at table 1 there are some noteworthy results. First of all, we observe that average contributions are noticeably higher in the second FPG game than in the first FPG game:

<table>
<thead>
<tr>
<th>Game</th>
<th>FPG1 (n=130)</th>
<th>FPG2 (n=130)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average contribution</td>
<td>2.28</td>
<td>2.86*</td>
</tr>
<tr>
<td>Avg contribution first round</td>
<td>1.26</td>
<td>2.53*</td>
</tr>
<tr>
<td>Avg contribution last round</td>
<td>0.68</td>
<td>0.70</td>
</tr>
<tr>
<td>% negative contributions</td>
<td>11.3%</td>
<td>4.6%</td>
</tr>
</tbody>
</table>

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

This is also illustrated by figure 1, that shows the 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 2 shows two pairs that, respectively, establish a cooperative and a sour relationship.

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 4.2 investigates the group level estimation results, 4.3 is devoted to individual level estimates, while section 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 eq. (3.1), neglecting for the moment forward looking behavior introduced in eq. (3.2). This leaves a model with only 3 parameters. Despite its simplicity, this model has been found quite succesful in explaining public good contributions (Bault et al., 2016).

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 eqs. (3.6a) and (3.6b).

The third model that we will investigate is the forward-looking model (FL), represented by eqs. (3.2) and (3.4). And, 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 2.
Table 3.2: Group level estimations

<table>
<thead>
<tr>
<th></th>
<th>$\gamma$</th>
<th>$\theta$</th>
<th>$\delta_1$</th>
<th>$\delta_2$</th>
<th>$\delta_{2P}$</th>
<th>$\delta_{2N}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myopic</td>
<td>0.16*(0.02)</td>
<td>0.54*(0.09)</td>
<td>0.10*(0.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M.NP</td>
<td>0.17*(0.02)</td>
<td>0.49*(0.10)</td>
<td>0.12*(0.02)</td>
<td>0.08*(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FL</td>
<td>0.06(0.06)</td>
<td>0.16*(0.02)</td>
<td>0.54*(0.10)</td>
<td>0.10*(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FL. NP</td>
<td>0.03(0.06)</td>
<td>0.17*(0.03)</td>
<td>0.49*(0.12)</td>
<td>0.11*(0.02)</td>
<td>0.08*(0.02)</td>
<td></td>
</tr>
</tbody>
</table>

Note: standard errors are between brackets; * indicates significance at the 1% level.

The myopic model, as well as the other models, estimates $\theta$ to be around 0.16. To give an idea about its interpretation, the predicted chance that a player with an $\alpha$-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 $\theta$ would be 0 all choices have a probability of $\frac{1}{15}$ (< 7%), while if $\theta$ goes to infinity the probability of a player choosing zero goes to 1. $\delta_1$ 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. $\delta_2$ is estimated to be 0.1 which means that a contribution of 5 would lead to an $\alpha$ of around 0.5, if the initial $\alpha$-value is zero.

If we allow for a dichotomy between positive and negative impulses (M.NP) we find similar $\theta$ and $\delta_1$ values. However, $\delta_{2P}$ seems larger than $\delta_{2N}$. The improvement in the likelihood is significant at the 10% level (p≈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).7

We next consider the forward-looking model (FL), but neglect the potential difference between positive and negative impulses for the moment. We find that $\gamma$ is insignificant. Only a modest improvement in the likelihood is obtained when compared with the improvement caused

---

7It should be noted that the difference between $\delta_{2P}$ and $\delta_{2N}$ reported in table 2 is a conditional result. Not all participants in our experiment where 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 then pairs that do not make such contributions.
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, $\theta$, $\delta_1$ and $\delta_2$ 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 $\delta_1$ is again around 0.5.

An interesting result is that $\delta_1$ 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., 2016). The fact that this finding is replicated indicates 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 $\alpha$ 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 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 $\delta_{2P}$ and $\delta_{2N}$ 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 $\delta_2$-parameter(s). If this was the case it would lead to a bias in the $\delta_2$-parameter(s), where the effect on $\delta_{2N}$ would be negative while the effect on $\delta_{2P}$ 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 $\delta_{2P}$ being higher, while $\delta_{2N}$ 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 $\gamma$, 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 six out of the last 10 rounds both contribute equally and either 4 or 5 tokens). This is in line with proposition 4, stating that $0.2 < \frac{\delta_{2i}}{1-\delta_{2i}} < 0.25$ 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 we find -as we did on the group level- 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 $\gamma$ that is significantly different from zero. It turns out that 79 out of 130 participants are indeed forward looking ($\gamma$ 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 behavior. Moreover, we still find that for 88 out of the 130 participants \(0.2 < \frac{\delta_{2i}}{1-\delta_{1i}} < 0.25\), while most players do not seem to be able to sustain negative relationships, as for only three participants we find that \(0.2 < \frac{\delta_{2i}}{1-\delta_{1i}} < 0.25\). There are also ten participants that build excessively strong negative ties (i.e., \(\alpha\) 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 2. Looking at the evidence presented at the group as well as at the individual level we must reject H4.

### 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 \(\alpha\)).

In order to get truly 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 insignif-

\(^8\)Not all participants encounter many negative impulses though, so this result is conditional on encountering enough negative impulses

\(^9\)For more on this see Hoyer et al. (2014), where the occurrence of such relationships is further analyzed with different experimental designs. The results of all the estimations mentioned above are available online.
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 $\alpha$-values (for $\alpha_0$ 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 $\alpha$ 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 $\alpha$ 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 $\alpha$ and $\beta$ 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):

\[
V_{ikt}^e = X_{ikt}^e - \alpha (X_{jt}^e - X_{it}^e)_{X_{it} < X_{jt}} - \beta (X_{it}^e - X_{jt}^e)_{X_{it} > X_{jt}}
\]  

(3.9)

Where $X_{ih}^e (h = i, j)$ 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 $\beta$ to be larger than $\alpha$ for both specifications of $X_{ih}^e (h = i, j)$, 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 $\alpha$ and $\beta$ 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, $\phi$ 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 $t$ ($\pi_{ikt}$) 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:

$$\pi_{ikt} = \frac{q_{ikt}}{\sum_h q_{ihkt}}$$

$$q_{ikt+1} = (1 - \phi_i)q_{ikt} + E_{ikt}$$

$$E_{ikt} = s_i R_{jt} (1 - \epsilon_i) \text{ if } k = j$$

$$E_{ikt} = s_i R_{jt} (\epsilon_i / 2) \text{ if } k = j \pm 1$$

$$E_{ikt} = 0 \text{ otherwise}$$

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 ($\alpha_i$), the belief about the altruism parameter of the other ($\alpha_j$) and a parameter $\lambda$ that weighs both, such that $i$’s utility ($u_i$) gets transformed into an extended utility, $v_i$:

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

(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 ($\alpha$). 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 presents the mean absolute error and the mean squared error of the predictions:

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean Absolute Error</th>
<th>Mean Squared Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ties Model</td>
<td>0.51 (0.53)</td>
<td>1.65 (2.39)</td>
</tr>
<tr>
<td>Fixed Alpha</td>
<td>1.92 (2.41)</td>
<td>6.97 (14.05)</td>
</tr>
<tr>
<td>Roth and Erev</td>
<td>1.64 (1.07)</td>
<td>4.39 (4.59)</td>
</tr>
</tbody>
</table>

Note: standard errors using average errors per individual between brackets
From our results in table 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 4 shows the results.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean Absolute Error</th>
<th>Mean Squared Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ties Model</td>
<td>1.00 (1.11)</td>
<td>3.58 (5.61)</td>
</tr>
<tr>
<td>Fixed Alpha</td>
<td>1.69 (1.72)</td>
<td>6.11 (12.68)</td>
</tr>
<tr>
<td>Roth and Erev</td>
<td>3.01 (1.47)</td>
<td>11.63 (7.21)</td>
</tr>
</tbody>
</table>

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 the reinforcement learning model, and especially the reinforcement learning model, performs 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 precise.

### 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 studies by Dal Bó and Fréchette (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, if cooperation becomes more profitable, people become ‘slower to anger’ and ‘faster to forgive’, this is, they are more willing to allow a defection and pick up cooperation after only a few cooperative choices of the counterpart. 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, and their estimates, can generate (or mimic) these strategies as well as a the shift towards more lenient and 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 $\alpha$-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 5a), we apply these other-regarding preferences to two specific games with benefit/cost (b/c) ratios of 2 (Table 5b) and 4 (Table 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 $\alpha$ is larger than 1/2 (in Table 5b) or 1/4 (in Table 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)

<table>
<thead>
<tr>
<th></th>
<th>(a) b/c</th>
<th>(b) b/c=2</th>
<th>(c) b/c=4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>D</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>b-c</td>
<td>-c</td>
<td>1+1α</td>
</tr>
<tr>
<td>D</td>
<td>b</td>
<td>0</td>
<td>2-α</td>
</tr>
</tbody>
</table>

Note: Table 5a gives the actual payoffs of player 1, while 5b and 5c 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 Fréchet (2011) and Fudenberg et al. (2012) find experimental evidence that many subjects in their experiments use either a tit-for-tat (TFT) or a tit-for-two-tats (TF2T) strategy although these strategies are often not evolutionary stable (in a evolutionary game theory context).\(^\text{10}\)

Below we will show that the simple and neurobiologically underpinned Ties model (Bault et al. (2017)) 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 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 game 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 between different specifications of the same game. While the previously mentioned studies do not attempt to explain why subjects use different strategies within the same game environment, 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 estimated Ties model allows us to see if the tie mech-

\(^{\text{10}}\)TFT 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. Dal Bó and Fréchet estimate that for their games with a continuation probability of 3/4 between 35% (if b/c≈2) and 56% (if b/c≈4) of subjects choose TFT. Fudenberg et al., who consider many more strategies and introduce noise, find between 19%, for b/c≈1.5, and 7%, for b/c=4, of players choosing TFT in contrast 20% players choose TF2T when b/c was 4, only 5% of players chose TF2T when b/c was 1.5 as can be found in table 7. Dal Bó and Fréchette do not consider TF2T.
anism and resulting strategies are consistent across different, albeit related, games regarding social dilemmas.

An example of a strategy that is easily generated by the tie mechanism is the well-known 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 following is required: First, a player should have a strong enough impulse parameter \( \delta_2 \) (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 tolerates deviations by the other player. Hence, the tie-persistence parameter \( \delta_1 \) must be sufficiently small. Finally, \( \alpha_0 \) 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 \( \delta_1 \) and \( \delta_2 \), play starts to mimic often reported strategies like AD, TFT, and TF2T. However, a much more challenging task is to use the parameter value estimates of this paper 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 strategy: ‘qualified’ tit-for-two-tats (QTF2T). This strategy is similar to TF2T in all but one respect: it requires more than one cooperative choice by the other before defection (D) is forgiven. In terms of the Ties model this means that first the value of \( \alpha \) has to be significantly built up; in this case, described in appendix C.3, until the theoretical maximum \( (\delta_2 / (1 - \delta_1)) \), but any value corresponding to any number of consecutive cooperative decisions can be chosen. The intuition behind this ‘qualification’ is that players 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 estimates found in this study and those
relevant for a PD environment we need to normalize the impulse, $I$. For, note that if we multiply the impulse by a factor $i$, the estimate of $\delta_2$ will change with 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 $I_n \equiv \frac{C_j - C_{ref}}{C_{eff} - C_{ref}}$. Thus, in order to translate the values we found for $\delta_2$ to values suitable for a PD game environment, where choosing $C$ (the efficient choice) is valued as 1, we multiply $\delta_2$ by 4. As before $\alpha_0$ will stand for the starting value of $\alpha$.

Below there are two graphs, the first for $b/c = 2$ and the second for $b/c = 4$, which show which parameter values of the tie mechanism ($\delta_1$ and $\delta_2$) correspond to which strategies\(^{11}\). 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 $\delta_2P$ (as only positive impulses are possible in this environment) and $\delta_1$ values, estimated with the myopic model that allows for a dichotomy between positive and negative impulses.

\(^{11}\)For characteristics of these strategies and a more elaborate analysis, see Appendix C.
Figure 3.3: Parameter estimates and strategies for b/c=2
\[ \delta_2 = \frac{1}{4} \]
\[ \delta_1 \delta_2 (1 - \delta_1) = \frac{1}{4} \]
\[ \delta_1 \delta_2 (1 - \delta_1) = 1 \]
\[ \delta_2 (1 - \delta_1) = \frac{1}{4} \]
\[ (\delta_1 \delta_2 (1 - \delta_1)) = \frac{1}{4} \]

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 in our Fragile Public Good Game to destroy the public good, which could make players reluctant to not contribute, out of fear for punishment. This is highlighted by the fact that if we use the estimates from the myopic model that does not allow for a difference in impulse impact there are some more AD players, as $\delta_2$ 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 $\alpha$-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 $\alpha$ value must have been relatively high, while if a player played D this value must have been relatively low. Now, because the tie ($\alpha$) of the former player will decay (as it gets multiplied by $\delta_1$) the chance that this player chooses C declines. In contrast, the tie of the other player will be reinforced (with $\delta_2$) by counterpart’s cooperative action in the previous round. Consequently, the $\alpha$ 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 converge, making the finding of Breitmoser 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 estimation results of our Fragile Public Good Game are very much in line with earlier studies of public goods games using the ties model. More specifically, we also observe that the memory component of the tie mechanism, represented by the tie persistence parameter $\delta_1$, is about equally important as the impulse component, represented by the parameter $\delta_2$, with the scalefree $\delta_1$ 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. Apperently, the bad is not stronger than the good in this context cf. (Baumeister et al., 2001). This asymmetry in the tie mechanism is helpfull 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 well-known reinforcement learning model as well as a model with constant social preferences. For both the learning model 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 to be 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 in the PD game. Our alternative explanation for the behavior in repeated PD games does not require people to switch strategies in a seemingly ad hoc way.
Appendix 3.A  Proof of Propositions

3.A.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:

\[
V(C_{it}) - V(C_{it} + 1) = 2C_{it} + 2 - 10\alpha + \gamma(-10 + \alpha(2\gamma C_{it} + 2(1 - \gamma)C_{jt} + \gamma + 1)) \tag{3.12}
\]

Where \(\gamma\) is between 0 and 1. For the proposition to be true this equation must be positive. We reformulate this condition to:

\[
2C_{it} + 2 + 2\alpha\gamma\left(\gamma C_{it} + (1 - \gamma)C_{jt} + \frac{1}{2}\gamma + \frac{1}{2}\right) > 10(\alpha + \gamma) \tag{3.13}
\]

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

\[
2C_{it} + 2 + 2\alpha\gamma\left(C_{it} + \frac{1}{2}\gamma + \frac{1}{2}\right) > 10(\alpha + \gamma) \tag{3.14}
\]

So at \(C_{it} = 5\) we have:

\[
12 + 2\alpha\gamma(5 + \frac{1}{2}\gamma + \frac{1}{2}) > 10(\alpha + \gamma) = > \quad 12 + \alpha(\gamma(11 + \gamma) - 10) - 10\gamma > 0 \tag{3.15}
\]

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

\[
12 + 11\gamma + \gamma^2 > 10(\gamma + 1) \tag{3.16}
\]
Which is always the case. If $\alpha$ is -1 we have:

\[
12 - 11\gamma - \gamma^2 > 10(\gamma - 1)
\]  

(3.17)

Now since (3.15) is a monotone function in $\alpha$ 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:

\[
-12 + 2\alpha\gamma(-6\frac{1}{2} + \frac{1}{2}\gamma) > 10(\alpha + \gamma)
\]  

(3.18)

This is never true for positive $\alpha$’s. For $\alpha$ is -1 we get:

\[
-12 + \gamma(13 - \gamma) > 10(\gamma - 1)
\]  

(3.19)

This cannot be for true for $\gamma$ 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:

\[
12 + 2\alpha\gamma(5\gamma + (1 - \gamma) + \frac{1}{2}\gamma + \frac{1}{2}) > 10(\alpha + \gamma)
\]  

(3.20)

Again this is always true for $0 \leq \alpha \leq 1$ (and if the other contributes positively, $\alpha$ 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.A.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 \(\alpha\) 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 (\(\gamma\) is zero).

**Myopic Agents**

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

\[
2C_{it} + 2 > 10\alpha \tag{3.21}
\]

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

\[
\alpha = \delta_1 \alpha + \delta_2 I \tag{3.22}
\]

For \(I\) we use the Nash equilibrium as a reference point as we did throughout the paper. This leads to:

\[
\alpha = \delta_1 \alpha + \delta_2 C \tag{3.23}
\]

or

\[
\alpha = \frac{\delta_2 C}{1 - \delta_1}
\]

From (3.21) we know that:

\[
0.2C < \alpha < 0.2(C + 1) \tag{3.24}
\]
Combining (3.23) and (3.24) we obtain:

\[0.2 < \frac{\delta_2}{1 - \delta_1} < 0.2 + \frac{0.2}{C}\]  \hspace{1cm} (3.25)

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

\[2C > 10\alpha\]  \hspace{1cm} (3.26)

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

\[0.2 > \frac{\delta_2}{1 - \delta_1} > 0.2 - \frac{0.2}{C}\]  \hspace{1cm} (3.27)

If $C=0$, then the stimulus is zero. This will lead to the value of $\alpha$ moving gradually towards zero as well. As the best response to an $\alpha$-value of zero is to play 0 we have that the [0,0] equilibrium can always exist regardless of the $\delta$-parameters. To the entire range of $\alpha$-values wherefor a contribution of zero is a vest 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 $\gamma$ is unequal to zero all values of $C_{it}$ are still part of symmetric equilibria, but the condition on $\delta_1$ and $\delta_2$ becomes stricter. Just as in the previous case we start from (3.14) and fill in (3.23):
\[2C + 2 + 2 \frac{\delta_2 C}{1 - \delta_1} \gamma(C + \frac{1}{2} \gamma + \frac{1}{2}) > 10\left(\frac{\delta_2 C}{1 - \delta_1} + \gamma\right)\] (3.28)

\[2C + 2 > (10 - 2\gamma(C + \frac{1}{2} \gamma + \frac{1}{2}))\frac{\delta_2 C}{1 - \delta_1} + 10\gamma\]

\[\frac{2C + 1 - 10\gamma}{(10 - 2\gamma(C + \frac{1}{2} \gamma + \frac{1}{2}))C} > \frac{\delta_2}{1 - \delta_1}\] (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:

\[\frac{2C - 10\gamma}{(10 - 2\gamma(C - 1 + \frac{1}{2} \gamma + \frac{1}{2}))C} > \frac{\delta_2}{1 - \delta_1}\] (3.30)

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

3.A.3 Proposition 3: Asymmetric equilibria exist if \(C_i C_j > 0\) and \(|C_i| \leq 5\) and \(|C_j| \leq 5\)

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

\[\alpha = \frac{\delta_2 C_j}{1 - \delta_1}\] (3.32)

And (3.24) changes into:

\[0.2C_i < \alpha < 0.2(C_i + 1)\] (3.33)
Leading to:

\[ 0.2 \frac{C_i}{C_j} < \frac{\delta_{2i}}{1 - \delta_{1i}} < 0.2 \frac{C_i}{C_j} + 0.2 \frac{C_j}{C_i} \]  \tag{3.34}

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

\[ 0.2 \frac{C_j}{C_i} < \frac{\delta_{2j}}{1 - \delta_{1j}} < 0.2 \frac{C_j}{C_i} + 0.2 \frac{C_i}{C_j} \]  \tag{3.35}

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

\[ 0.04 < \frac{\delta_{2i}}{1 - \delta_{1i}} < 0.08 \]  \tag{3.36}

And:

\[ 1 < \frac{\delta_{2j}}{1 - \delta_{1j}} < 2 \]  \tag{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 \( \alpha \) in the other and contributing positively can not be an optimal choice under a negative \( \alpha \).

3.A.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 a 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.
Appendix 3.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 impulses:

\[ I_{ijt} = C_{jt} - C_{ref}^i \]  

(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 text, 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: \( C_{ref}^i = C_{it} \) if an agent’s own contribution is used and \( C_{ref}^i = C_{jt-1} \) if one looks at the contribution of the other in the previous round. In the last two cases we initialize the system using \( C_{ref}^i = 0 \). In the table below the parameters from estimating the model using different reference points are shown.

<table>
<thead>
<tr>
<th>Reference point</th>
<th>( \delta_1 )</th>
<th>( \delta_2P )</th>
<th>( \delta_2N )</th>
<th>( \sum LL )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nash</td>
<td>0.490</td>
<td>0.115</td>
<td>0.080</td>
<td>-8545</td>
</tr>
<tr>
<td>Pareto</td>
<td>0.975</td>
<td>0.108</td>
<td>0.000</td>
<td>-11265</td>
</tr>
<tr>
<td>Own Contribution</td>
<td>1</td>
<td>0.003</td>
<td>-0.006</td>
<td>-11598</td>
</tr>
<tr>
<td>Contribution other</td>
<td>0.989</td>
<td>0.179</td>
<td>0.160</td>
<td>-10826</td>
</tr>
</tbody>
</table>

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 \( \delta_1 \) 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 \( \alpha \) 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 $\alpha$ high, which is necessary for positive contributions to occur, is for $\delta_1$ 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 $\alpha$ 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 $C^{\text{ref}} = 0$ 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 $\alpha$ is basically determined by the change in contributions since the initial value of $\alpha$ diminishes over time. If we start from the basic tie mechanism described in (3.5) we have (with $I = C_{jt}$):

$$
\alpha_{ij2} = \alpha_{ij1}\delta_{1i} + \delta_{2i}C_{j1} \implies \alpha_{ij3} = \alpha_{ij1}\delta^2_{1i} + \delta_{1i}\delta_{2i}C_{j1} + \delta_{2i}C_{j2} \implies \\
\alpha_{ijt} = \alpha_{ij1}\delta^{t-1}_{1i} + \delta^{t-2}_{1i}\delta_{2i}C_{j1} + \ldots + \delta_{1i}\delta_{2i}C_{jt-2} + \delta_{2i}C_{jt-1} \implies
$$

$$
\alpha_{ijt} \approx \frac{\delta_{2i}C}{1 - \delta_{1i}} \quad (\text{for } t \to \infty)
$$

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

$$
\Delta \alpha_{it+1} = \alpha_{it+1} - \alpha_{it} = \delta_{1i}\alpha_{it} + \delta_{2i}C_{jt} - \alpha_{it} \implies \\
\Delta \alpha_{it+1} = \delta_{2i}C_{jt} - (1 - \delta_{1i})\alpha_{it} \approx \delta_{2i}C_{jt} - (1 - \delta_{1i})\frac{\delta_{2i}C^{eq}}{1 - \delta_{1i}} = \delta_{2i}(C_{jt} - C^{eq})
$$

59
In the symmetric equilibria (the most commonly observed equilibria in our experiment) we have that, when the reference point with respect to $\alpha$ is fixed, the change in $\alpha$ approximates a linear function of the change in the other player’s contribution.

**Appendix 3.C  Tie Model Parameters for different Repeated PD Strategies**

**3.C.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 $\alpha$-values will change in the analysis below if one uses other $b/c$ ratios):

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

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>3,3</td>
<td>-1, 4</td>
</tr>
<tr>
<td>D</td>
<td>4,-1</td>
<td>0,0</td>
</tr>
</tbody>
</table>

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

$$V_{it} = U_{it} + \alpha_{ijt}U_{jt}$$  \hspace{1cm} (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 $\alpha_{ijt}$ represents the weight $i$ attaches
to the utility of $j$ ($i$’s tie with $j$) in period $t$, which is updated as follows:

$$\alpha_{ijt} = \delta_1 \alpha_{ijt-1} + \delta_2 I_{t-1}$$

(3.42)

With $\alpha_{ij1}$ 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 $\alpha$. If $\alpha$ is larger than $1/4$ cooperating is a dominant choice while for $\alpha$ 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 neurological underpinned Ties model (Bault et al., 2017, 2016) could explain the behavior described by these strategies.

### 3.C.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 $\delta_1$, $\delta_2$, and $\alpha_{ij1}$ for which this behavior is sustained. First, we investigate what levels of $\alpha$ can be reached after continuous play of either $C$ or $D$. Since $\delta_1 < 1$, being exposed to infinitely repeated defection leads to $\alpha$ going to zero. Now, what about continuous play of $C$? Noting that the initial value of $\alpha$ vanishes as $t$ goes to infinity, we start from the expression for $\alpha$ in period 2 and iterate:

$$\alpha_{ij2} = \delta_1 \alpha_{ij1} + \delta_2 I_1 \implies \alpha_{ij3} = \delta_1^2 \alpha_{ij1} + \delta_1 \delta_2 I_1 + \delta_2 I_2 \implies$$

$$\alpha_{ijt} = \delta_1^{t-1} \alpha_{ij1} + \delta_1^{t-2} \delta_2 I_1 + \ldots + \delta_1 \delta_2 I_{t-2} + \delta_2 I_{t-1}$$

(3.43)

Furthermore, if $C$ is played, $I$ is always equal to one, so that:
\[ \alpha_{ijt} = \delta_{1i}^{t-1} \alpha_{ij1} + \delta_{1i}^{t-2} \delta_{2i} + \ldots + \delta_{1i} \delta_{2i} + \delta_{2i} :=> \]
\[ \alpha_{ijt} \approx \frac{\delta_{2i}}{1 - \delta_{1i}} \quad \text{(as } t \rightarrow \infty) \]  

(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 \( \alpha \), 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):

\[ \alpha_1 \geq 1/4 \]  

(3.45)

We also need that, no matter how high the current level of \( \alpha \) 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:

\[ \frac{\delta_1 \delta_2}{1 - \delta_1} \leq 1/4 \]  

(3.46)

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

\[ \delta_2 \geq 1/4 \]  

(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.C.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:

$$\delta_1 \delta_2 \geq 1/4$$

(3.48)

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

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

(3.49)

Combining (3.48) and (3.49) gives:

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

(3.50)

So we have:

$$\delta_2 \geq 1/2$$

(3.51)
A potential problem for the result above is that the upper bound of $\alpha$, $\frac{\delta_2}{1-\delta_1}$, 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):

$$\frac{\delta_2}{1-\delta_1} = \frac{\delta_2\delta_1}{\delta_1(1-\delta_1)} \geq \frac{1}{4(\delta_1(1-\delta_1))} => \frac{\delta_2}{1-\delta_1} \geq 1$$

(3.52)

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

$$\frac{\delta_1\delta_2}{1-\delta_1} \geq 1/4$$

(3.53)

**Appendix 3.D Payoff Matrix and SVO Example**

![Payoff Matrix](image)

Figure 3.5: Payoff Matrix
Figure 3.6: SVO example
Chapter 4

Emotional Leadership in an Intergroup Conflict Game Experiment*

4.1 Introduction

The ability of leaders to make others follow their lead is of utmost importance to them. Whether it concerns the ability of politicians leading their citizens into, or away from, war or that of managers leading their teams into competition, it often requires not just a compelling rationale but also an emotional address as well as an example of direction by the leader. For example, many politicians combined these latter two responses in their speeches after the recent tragic shootings in Paris and San Bernadino. They evoke emotions of anger towards the enemy or other party, while they (try to) convey a message of personal commitment, restraint, and hope. Perhaps the best example of a leader leading-by-example as well as using emotional leadership might be Dr. Martin Luther King, Jr. who not only inspired and advanced the U.S. civil rights movement with his speeches but also led activists in marches, demonstrations and sit-ins. These

*This chapter is based on joint work with Frans van Winden, published in the Journal of Economic Psychology; see Loerakker and van Winden (2017). Financial support of the Research Priority Area Behavioral Economics of the University of Amsterdam is gratefully acknowledged.
two forms of leadership, leading-by-example and emotional leadership are investigated in this experimental study.

Hermalin (1998) shows that leading-by-example might be a useful signaling tool if the leader knows more about the environment than his or her followers. For sequential public good games Potters et al. (2007), Gächter et al. (2010) and Arbak and Villeval (2013) find experimental evidence that leading-by-example indeed leads to more cooperative outcomes. This effect is less strong however if the leader does not possesses more information than the follower(s). To the best of our knowledge, as yet no experimental work has been done on the effect of leadership in the game that we will study in this paper, a conflict game (related to Hirshleifer (1988) and Skaperdas (1992); for surveys, see: Garfinkel and Skaperdas (2007) and Abbink (2012)).

This is somewhat surprising given the natural link between group conflict and leadership.

Humphrey (2002) describes the importance of emotional leadership and states that "a key leadership function is to manage the emotions of group members" (Humphrey, 2002, p.498). Furthermore, this meta-study finds that the emotional display of leaders has a significant influence on the behavior of followers. This research concers a work environment, showing the importance of these mechanisms in that environment, but leaves open questions regarding causality. It makes clear, though, that if we want to get a better understanding of real leader-follower relationships, it is important to know how leadership is obtained. As Humphrey notes, leaders have (or are perceived to have) different characteristics than followers. As Hermalin does, he stresses the importance of leadership being a voluntary activity that is often costly to obtain. To mimic such a situation, we will use the auction design of Concina and Centorrino (2013). In their design the position of a leader is auctioned off, using a second-price sealed-bid auction, so that the one who wants to be the leader most will be the leader during the experiment.

In our experiment subjects play a repeated conflict game as member of a fixed group. We

---

1Conflict games are related to rent-seeking games based on Tullock (1980), where typically a fixed prize is contested and the winner (only) loses the resources that s/he spent on the contest (for surveys, see: Konrad (2009); Sheremeta (2017) (on group contests), and Dechenaux et al. (2015) (on experiments). The main difference is that in a conflict game all the endowments are at stake (no private account is available as a safe haven) as the winner can take all net of conflict expenditures. This makes the prize endogenous, because both the winner’s and the loser’s resources spent on conflict are wasted.
allow leaders to have some emotional control over their followers by having them evoke a basic (or neutral) emotion in their followers. These emotions are evoked with the use of specially selected and validated video clips that are displayed in between the repetitions of the game. Besides this emotional control, leaders can also lead by example as they move first in every period of the conflict game. To get some better insights in the experimental results we develop and analyze different behavioral models and compare their predictions with the data.

Overall, we do not find a significant treatment effect as groups with a leader who can use this emotion evoking mechanism do not contribute more (or less) than groups without this emotional control. Besides this, our main findings are that emotional leadership, but especially leading-by-example is an effective instrument for a leader to influence the behavior of group members, and, furthermore, that the behavioral data can be best explained by models including either imitation (Cartwright and Patel, 2010), psychological costs (Dufwenberg et al., 2011) or affective ties (van Dijk and van Winden, 1997).

The organisation of the paper is as follows. In section 2 the game will be introduced and the main theory is developed and analyzed. Section 3 presents the experimental design, while section 4 goes into the results. Section 5 closes with a concluding discussion.

4.2 The Game

In comparison to extant conflict games (see, e.g., Abbink, 2012), our game is novel in that it introduces groups (of four) with leaders. All players are endowed with $Y$ and have to decide on their contribution to their group’s conflict (where we will focus on the contribution of a generic member $i$ of one of the groups: $C_i$). One player – the leader – first makes a contribution that is observed by the other group members – the followers – who subsequently and simultaneously make their contribution decisions. The winner of the conflict is determined by a lottery, where the odds that the group of $i$ wins, $P(Win_i)$, is determined by the relative investment of resources into the conflict: $P(Win_i) = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j}$. The group that wins the conflict gets all the
remaining resources. More specifically, player \( i \) of the winning group keeps: \( Y - C_i \) and gets, in addition, one-fourth of the remaining resources of the other group: \( Y - \frac{1}{4} \sum_{j=1}^{4} C_j \). Members of the losing group are left with nothing. If both groups contribute zero to the conflict (\( C = 0 \) for all players), however, there is no conflict and every player keeps his or her endowment \( Y \). This outcome is Pareto efficient, but not a best response as every player would guarantee its group to win by making a minimal contribution. The expected payoff of player \( i \) \((E\pi_i)\) thus equals (if \( C \neq 0 \) for at least one of the players):

\[
E\pi_i = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} (Y - \frac{1}{4} \sum_{j=1}^{4} C_j + Y - C_i)
\]

(4.1)

Appendix A.1 shows that the standard Nash equilibrium, assuming risk neutrality, entails that a leader’s contribution \((C_L)\) equals: \( C_L = 0 \), while a follower’s contribution \((C_F)\) equals: \( C_F = \frac{8}{31} Y \). As a consequence, the expected payoff of a leader is: \( \frac{28}{31} Y \), while that of a follower equals: \( \frac{24}{31} Y \). It might seem counterintuitive that the leader contributes less than the followers, let alone zero. By contributing first, however, the leader is able to free-ride on the contributions of the others, who will have an incentive to contribute more due to the low (zero) contribution made by the leader. This is an interesting finding in itself, as leaders are typically seen as giving the good example. These predicted contributions (labeled Nash below) are notably different from the ones predicted for a completely simultaneous game with no leader (labeled Nash Simultaneous), where all players contribute \( C = \frac{1}{5} Y \) and player’s expected payoffs are \( \frac{4}{5} Y \).2

In the rest of this section we further discuss the role of leaders and derive predictions for different behavioral models.

2Appendix A.2 presents an equilibrium analysis for risk-averse players. We do so for the more tractable case of the simultaneous game, as the intuition is the same for both the sequential as well as the simultaneous game, namely, that agents will contribute more when they dislike risk, to decrease the chance of losing everything.
4.2.1 Leaders, ties, and psychological preferences

Typical findings of group contest experiments are that players contribute more to the contest than the Nash equilibrium prediction and, in a repeated game, players start to contribute more after a loss. Both of these findings are confirmed in our experiment. Interestingly, also the effect of leadership in a contest game has been studied before. Eisenkopf (2014) finds that the advise of managers that have only an advisory role, as a third party, in a two person contest game has a significant effect on contributions to the contest. This might be an indication that also in our experiment followers may not contribute less when they observe a higher contribution by their leader, as is derived from the best response function above, but more.

The fact that in this study the leader contributes before the others and that this contribution is observable for followers (own group members) allows the leader not only to free-ride at the expense of the followers, as predicted by the standard Nash equilibrium, but might also allow him or her to lead-by-example. By leading-by-example we mean that higher contributions from the leader could potentially influence the contributions made by their followers in a positive way as is, for instance, observed by Güth et al. (2007), Potters et al. (2007), Gächter et al. (2012), and Concina and Centorrino (2013) in public good game experiments, who find that if leaders contribute more to a public good so do the followers. In these settings the leader typically contributes more than his or her followers, in contrast to the Nash prediction for these games. This may be due to followers believing that the leader has more or better information about the situation, due to a social norm induced by the leader, or because the followers feel emotionally attached to their leader (Hermalin, 1998, Popper, 2000).

Not only do followers bond with their leader there is also evidence that leaders care for the outcomes of their group and feel pride, when their group performs well, or guilt, when their group performs poorly (Arbak and Villeval, 2013, Concina and Centorrino, 2013). To understand why following the leader might be profitable for the followers and why the leader has an incentive to not just free ride on the contributions of the followers by letting them compensate for his or her lack of contributions, it is important to note that the game, when observed from
within the group, is like a public good game with an interior Nash equilibrium. Beyond a certain point, further contributing is good for the other members of the group but not a best-response for an individual. In the rest of this subsection we will analyze different behavioral models that could be relevant for the game described above and indicate what behavior they predict. The exact predictions as well as the derivation of these predictions can be found in appendix A.

In sequential games, like ours, reciprocity can be an important element; see, for instance, Cox (2004) and Fehr and Gächter (2000). Dufwenberg and Kirchsteiger (2004), elaborating on the work of Rabin (1993), developed a model for sequential games where players reciprocate kind actions. In this model whether or not an action is seen as kind depends on a player’s belief of what the other player(s) will do, and a reference point or action (that can be seen as a neutral action). Given the large number of potential choices by all the players it is unclear to us, though, what exactly the predictions of this model would be for our game.

Another model that allows for the dynamics described above is a simplified version of the ties model by van Dijk and van Winden (1997). This model is especially relevant in the emotional context of our experiment as the reciprocity (or lack thereof) is caused by a mechanism of affective impulses. These impulses are generated by the behavior of the other player(s). If their behavior is better than what is taken as a reference point, this player will start to weigh the utility of the respective player more when taking a decision. This weight is typically represented by a value $\alpha$. This fits in nicely with the followers in our game, who might not only care about their own payoff but may also (start to) care about the payoffs of the rest of the group, through a tie with a caring leader. How much they would care depends on the contribution of the leader.

A model that can, in principle, generate similar behavioral predictions as the affective ties model would be a model with psychological preferences as in Dufwenberg et al. (2011) and McCannon (2015). In these models players are assumed to suffer from guilt aversion or psychological costs from deviating from a certain norm. The leader faces costs that are a (linear) function of the distance between his or her contribution and an exogenous reference point. The followers face similar cost but their reference point is the contribution of the leader. Social pref-
ference models like the inequality aversion models of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000), could potentially also lead to higher contributions by the leader but would draw the equilibrium towards the equilibrium of the simultaneous game as leaders would like to contribute more (to diminish inequality), while followers would like to contribute less.

What makes both the social ties model and the psychological preferences model distinct from these fixed social preference models in this context is that the contribution of the leader could positively affect the contribution of the followers as it creates a bonding (in the ties model) or sets a reference point (in the psychological preferences models). This could potentially overcome the ‘substitution effect’ of followers contributing more if the leader contributes less, and result in leaders contributing more instead of less. Another finding that would support these models over the other models of social preferences would thus be a positive relationship between the change in the contribution of the leader and that of the followers (when controlling for history), as a larger impulse generated or a higher reference point set by the leader should lead to higher contributions by the followers.

Another potentially relevant model is by Cartwright and Patel (2010). Their model distinguishes three types of players: strategists, imitators and independents, in a sequential public good game, but the model translates to our game as well. In their model first movers are assumed to have an incentive to contribute if there are enough imitators in the subject pool. Strategists will use this and would want to contribute early to set a high standard. This can be translated to our game assuming that the strategists are the leaders who contribute relatively much. Some of the followers (the imitators) will then follow suit, while independents always do the same regardless of the contribution of the leader. This could generate high contributions by leaders and a positive relationship between the contribution of the leader and that of the followers.

4.2.2 Emotional leadership

Salovey and Mayer (1990) state that managing the emotions of followers is an important task of managers. One reason they mention for this is that followers look at leaders for a sense of
direction in uncertain times. As participants in the conflict game may find it difficult to define an optimal strategy, this may hold for our setting as well. Law et al. (2004) find that positive emotions appear to correlate with better job satisfaction. Furthermore, it is interesting to note that more empathetic team members seem more likely to become a leader. Together with the observation by Myers (2015) that prosocials are more likely to participate in collective action, this is suggestive of leaders caring more for the group, and in line with a higher $\alpha$-value in the affective social ties model. For a review on (strategic) emotional leadership, see Humphrey (2002).

### 4.3 Experimental Design

Participants played 12 periods of the conflict game in fixed groups, of 4 participants. In every period they started with an endowment of 20 tokens. The exchange rate for tokens was 10 tokens = 1 euro. Before the first, fourth, seventh and tenth period a short video clip, taking less than 4 minutes, was shown. The content of those clips varied per treatment. After the instructions (reproduced in appendix B) but before the twelve periods, an auction determined who would be the leader of the group. This was a second-price sealed-bid auction, as in Concina and Centorrino (2013), with all participants getting 20 tokens to bid. If a bid was successful 20 tokens minus the number of tokens bid by the second highest bidder were added to the earnings of this participant, otherwise 20 tokens were added. Players were paid for every round they played.

The experiment took place in April and May 2015 at the CREED laboratory of the University of Amsterdam. In total, 240 participants took part in 16 sessions. Another session had to be ended early because of software problems and is therefore neglected here. Average earnings amounted to 14.90 euros.
4.3.1 Treatments

In the Baseline treatment (Baseline) the video clips shown to all subjects were neutral. Leaders and followers saw the same clips in the same period and leaders had no other task then to contribute first. This allowed them to (only) lead by example in a neutral environment. As in all other treatments, players were informed after every round about the contributions of their individual team members, the total contribution of the other group as well as their own earnings. All treatments had 80 participants.

Besides contributing first, leaders in the Strategic Emotion (SE) treatment could also decide what emotion to induce in their followers. They could choose a basic emotion or the neutral state, but could not see or choose the actual video clip used to induce this emotion, before the first, fourth, seventh, and tenth period. The leader always saw neutral movie clips, regardless of the choice he or she made. This allowed the leader to not only lead by example but to also strategically induce emotions in an attempt to ’steer’ followers.4

To see if the effects found in SE, as compared to Baseline, were due to the fact that followers were emotionally aroused or because the leader indeed manipulated the behavior of the followers by selecting specific emotions, we used a Random Emotion (RE) treatment as control. In this treatment the role of the leader was exactly the same as in Baseline, the only difference being that the followers would see emotion inducing movie clips, which followed the same sequences as were chosen by leaders in the SE treatment. This way we could disentangle the effect of a more emotional environment from the effects of emotions strategically used by leaders. It is important to keep in mind that subjects in all treatments were informed about their roles and the kind of videos they could get to see or choose.

---

4In the instructions of the experiment all participants got to see a trailer that included short samples (samples took about 10 to 12 seconds per sample, with one sample for every emotion and one for the neutral state) of some of the clips used in the experiment in order to give the participants an idea about the nature of the clips used.
4.3.2 Emotion evocation

According to Gross and Levenson (1995), the two main criteria for a good emotion evoking movie are *intensity* and *discreteness*. The former indicates how well the targeted emotion is induced, while the latter informs us about the difference in strength between the inducement of the targeted emotion and that of the second strongest (basic) emotion induced. The idea is that it is not so bad if next to happiness also similar emotions as joy and amusement are evoked, but that things become messy when together with anger also sadness and disgust are induced, as these are also basic emotions. Most of our tested clips reached the criteria of both *intensity* and *discreteness*. There were problems, however, with anger as well as the neutral state. Most anger clips did either not evoke enough anger or induced a wide variety of negative emotions, while neutral movies evoked too much happiness in 2 out of 4 instances. The exact criteria for *intensity* and *discreteness* as well as the full list of clips and the instructions for the validation experiment can be found in appendix C.5.

As mentioned before, video clips were shown that intended to evoke a certain basic emotion in both the Strategic Emotion (SE) treatment as well as the Random Emotion (RE) treatment. In order to test what videos indeed evoked the targeted basic emotion a test protocol was developed. Our fully computerized protocol (see Appendix C) was based on the procedures described by Ray (2007) and Gross and Levenson (1995), from whom we also used many of their recommendations for video clips. Other video clips were suggested by Bartolini (2011). This testing procedure took place in April 2015 at the CREED laboratory. In total 87 participants saw 12 movies each. This lasted for around 45 to 50 minutes for which they were compensated with 10 euros.

---

5In the end we used only two videos for both anger (although the second anger video ‘Cry Freedom’ did technically not fulfill the *discreteness* criterion) as well as the neutral state. This meant that leaders saw the same (very short) video twice and that if a leader would have chosen ‘Anger’ three times, which never happened in our experiment, his or her followers would have seen the first movie again after the third ‘Anger’ choice.
4.3.3 Hypotheses

Hypotheses are presented first, followed by a rationale. Our first hypothesis concerns the difference in behavior between the different treatments.

**Hypothesis 4.1.** Contributions in SE will be larger than contributions in RE, which in turn will be larger than contributions in Baseline.

This hypothesis is driven by the assumption that the goal of most leaders will be to increase the contributions of the followers. SE gives the leader the most opportunities to influence the followers. In the more emotional environment of RE the contributions of the leader may again have a larger (positive) effect than in Baseline.

**Hypothesis 4.2.** Bids in SE will be higher than bids in RE, which in turn will be higher than bids in Baseline.

Because the amount of control the leader has is greatest in SE, our hypothesis is that bids for the leadership position will be highest in SE. Furthermore we hypothesize that participants value the role of leader more in the more emotional RE treatment, when compared with Baseline.

**Hypothesis 4.3.** Leaders will contribute more than followers.

In subsection 2.1 we presented three different behavioral models that predict that leaders may contribute more: the affective ties model of van Dijk and van Winden (1997), the imitation model of Cartwright and Patel (2010), and a psychological games model of Dufwenberg and Kirchsteiger (2004). There is also empirical evidence suggesting that in our conflict game leaders will contribute more than followers. For example, as Arbak and Villeval (2013) and Reuben et al. (2015) show that leaders are more competitive than followers, this would make them to contribute more in our game in order to win the conflict. Furthermore, the findings by Potters et al. (2007) and Gächter et al. (2014), that leaders contribute more to public good games than their followers when given a first-mover advantage, indicate that players contribute more when given the leadership position. Another finding that supports this hypothesis comes
from Myers (2015), who shows that self-selected leaders care more about the welfare of others. This should lead to higher contributions by leaders as their contributions positively affect the payoff of their followers.

Our last hypothesis focuses on the effect of the different emotions.

**Hypothesis 4.4.** Positive emotions have a positive influence on contributions.

Telle and (2015) present an overview of the evidence for a positive influence of positive affect on prosocial behavior. In our game this would translate to positive emotions improving the relationship between the leader and followers, and making followers act more prosocial towards the leader. This should increase contributions. Negative emotions, on the other hand, would deteriorate the relationship between the leader and the followers, decreasing the contributions. A similar process is described by Humphrey (2002), who shows that positive emotions signaled by leaders correlate with better job satisfaction as well as positive emotions in followers.

### 4.4 Results

In this section we first look if there are differences in the average contribution between the treatments. We then turn our attention to the behavior in the auction, before we look at the potential differences between the contributions of leaders and followers. Next, we study the choices of emotions by leaders and its direct/isolated effect on the contribution behavior of the followers. Finally, we investigate what exactly drives the (change in) contribution by leaders and followers. Regarding the latter, three key determinants will be focused on: the recent history, the (change in) contribution of the leader and the choice of emotion by the leader.

Turning now to the contribution levels in the three treatments, table 1 shows that they are very similar and not significantly different.
Table 4.1: Average contribution per treatment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Contribution</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>8.26 (0.61)</td>
<td>20</td>
</tr>
<tr>
<td>SE</td>
<td>8.68 (0.52)</td>
<td>20</td>
</tr>
<tr>
<td>RE</td>
<td>8.38 (0.51)</td>
<td>20</td>
</tr>
</tbody>
</table>

Note: Robust standard errors clustered on the group level in parentheses. No significant differences between treatments at the 5% level, using a Mann-Whitney-Wilcoxon (MWW) test.

If we compare the actual contributions it is clear that players do not contribute according to the Nash predictions. Contributions range between 40% and 45% of the endowments, which is more than twice the predicted 20% (3 followers contributing 26% and the leader zero). Instead, these results seem to be more in line with Perfect Imitation or the Ties model.

It is interesting to note that the difference in emotional environment as well as a different role for the leader of a group does not lead to a significant change in the overall contribution level. This does not imply however that the environment does not matter. Positive and negative emotional stimuli might offset each other, and the extent to which leaders are followed and the way they behave could potentially differ between the different treatments. Another potential explanation for the limited effect of the emotional stimuli, is the fact that these stimuli are presented completely out of context, in stark contrast to the use of the stimuli outside of the lab. Overall, though, as there is no significant difference in average contribution between the treatments, using Mann-Whitney-Wilcoxon (MWW) tests, we do not find much support for our first hypothesis (H1). Also in the development of contributions over time there is little difference observable, as can be seen from figure 1.

---

6We thank an anonymous referee for this observation. We wanted to start our first approach to emotional leadership with a defendable but simple design, however, focusing on a small set of generally agreed upon basic emotions, using a standard technique of inducing emotions and only one choice of emotion for leaders.
4.4.1 Auction

If subjects have a different valuation for the leadership position in the different treatments, we would expect that the leadership position would be more valuable in the SE treatment as it gives leaders not only the opportunity to be the first mover but also allows them to select the emotion to be induced in followers. When we test, using MMW-tests, for differences in the bids between the treatments (see tables 2 and 3) we find no significant differences, however. If anything, subjects were offering more for the leadership position in Baseline, although leaders have seemingly more control in SE. This could be due to the fact that many subjects are uncomfortable with such ‘emotional control’ over others, or because subjects might dislike not being the leader less in SE. Based on these results we have to reject our second hypothesis (H2).
Table 4.2: Average bid per treatment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Bid</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>6.39 (0.65)</td>
<td>80</td>
</tr>
<tr>
<td>SE</td>
<td>4.96 (0.62)</td>
<td>80</td>
</tr>
<tr>
<td>RE</td>
<td>5.39 (0.65)</td>
<td>80</td>
</tr>
</tbody>
</table>

Note: Robust standard errors clustered on the group level in parentheses. No significant differences between treatments at the 5% level, using a Mann-Whitney-Wilcoxon (MWW) test.

Table 4.3: Average winning bid per treatment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Winning Bid</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>12.45 (1.08)</td>
<td>20</td>
</tr>
<tr>
<td>SE</td>
<td>11.20 (1.23)</td>
<td>20</td>
</tr>
<tr>
<td>RE</td>
<td>12.45 (1.34)</td>
<td>20</td>
</tr>
</tbody>
</table>

Note: Robust standard errors clustered on the group level in parentheses. No significant differences between treatments at the 5% level, using a Mann-Whitney-Wilcoxon (MWW) test.

4.4.2 Contributions of leaders and followers

To investigate whether leaders and followers contribute differently it is important to see if leaders either lead-by-example or try to exploit the fact that followers might want to make up for his or her lack of contributions. Table 4 shows that overall leaders seem to contribute (between 11% and 17%) more than followers. Moreover, for all treatments the differences between the contributions of leaders and followers are significant at the 5% level when we take the average total group contribution (excluding leaders) as one observation using a MWW-test. There is, however, no apparent difference between the treatments, as none of the differences is significantly different from zero at the 5% level. The fact that leaders contribute more than followers might be a sign that leaders indeed try to lead by example. In any case, it confirms our third hypothesis (H3). Furthermore, it provides evidence that the results found when looking at the

7The difference between leaders and followers is also significant in all treatments when using a parametric t-test (if we pool the data for all treatments, p<0.01).
group level are not due to perfect imitation, but could be more in line with the Ties model, models with psychological preferences or the imitation model of Cartwright and Patel, as discussed in section 2.1.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Leaders</th>
<th>Followers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>9.15 (0.80)</td>
<td>8.00 (0.45)</td>
</tr>
<tr>
<td>SE</td>
<td>9.77 (0.61)</td>
<td>8.32 (0.39)</td>
</tr>
<tr>
<td>RE</td>
<td>9 (0.59)</td>
<td>8.18 (0.40)</td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. Differences between roles are significant at the 5% level, using a Mann-Whitney-Wilcoxon (MWW) test.

### 4.4.3 Emotions

Now that we have established that leaders contribute more than their followers we turn our attention to the choice of emotion in the SE treatment. Figure 2 shows a histogram with the choice frequencies of all different emotions. We find that happiness is chosen most often. Some leaders indicated in the questionnaire that they selected happiness to keep their followers in a good mood, which would lead to higher contributions, while others chose happiness to express their own happy emotional state. Given that we have 20 leaders of which only 6 made a remark about happiness it is hard to draw strong conclusions from this, though.
To see if the choice of emotion has an effect on followers and can, thus, be used by leaders as a form of strategic emotional leadership, we first investigate the effect of emotion on the change in contribution of followers. We use first differences to correct for the fact that some individuals or groups start out with higher contribution levels than others. Let: \( \Delta \text{Contribution} = \text{Contribution}_{it} - \text{Contribution}_{it-1} \), where \( \text{Contribution}_{it} \) represents the contribution of participant \( i \) at time \( t \) and \( \text{Contribution}_{it-1} \) the contribution of participant \( i \) at time \( t - 1 \). In this estimation are used as dummy variables, where the neutral state is omitted. For the results, see table 5.
Table 4.5: Effect of emotions on the change in contribution of followers

<table>
<thead>
<tr>
<th></th>
<th>SE</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Δ Contribution</td>
<td>Δ Contribution</td>
</tr>
<tr>
<td>Anger</td>
<td>-0.353</td>
<td>-0.0263</td>
</tr>
<tr>
<td>Disgust</td>
<td>0.647</td>
<td>-1.328</td>
</tr>
<tr>
<td>Fear</td>
<td>0.478</td>
<td>-1.016</td>
</tr>
<tr>
<td>Happiness</td>
<td>0.430</td>
<td>-0.237</td>
</tr>
<tr>
<td>Sadness</td>
<td>-0.524</td>
<td>-0.714</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0909</td>
<td>0.447</td>
</tr>
</tbody>
</table>

Observations 660 660
R-squared 0.009 0.007

Note: OLS estimation, robust standard errors in parentheses.
None of the coefficients is significant at the 5% level.
In SE the coefficients are jointly significant at the 5% level.

When studied in isolation it seems as if emotions do not affect the change in contribution much (jointly they do have a significant effect in SE). Because some of the negative emotions were not picked that often and because the clips inducing a specific negative emotion typically also evoked other negative emotions\(^8\) we continue to work with the *valence* of an emotion from now on, which is constructed as follows: It takes the value of 1 if a positive emotion (happiness) has been selected, a value of 0 if it concerns the neutral state, and -1 in case of a negative emotion. Valence turns out to have a significant positive effect on the change in contribution in SE (0.65, \(p \approx 0.04\)), but not in RE(0.4, \(p > 0.1\), using OLS regression. This supports our fourth hypothesis (H4). We have to be careful, however, as this relationship might not be causal.

A confounding factor that comes to mind (only in the SE treatment) is that the result of the previous game might influence both the contributions in the next period as well as the choice of emotion. If we look at the SE treatment in periods 4, 7 and 10 we find a correlation of

\(^8\)in the case of *Cry Freedom* even so much so that it violated the Discreetness criterion, see Appendix C.5.
only -0.19 between valence and loss (a variable that indicates if a loss occurred in the previous period). Loss, however, does not seem to be a significant predictor of valence, as loss is not significant if we use a ordered logit regression ($p>0.1$).

### 4.4.4 Behavior of followers

Leaders can try to influence their followers in two different ways. They can try to change the behavior of the followers by the choice of emotion and by making a contribution, that is observable for the followers before they make their decision. These publicly observable contributions can be seen as a form of leading by example, as the leader might be trying to set a norm or show his or her affection towards the followers. As discussed in subsection 2.1, leading-by-example is often found to be important in settings involving public goods. In our game a contribution generates a public good within the group, but across groups it provides a public bad. So it is not immediately clear what leading-by-example should imply. Before we saw that leaders contribute more than followers, but do their contributions influence the followers? When regressing the change in contribution of the leader on the change in contributions of followers we find an effect size of around 0.2 for all treatments, and this effect is always significant at the 5% level, and in Baseline even at the 1% level. Once again we have to be careful with drawing conclusions as the shared history (winning or losing) might influence both decisions.

To come to firmer results we need to add variables to control for the shared history. Based on Lacomba et al. (2014), who analyzed a 2-person version of this game, we assume that it is not necessarily only the fact that one won or lost the previous period that might play a role, but also the difference in contribution levels between the two groups (or two individuals in their case).

Furthermore, we will control for valence, and the change in contribution of the leader. Using ordinary least squares (OLS) regression, we first estimate the effect of (recent) history on the change in contribution of followers (model 1) and subsequently add other variables in new models. Model 2 adds the change in contribution of the leader, models 3 adds valence, while
both valence as well as the change in contribution of the leader are added in model 4. Tables 6, 7, and 8 show the estimates of these four models for the different treatments.

Table 4.6: The change of contribution of followers in Baseline

<table>
<thead>
<tr>
<th>Models</th>
<th>Baseline</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ΔContribution</td>
<td>ΔContribution</td>
<td>ΔContribution</td>
</tr>
<tr>
<td>Loss</td>
<td>1.409**</td>
<td>1.250**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.406)</td>
<td>(0.419)</td>
<td></td>
</tr>
<tr>
<td>( \frac{\sum C_i - \sum C_j}{4} )</td>
<td>-0.0284</td>
<td>0.000551</td>
<td></td>
</tr>
<tr>
<td>( \sum C_i - \sum C_j &gt; 0 ) &amp; (0.153)</td>
<td>(0.152)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \frac{\sum C_i - \sum C_j}{4} )</td>
<td>0.309</td>
<td>0.405*</td>
<td></td>
</tr>
<tr>
<td>( \sum C_j - \sum C_i &gt; 0 ) &amp; (0.218)</td>
<td>(0.198)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔLeader</td>
<td>0.207**</td>
<td></td>
<td>(0.0717)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.399</td>
<td>-0.692</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.392)</td>
<td>(0.352)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>660</td>
<td>660</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.042</td>
<td>0.083</td>
<td></td>
</tr>
</tbody>
</table>

Note: OLS estimation, robust standard errors clustered on the individual level in parentheses.

** p<0.01, * p<0.05
Table 4.7: The change of contribution of followers in SE

<table>
<thead>
<tr>
<th>SE</th>
<th>Models</th>
<th>(1) ΔContribution</th>
<th>(2) ΔContribution</th>
<th>(3) ΔContribution</th>
<th>(4) ΔContribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>0.771***</td>
<td>0.541</td>
<td>0.780***</td>
<td>0.550</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.331)</td>
<td>(0.33)</td>
<td>(0.33)</td>
<td>(0.333)</td>
<td></td>
</tr>
<tr>
<td>(\sum C_i - \sum C_j)</td>
<td>0.199*</td>
<td>0.189*</td>
<td>0.168</td>
<td>0.157</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.106)</td>
<td>(0.112)</td>
<td>(0.109)</td>
<td></td>
</tr>
<tr>
<td>Valence</td>
<td>0.265**</td>
<td>0.265**</td>
<td>0.272**</td>
<td>0.272**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.122)</td>
<td>(0.122)</td>
<td>(0.122)</td>
<td>(0.117)</td>
<td></td>
</tr>
<tr>
<td>ΔLeader</td>
<td>0.211***</td>
<td>0.211***</td>
<td>0.212***</td>
<td>0.212***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0755)</td>
<td>(0.0755)</td>
<td>(0.0755)</td>
<td>(0.0755)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.844***</td>
<td>-0.698**</td>
<td>-0.772***</td>
<td>-0.624**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.269)</td>
<td>(0.279)</td>
<td>(0.269)</td>
<td>(0.280)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>660</td>
<td>660</td>
<td>660</td>
<td>660</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.027</td>
<td>0.071</td>
<td>0.031</td>
<td>0.079</td>
<td></td>
</tr>
</tbody>
</table>

Note: OLS estimation, robust standard errors clustered on the individual level in parentheses. 
*** p<0.01, ** p<0.05, * p<0.10
Table 4.8: The change of contribution of followers in RE

<table>
<thead>
<tr>
<th>Models</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔContribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss</td>
<td>1.001**</td>
<td>0.804*</td>
<td>0.993**</td>
<td>0.801*</td>
</tr>
<tr>
<td>(0.362)</td>
<td>(0.370)</td>
<td>(0.369)</td>
<td>(0.375)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>∑Ci - ∑Cj</td>
<td>-0.0877</td>
<td>-0.0984</td>
<td>-0.0862</td>
<td>-0.0975</td>
</tr>
<tr>
<td>(0.198)</td>
<td>(0.193)</td>
<td>(0.197)</td>
<td>(0.193)</td>
<td></td>
</tr>
<tr>
<td>Valence</td>
<td>0.103</td>
<td>0.128</td>
<td>0.183</td>
<td>0.122</td>
</tr>
<tr>
<td>(0.191)</td>
<td>(0.191)</td>
<td>(0.191)</td>
<td>(0.191)</td>
<td></td>
</tr>
<tr>
<td>ΔLeader</td>
<td>0.172**</td>
<td>0.172**</td>
<td>0.172**</td>
<td>0.172**</td>
</tr>
<tr>
<td>(0.0603)</td>
<td>(0.0603)</td>
<td>(0.0608)</td>
<td>(0.0608)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.491</td>
<td>-0.344</td>
<td>-0.464</td>
<td>-0.329</td>
</tr>
<tr>
<td>(0.416)</td>
<td>(0.415)</td>
<td>(0.424)</td>
<td>(0.422)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>660</td>
<td>660</td>
<td>660</td>
<td>660</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.025</td>
<td>0.076</td>
<td>0.026</td>
<td>0.076</td>
</tr>
</tbody>
</table>

Note: OLS estimation, robust standard errors clustered on the individual level in parentheses. ** p<0.01, * p<0.05

First we notice that the (recent) history does play an important role in the behavior of followers. This can be seen from the fact that either losing in the previous round or the difference in contribution between groups (if this difference is negative) is highly significant in all model specifications of all treatments. As a result, groups that lose, in a certain period, typically increase their contribution the next period. Only in the SE treatment, though, it seems more important whether or not the group contributed more than their opponents, as was also found by Lacomba et al. (2014) in a two-player setting. These two events, contributing less and losing, are obviously positively correlated.

Furthermore it becomes clear, from tables 6, 7, and 8, that the contribution of the leader is of significant importance in all three treatments. The effect of ΔLeader on ΔContribution is quite constant over the treatments, as the effect size is between 0.17 in RE and 0.21 in SE. This means that for every additional token contributed by the leader every one of his or her followers
will contribute around 0.2 additional tokens. So, if the leader contributes 5 tokens more, his or her followers will contribute one additional token. From the leader’s perspective contributing (more) is thus more attractive, as additional contributions are basically matched with 60%. These results – that a contribution of the leader has a positive effect on the contributions of followers, and that leaders contribute more than followers (see section 4.2) – can be explained by the tie model, as well as the imitation model and a model of psychological preferences, which are all discussed in section 2.1. We also note that $\Delta$Leader seems to mediate the effect of loss, as the effect of loss becomes smaller in all model specifications when $\Delta$Leader is introduced and both the effect of loss on $\Delta$Leader as well as the effect of $\Delta$Leader on $\Delta$Contribution is significant. This represents the fact that followers and leaders are affected in a similar way by a loss, as we will see in the next subsection.

When controlling for (recent) history, one can observe from table 7 and 8 that a leader has the ability to both lead by example as well as lead by emotional influence. Regarding the latter s/he needs control over these emotions (of others), though. The showing of emotional video clips as such does not help a leader, as we can see from the absence of a significant effect of valence and the smaller effect of $\Delta$Leader in the RE treatment. In an emotional environment that is caused by outside events these emotions do not seem to effect behavior of followers nor does this effect enhances the ability of a leader to lead by example. This means that our fourth hypothesis (H4), that positive emotions have a positive effect on contributions, is only true under the condition where these emotions are evoked by the leader of the group. Whether this is due to a timing issue – with some leaders knowing when it is useful to induce these emotions – or due to followers caring about the source of the emotions, is still unclear. Future research could help us find out the relative importance of these two channels.

---

9For more on the mechanism of mediation, see Baron and Kenny (1986).
10It is interesting to note that the correlation between $\Delta$Leader and valence is -.01 in this treatment.
4.4.5 Behavior of Leaders

The finding that leaders have a big influence on followers with their contribution raises the question what drives the changes in the contribution of leaders. We use the same recent history variables as before in model 1 for the change in contributions. Furthermore, we pool the results for all leaders as they all see the same clips.

Table 4.9: Contribution of leaders

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>SE</th>
<th>RE</th>
<th>Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ Leader</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss</td>
<td>1.734**</td>
<td>1.084**</td>
<td>1.607**</td>
<td>1.482***</td>
</tr>
<tr>
<td></td>
<td>(0.808)</td>
<td>(0.516)</td>
<td>(0.651)</td>
<td>(0.381)</td>
</tr>
<tr>
<td>(\frac{</td>
<td>\sum C_i - \sum C_j</td>
<td>}{4})</td>
<td>-0.0365</td>
<td>0.0460</td>
</tr>
<tr>
<td></td>
<td>(0.299)</td>
<td>(0.255)</td>
<td>(0.299)</td>
<td>(0.161)</td>
</tr>
<tr>
<td>(\frac{</td>
<td>\sum C_i - \sum C_j</td>
<td>}{4})</td>
<td>0.376</td>
<td>0.325</td>
</tr>
<tr>
<td></td>
<td>(0.297)</td>
<td>(0.255)</td>
<td>(0.300)</td>
<td>(0.160)</td>
</tr>
<tr>
<td>(\frac{</td>
<td>\sum C_j - \sum C_i</td>
<td>}{4})</td>
<td>-1.111</td>
<td>-0.690</td>
</tr>
<tr>
<td></td>
<td>(0.720)</td>
<td>(0.535)</td>
<td>(0.619)</td>
<td>(0.358)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.111</td>
<td>-0.690</td>
<td>-1.095*</td>
<td>-0.955***</td>
</tr>
<tr>
<td></td>
<td>(0.720)</td>
<td>(0.535)</td>
<td>(0.619)</td>
<td>(0.358)</td>
</tr>
</tbody>
</table>

Observations: 220 220 220 660
R-squared: 0.036 0.031 0.037 0.034

Note: OLS estimation, robust standard errors clustered on the individual level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 9 shows that leaders are clearly driven by the results in the previous period. Their strategy seems to be rather simple: if the group won (and contributed more than the other group) the leader typically contributes less to increase potential profit; after a loss, though, the leader contributes more in order to increase the chance of winning. Moreover, their contributions are also increased when their group contributed less than the other group.

4.5 Concluding discussion

We introduce leadership in a conflict environment by using a sequential group conflict game. In this game there are two ways to lead for leaders: leading-by-example and leading by the
strategic use of emotions. The former is possible by giving leaders the opportunity to contribute before their followers. Evoking basic emotions in followers relates to the latter form of leadership. To that purpose, specially selected video clips are available to leaders. Every three periods leaders choose an emotion (not a clip) that is then induced in followers by the use of one of these video clips. Although no significant treatment effect on the level of contribution has been found, potentially due to the very minimal setting and the use of emotional movie clips without any context, we do find evidence that leading-by-example is an important channel for leaders in our environment(s). Emotional leadership is found to have a significant effect as well, but seems less potent in this environment, as leaders do not appear to be effective in using it strategically. Note, however, that our design only allows for minimal emotional interaction between leaders and followers, as leaders are only able to evoke a basic emotion and have no other means of verbal communication. Therefore this research is by no means evidence that (strategic) emotional leadership is more generally of less importance.

In a theoretical analysis of the game we find (very) different behavioral predictions. In sharp contrast to the Nash equilibrium predictions, in which leaders should not contribute at all and followers contribute about a quarter of their endowment, we find that leaders contribute more than followers and both leaders and followers contribute much more than predicted. The Ties model of van Dijk and van Winden (1997), the Imitation model of Cartwright and Patel (2010) and the psychological preferences model of Dufwenberg et al. (2011) can in principle explain our findings. All these models share one important characteristic, namely that they can explain the positive relationship between the contribution of the leader and that of the followers. In the Ties model this positive relation is created by affective bonds created by the contribution of the leader, in the psychological preferences model the contribution of the leader creates a norm that (through guilt aversion) followers prefer to follow, while in the imitation model this relationship is driven by the existence of imitators that simply follow the leader. We leave for future research the question what the precise motives are for this leading by example. Are leaders motivated by higher payoffs for themselves (as the strategists in the imitation model), or initially (at least)
partly driven by an intrinsic prosocial motivation related to social value orientation or norms?

Appendix 4.A  Theory

In this section we elaborate on the intergroup conflict game discussed in section 2. More specifically, we start by giving an overview of the equilibrium predictions for different behavioral models and in the subsequent subsections derive these predictions.

We start our analysis with the Ties model. For the rest of this analysis we normalize the endowment, putting $Y$ equal to one, which is equivalent to interpreting $C$ as the fraction of $Y$ that is contributed. We start by modeling the payoff function of the follower. Assuming that the followers initially do not care about the other group members, according to the social ties model, the weight attached to the payoff of the others will then be determined by the impulse. As behavioral reference point we take the contribution predicted by the standard Nash equilibrium $(0)$, based on experimental evidence (Loerakker et al., 2016). Then, the weight attached to the utility of the leader – perceived as made up by the wellbeing of the group – is determined by the contribution of the leader $C_{Li}$ (which determines the impulse) multiplied by an impulse parameter $\delta$. As a result, the expected payoff of a follower (eq.4.1) is now assumed to include $\delta C_{Li}$ times the total payoff of the other group members:

\[
E\pi_F = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} \left(2 - \frac{1}{4} \sum_{j=1}^{4} C_j - C_F + \delta C_{Li}(6 - \frac{3}{4} \sum_{j=1}^{4} C_j - 2C_F - C_{Li})\right)
\]

(4.A.1)

We further assume that a leader anticipates this when making a contribution and may care for the outcomes of the followers from the onset. Many experimental studies have shown, using social value orientation measures, that there are prosocial individuals who care for the outcomes of others (McClintock and Liebrand, 1988, van Lange, 1999). This might be especially relevant for self-selected leaders, because prosocials are also found to participate more in voter participation games (Myers, 2015). This is captured in eq. 4.A.2, representing the expected payoff
of the leader ($E\pi_{Li}$), by the parameter $\alpha_i$ that represents the weight the leader attaches to the payoffs of the followers.

$$E\pi_{Li} = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} \left( 2 - \frac{1}{4} \sum_{j=1}^{4} C_j - C_L + 3\alpha_i(2 - \frac{1}{4} \sum_{j=1}^{4} C_j - \frac{1}{3} \sum_{i=1}^{3} C_F) \right)$$

(4.A.2)

Table A.1 shows the predicted contributions of leaders and followers for different values of $\alpha$ and $\delta$ (for a more in-depth analysis of this model, see appendix A.4):

Table 4.A.1: Predicted contributions for the social ties model, using different parameter values

<table>
<thead>
<tr>
<th>$\delta = 0$</th>
<th>$\delta = \frac{1}{2}$</th>
<th>$\delta = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha = 0$</td>
<td>0, 0.26</td>
<td>0.09, 0.26</td>
</tr>
<tr>
<td>$\alpha = 0.1$</td>
<td>0, 0.26</td>
<td>0.20, 0.26</td>
</tr>
<tr>
<td>$\alpha = 0.2$</td>
<td>0, 0.26</td>
<td>0.28, 0.26</td>
</tr>
<tr>
<td>$\alpha = 0.3$</td>
<td>0, 0.26</td>
<td>0.39, 0.25</td>
</tr>
</tbody>
</table>

Note: The first number is the predicted contribution for the leader, while the second number is the prediction for a follower.

From the figures in this table we learn that total contributions rise when $\alpha$ and $\delta$ go up. Furthermore, it is interesting to see that if both $\alpha$ and $\delta$ are substantially positive we have a situation were a leader contributes more than his followers and that this does not happen when only one of them is positive.

Furthermore, note that if both leaders anticipate to be perfectly followed (Perfect Imitation) the game becomes identical to a game between 2 players. The equilibrium for that game was already derived by Lacomba et al. (2014). They show that in a Nash equilibrium both players (the leaders in our game) will contribute $C = 1/2$, earning in expectation $1/2$. If only one of the leaders is followed while the other leader does not have any influence on the contributions of his or her followers (Perfect Imitation vs. No Influence) it can be shown that in equilibrium the group that follows their leader contributes $C = 4/9$, while the group that does not contributes $C = 2/9$ (see appendix A.3). As a consequence, the ‘following’ group also has expected
earnings that are twice as high in equilibrium, \( \frac{5}{9} \) versus \( \frac{4}{9} \).

Table A.2, below, summarizes the predictions of the beforementioned behavioral models regarding our game. The table starts by displaying the Nash equilibria of our sequential conflict game (Nash) and its simultaneous version (Nash Simultaneous). Then, the contributions predicted by a model wherein both groups perfectly follow their leader (Perfect Imitation) are shown, followed by the prediction for a group that perfectly follows their leader while the other group completely neglects its leader (Perfect Imitation vs. No Influence). Predictions for a group that neglects its leader while playing against a group that does not do so (No Influence vs Perfect Imitation) are shown next. Lastly, the contributions predicted by the social ties model (Ties) for two different parameter combinations are presented. Finally, it is noted that models that include psychological costs or the imitation model developed by Cartwright and Patel could in principle give roughly the same predictions as the Ties model in this table; the latter, for instance, depending on the distribution of strategists, imitators and independents in the population.

<table>
<thead>
<tr>
<th>Models</th>
<th>Leaders</th>
<th>Followers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nash</td>
<td>0</td>
<td>0.26</td>
</tr>
<tr>
<td>Nash Simultaneous</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Perfect Imitation</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Perfect Imitation vs. No Influence</td>
<td>0.44</td>
<td>0.44</td>
</tr>
<tr>
<td>No Influence vs. Perfect Imitation</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Ties (( \alpha = 0.1, \delta = \frac{1}{2} ))</td>
<td>0.20</td>
<td>0.26</td>
</tr>
<tr>
<td>Ties (( \alpha = 0.3, \delta = 1 ))</td>
<td>0.43</td>
<td>0.34</td>
</tr>
</tbody>
</table>

4.A.1 Equilibrium analysis

We start our formal analysis by deriving the equilibria for risk neutral as well as for different types of risk averse players. The game has 8 players that are divided into 2 groups. Each player is endowed with \( Y \) and has to decide how much to contribute \( (C_i) \) to their group’s conflict. The
group that wins the conflict gets all that is not invested in the conflict, while members of the losing group get nothing. More specifically, each member of the winning group keeps its own remaining resources \((Y - C_i)\) and get, in addition, one-fourth of the remaining resources of the other group \((\sum_{j=1}^{4} Y - C_j)\). The winner of the conflict is determined by a lottery, where the odds that the group of \(i\) wins are determined by \(P(Win_i) = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j}\). The expected payoff of \(i\), thus, is:

\[
E\pi_i = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} \left( Y - 1/4 \sum_{j=1}^{4} C_j + Y - C_i \right) \quad (4.A.3)
\]

For expositional reasons we start here with the simultaneous game (Nash Simultaneous) followed by the sequential game of our intergroup conflict (Nash).

**Simultaneous game**

We take the first derivative of the expected payoff described by eq. (4.A.3) to determine a Nash equilibrium:

\[
\frac{dE\pi_i}{dC_i} = -\frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} + \left( 2Y - 1/4 \sum_{j=1}^{4} C_j - C_i \right) \frac{\sum_{j=1}^{4} C_j}{(\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j)^2} \quad (4.A.4)
\]

A Nash equilibrium requires \(dE\pi_i/dC_i\), for all \(i\), and thus:

\[
\frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} = \left( 2Y - 1/4 \sum_{j=1}^{4} C_j - C_i \right) \frac{\sum_{j=1}^{4} C_j}{(\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j)^2} \quad (4.A.5)
\]

The term left of the equality sign is the same for all members of the same group. The term right of the equality sign, however, is only equal for all group members if the contributions of all group members are identical (Symmetry). We simplify with \(X_1 \equiv \sum_{i=1}^{4} C_i\) and \(X_2 \equiv \sum_{j=1}^{4} C_j\), rendering:
\[
\frac{X_1}{X_1 + X_2} = \frac{2YX_2 - 1/4X_2^2 - 1/4X_1X_2}{(X_1 + X_2)^2} \quad (4.A.6)
\]

A similar condition holds for the other group:

\[
\frac{X_2}{X_1 + X_2} = \frac{2YX_1 - 1/4X_1^2 - 1/4X_1X_2}{(X_1 + X_2)^2} \quad (4.A.7)
\]

Combining and elaborating gives:

\[
\frac{X_1}{X_2} = \frac{2YX_2 - 1/4X_2^3 - 1/4X_1X_2}{2YX_1 - 1/4X_1^3 - 1/4X_1X_2} => \]

\[
2YX_1^2 - 1/4X_1^3 - 1/4X_1^2X_2 = 2YX_2^2 - 1/4X_2^3 - 1/4X_2^2X_1 => \quad (4.A.8)
\]

\[
X_1^2(2Y - 1/4X_1 - 1/4X_2) = X_2^2(2Y - 1/4X_1 - 1/4X_2) => \]

\[
X_1^2 = X_2^2 => X_1 = X_2
\]

Having established that all contributions must be equal, using eq. (4.A.5), this leads to the following symmetric equilibrium contribution for Nash Simultaneous:

\[
\frac{1}{2} = ((2Y - 2C_i)^4C_i^4) = \frac{4C_i}{64C_i^2} => \]

\[
\frac{1}{2} = ((Y - C_i)^C_i^8C_i^2) => \quad (4.A.9)
\]

\[
8C_i^2 = 2YC_i - 2C_i^2 => \]

\[
10C_i^2 = 2YC_i => C_i = 1/5Y
\]

**Sequential game (Nash)**

To analyze the sequential game of our intergroup conflict we start from eq. (4.A.5). First using the same arguments as for the simultaneous game, we conclude that all followers will contribute the same amount. Next, note that \(\sum_{i=1}^{4} C_i = \sum_{j=1}^{4} C_j\) should hold, since the reward is the same for both groups and if \(\sum_{i=1}^{4} C_i \neq \sum_{j=1}^{4} C_j\) the chance of winning is different for
the two groups. If this would be the case it cannot be that \( \frac{dE\pi_i}{dC_i} = 0, \frac{dE\pi_i}{dC_F} = 0, \frac{dE\pi_i}{dC_{Fj}} = 0 \) and \( \frac{dE\pi_i}{dC_{Li,j}} = 0 \). We continue by finding the best response of the followers from eq. (4.A.5) and filling this in directly in eq. (4.A.3) for the leaders. With \( C_F \) and \( C_L \), respectively, denoting the contribution of a follower and a leader eq. (4.A.5) can now be rewritten in the following way:

\[
\sum_{i=1}^{4} C_i = \left( \frac{2Y - 1}{4} \sum_{j=1}^{4} C_j - C_i \right) \left( \frac{2Y - 1}{4} \sum_{j=1}^{4} C_j \sum_{j=1}^{4} C_j \right) = > \\
\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j = 2Y - 1/4(\sum_{j=1}^{4} C_j) - C_i = > (4.A.10)
\]

\[
2C_L + 6C_F = 2Y - 7/4C_F - 1/4C_L = > \\
C_F = 8/31Y - 9/31C_L \text{ if } C_L \leq 8/9 \text{ else } C_F = 0
\]

Inserting this condition in eq. (4.A.3), the expected payoff for a leader can now be written as:

\[
E\pi_i = \frac{4/31C_L + 24/31Y}{4/31C_L + 24/31Y + \sum_{j=1}^{4} C_j}(2Y - 1/4 \sum_{j=1}^{4} C_j - C_L)
\]

Taking the first derivative with respect to \( C_L \) and setting it equal to zero, gives:

\[
\frac{4/31C_L + 24/31Y}{4/31C_L + 24/31Y + \sum_{j=1}^{4} C_j} = \frac{4}{31} \frac{\sum_{j=1}^{4} C_j}{(4/31C_L + 24/31Y + \sum_{j=1}^{4} C_j)^2}(2Y - 1/4 \sum_{j=1}^{4} C_j - C_L)
\]

(4.A.12)

Using \( \sum_{i=1}^{4} C_i = \sum_{j=1}^{4} C_j \) we obtain:

\[
1 = \frac{4}{31} \frac{2Y - (1/31C_L + 6/31Y) - C_L}{2(4/31C_L + 24/31Y)} = > \\
2(4/31C_L + 24/31Y) = \frac{4}{31}(2Y - (1/31C_L + 6/31Y) - C_L) = > (4.A.13)
\]

\[
C_L = \frac{-316}{94}Y
\]

As this is outside of the range of feasible contributions, we continue by investigating whether \( C_L = 0 \) is a best response. Given that leader \( j \) contributes zero:
\[ E\pi_i = \frac{4/31 C_L + 24/31 Y}{4/31 C_L + 48/31 Y} \cdot (56/31 Y - C_L) \Rightarrow \\
E\pi_i = \frac{-4/31 C_L^2 - 520/961 C_L Y + 1344/961 Y^2}{4/31 C_L + 48/31 Y} \]

(4.A.14)

As the leader’s expected payoff is clearly decreasing in \( C_L \) in the positive domain, we conclude that there is a pure Nash equilibrium where both leaders contribute 0 and all followers contribute \( 8/31 Y \).

4.A.2 Risk aversion

This section provides an equilibrium analysis for risk averse, instead of risk neutral players. For convenience we will focus here on symmetric equilibria in the simultaneous game (see A.1.), with all players having known and identical utility functions. Two often used utility functions will be evaluated, one that assumes constant absolute risk aversion (CARA) and one that assumes constant relative risk aversion (CRRA).

CARA

We begin by looking at a situation where all players have the following CARA utility function:

\[ u_i = 1 - e^{-\alpha \Pi_i} \]

(4.A.15)

with \( \Pi_i \) denoting the monetary payoff. In that case, the expected payoffs equals:

\[ E u_i = \frac{X_1}{X_1 + X_2} \cdot \left(1 - e^{-\alpha(2Y^{1/4} \sum_{j=1}^{4} C_j - C_i)}\right) \]

(4.A.16)

Using the first-order condition: \( \frac{dE u_i}{dC_i} = 0 \), we get:
\[
\frac{X_2}{(X_1 + X_2)^2} (1 - e^{-\alpha(2Y - 1/4 \sum_{j=1}^4 C_j - C_i)}) = \frac{X_1}{X_1 + X_2} \alpha e^{-\alpha(2Y - 1/4 \sum_{j=1}^4 C_j - C_i)}
\]

\[
\frac{1}{2} (1 - e^{-\alpha((2Y - 2C_i))}) = 4C_i \alpha e^{-\alpha((2Y - 2C_i))}
\]

(4.A.17)

\[
\ln(1 + 8\alpha C_i) = 2\alpha(Y - C_i)
\]

As there is no nice expression that relates \(C_i\) to \(Y\), we first look at what happens if \(\alpha\) approaches zero. As this leads to zeros on both sides we need to use L'Hôpital’s rule to get:

\[
\frac{d}{d\alpha} \ln(1 + 8\alpha C_i) = \frac{d}{d\alpha} 2\alpha(Y - C_i) = > \\
\frac{8C_i}{1 + 8\alpha C_i} = 2(Y - C_i)
\]

(4.A.18)

\[
8C_i = 2(Y - C_i) \\
C_i = \frac{1}{5}Y
\]

This confirms our analysis above that without risk aversion the equilibrium is \(C_i = \frac{1}{5}Y\).

Figure A.1 shows how the symmetric equilibrium contribution changes as the level of risk aversion (\(\alpha\)) changes, using \(Y = 20\) as in the experiment.
Figure 4.A.1: Equilibrium contributions for different values of constant absolute risk aversion

CRRA

For the analysis of CRRA utility functions we take the following widely used specification:

\[ u_i = \frac{\pi_i^{1-\beta}}{1-\beta} \]  

(4.A.19)

This leads to the following expected utility function:

\[ Eu_i = \frac{X_1}{X_1 + X_2} \frac{(2Y - 1/4 \sum_{j=1}^{4} C_j - C_i)^{1-\beta}}{1-\beta} \]  

(4.A.20)

We continue analogously to the CARA case above:
\[
\frac{X_2}{(X_1 + X_2)^2} \frac{(2Y - 1/4 \sum_{j=1}^{4} C_j - C_i)^{1-\beta}}{1 - \beta} = \frac{X_1}{X_1 + X_2} (2Y - 2C_i)^{-\beta} \\
\frac{1}{2} \frac{(2Y - 2C_i)^{1-\beta}}{1 - \beta} = 4C_i(2Y - 2C_i)^{-\beta} \quad (4.A.21)
\]

\[
8C_i = \frac{1}{1 - \beta} 2(Y - C_i)
\]

\[
C_i = \frac{2}{(1 - \beta)(8 + \frac{2}{1-\beta})} Y
\]

Here it is immediately clear that if \( \beta \) approaches zero the symmetric best response moves to \( C_i^{br} = \frac{1}{5} Y \). Figure A.2 depicts the equilibrium contribution for different levels of risk aversion, using again \( Y = 20 \).
4.A.3 On (a)symmetric perfect following of leaders

If the leader is not followed (we use Nash Simultaneous to model this situation) this does not affect the Nash equilibrium, but what if they do follow him or her? If both leaders are perfectly followed (Perfect Imitation) they basically play as if they are in a two-player version of the game. Lacomba et al. (2014) showed that in this game both players contribute half of their endowment in the Nash equilibrium.

Let us now consider a situation where only one of the groups perfectly imitates their leader, while the other group neglects its leader (No Influence vs. Perfect Imitation). For the group
neglecting the leader we start with eq. (4.1):

\[ E_{\pi_i} = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} (2Y - \frac{1}{4} \sum_{j=1}^{4} C_j - C_i) \]  

(4.A.22)

And again as in eq. 4.A.5 a Nash equilibrium requires:

\[ \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} \left( 2Y - \frac{1}{4} \sum_{j=1}^{4} C_j - C_i \right) \left( \frac{\sum_{j=1}^{4} C_j}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} \right) \]  

(4.A.23)

Simplifying with \( X_1 \equiv \sum_{i=1}^{4} C_i \) and \( X_2 \equiv \sum_{j=1}^{4} C_j \), one obtains:

\[ \frac{X_1}{X_1 + X_2} = \frac{2YX_2 - \frac{1}{4}X_2^2 - \frac{1}{4}X_1X_2}{(X_1 + X_2)^2} \]  

(4.A.24)

For the group perfectly following its leader, a different condition holds. First of all, this leader’s expected payoff will be (with \( C_j \equiv 4C_{Lj} \)):

\[ E_{\pi_{Lj}} = \frac{1}{4} \frac{C_j}{\sum_{i=1}^{4} C_i + C_j} (8Y - \sum_{j=1}^{4} C_i - C_j) \]  

(4.A.25)

Using the first-order condition for a maximum, \( dE_{\pi_{Lj}} / dEC_j = 0 \), the following condition for an equilibrium is arrived at:

\[ \frac{C_j}{\sum_{i=1}^{4} C_i + C_j} = (8Y - \sum_{j=1}^{4} C_i - C_j) \left( \frac{\sum_{j=1}^{4} C_i}{\sum_{i=1}^{4} C_i + C_j} \right)^2 \]  

(4.A.26)

We now follow the same procedure as described above (after noting that the behavior of the members of the first group must be the same in equilibrium and setting \( X_1 \equiv \sum_{i=1}^{4} C_i \) and \( X_2 \equiv C_j \)). For an individual in the group where the leader has no influence condition eq. (4.A.24) holds:

\[ \frac{X_1}{X_1 + X_2} = \frac{2YX_2 - \frac{1}{4}X_2^2 - \frac{1}{4}X_1X_2}{(X_1 + X_2)^2} \]  

(4.A.27)
, while for the leader who is perfectly imitated eq. (4.A.26) can be rewritten as:

\[ \frac{X_2}{X_1 + X_2} = \frac{8YX_1 - X_1^2 - X_1X_2}{(X_1 + X_2)^2} \]  

(4.A.28)

Combining eq. (4.A.27) and eq. (4.A.28) gives:

\[ \frac{X_1}{X_2} = \frac{2YX_2 - \frac{1}{4}X_2^2 - \frac{1}{4}X_1X_2}{8YX_1 - X_1^2 - X_1X_2} \Rightarrow \]

\[ 8XY_1^2 - X_1^3 - X_2^2X_2 = 2YX_2^2 - \frac{1}{4}X_2^3 - \frac{1}{4}X_2X_2^1 \Rightarrow \]

\[ X_1^2(8Y - X_1 - X_2) = X_2^2(2Y - \frac{1}{4}X_1 - \frac{1}{4}X_2) \Rightarrow \]

\[ 4X_1^2(8Y - X_1 - X_2) = X_2^2(8Y - X_1 - X_2) \Rightarrow \]

\[ 4X_1^2 = X_2^2 \Rightarrow 2X_1 = X_2 \]  

(4.A.29)

We use eq. (4.A.29) to obtain an explicit expression for \( C_j \):

\[ \frac{2}{3} = (8Y - \frac{3}{2}C_j) \frac{1}{2}C_j \Rightarrow \]

\[ C_j = \frac{16}{9}Y \Rightarrow C_j = C_L = \frac{4}{9}Y \]  

(4.A.30)

For an individual in the no influence group we thus have:

\[ \sum_{i=1}^{4} C_i = \frac{8}{9}Y \Rightarrow C_i = \frac{2}{9}Y \]  

(4.A.31)

We, thus, find that the group where the leader has no influence slightly increases its contribution, as compared to Nash Simultaneous, when playing against a group that perfectly imitates its leader. On the other hand, the perfect imitation group slightly decreases its contribution (when compared to playing against a group that has a similar structure).

As (expected) payoffs we find for the imitating group members:
\[ E\pi_{PerfectImitation} = \frac{2}{3}(2Y - \frac{4}{9}Y - \frac{2}{9}Y) = \frac{2}{3}(2Y - \frac{2}{3}Y) = \frac{8}{9}Y \quad (4.A.32) \]

and for members of the non-imitation group:

\[ E\pi_{NoInfluence} = \frac{1}{3}(2Y - \frac{2}{3}Y) = \frac{4}{9}Y \quad (4.A.33) \]

So in terms of expected payoff the group that follows their leader perfectly is much better off than the group that neglects their leader. The intuition here is that in a perfect imitation group the players (implicitly) take the positive externalities towards their team members into account (but not the negative externalities towards the members of the other group), this leads to higher contributions and more profit. However, free-riding on the following of others is still profitable.

### 4.A.4 Ties model

As in the main text we start by analyzing modeling the payoff function of the follower. In this model, the distance between the contribution of the leader and a reference point (here the contribution in the Nash equilibrium, zero, is taken as the reference point) generates an affective impulse that is multiplied by a impulse parameter. This impulse results in an interpersonal emotional tie (here expressed as \( \delta C_L \)), that makes a follower care about its leader. Expected payoff of a follower, thus, looks like:

\[
E\pi_F = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j}\left(2Y - \frac{1}{4} \sum_{j=1}^{4} C_j - C_F + \delta C_L (6Y - \frac{3}{4} \sum_{j=1}^{4} C_j - 2C_F - C_L)\right) \quad (4.A.34)
\]

It is assumed that a leader anticipates the effect its contribution has on its followers and cares for the outcomes of the followers from the onset. The latter is captured by a parameter \( \alpha \) that represents the weight the leader attaches to the payoffs of follower.
\[ E\pi_L = \frac{\sum_{i=1}^{4} C_i}{\sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j} \left( 2Y - \frac{1}{4} \sum_{j=1}^{4} C_j - C_L + 3\alpha (2Y - \frac{1}{4} \sum_{j=1}^{4} C_j - C_F) \right) \] (4.A.35)

Now that we have established the expected payoff functions we can use backward induction to solve for the equilibrium. We start by analyzing the followers best response function. We can find it as we did for the sequential game in eq. (4.A.10) using the first-order condition, \[ \frac{dE\pi_F}{dC_F} = 0 \] (from here on we set \( Y = 1 \) for tractability):

\[
\begin{align*}
\sum_{i=1}^{4} C_i &= \left( \frac{2}{3} \sum_{j=1}^{4} C_j - C_F + \delta C_L (6 - \frac{3}{4} \sum_{j=1}^{4} C_j - 2C_F - C_L) \sum_{j=1}^{4} C_j \right) \left( \sum_{i=1}^{4} C_i + \sum_{j=1}^{4} C_j \right)^2 \\
2C_L + 6C_F &= 2 - \frac{7}{4} C_F - \frac{1}{4} C_L + 6\delta C_L - \frac{17}{4} \delta C_F C_L - \frac{7}{4} \delta C_L^2 \\
C_F &= \frac{2 - 2\frac{1}{4} C_L + 6\delta C_L - \frac{1}{4} \delta C_L^2}{\frac{7}{4} + 4\frac{1}{4} \delta C_L} \text{ if } 2 - 2\frac{1}{4} C_L + 6\delta C_L - \frac{1}{4} \delta C_L^2 \geq 0 \text{ else } C_F = 0 \\
\end{align*}
\] (4.A.36)

For the leader we repeat this procedure, but have to keep in mind that its contribution will affect that of its followers:

\[
\begin{align*}
(1 + 3\alpha \frac{\Delta C_F}{\Delta C_L}) \frac{C_L + 3C_F}{C_L + 3C_F + \sum_{j=1}^{4} C_j} &= \\
(1 + 3\frac{\Delta C_F}{\Delta C_L}) \sum_{j=1}^{4} C_j (2 - \frac{1}{4} \sum_{j=1}^{4} C_j - C_L + 3\alpha (2 - \frac{1}{4} \sum_{j=1}^{4} C_j - \frac{1}{3} \sum_{i=1}^{3} C_F) \\
(C_L + 3C_F + \sum_{j=1}^{4} C_j)^2
\end{align*}
\] (4.A.37)

Now we need to find the effect of the contribution of the leader on that of the followers \( \frac{\Delta C_F}{\Delta C_L} \), this effect is obtained by taking the derivative of eq. (4.A.36) to the contribution of the leader \( C_L \):
\[ \frac{\Delta C_F}{\Delta C_L} = \frac{\delta(608 - 434C_L) - 279 - 119\delta^2C_L^2}{(31 + 17\delta C_L)^2} \] (4.A.38)

To find solutions we plug in eq. (4.A.38), as well as values for \( \alpha \) and \( \delta \) into eq. (4.A.37) and solve numerically.

Appendix 4.B Instructions

Below the instructions of the experiment are shown. This example comes from the SE treatment.
Instructions

Welcome to this experiment on decision-making. In this experiment your decisions will influence the money you will earn as well as the earnings of others. All your decisions will be recorded anonymously and you will be paid confidentially. During this experiment you can earn tokens. You will start with 20 tokens as a starting fee. At the end of this experiment your tokens are substituted for Euros at an exchange rate of 10 tokens = 1 Euro.

During this experiment you will be part of a four-person group. Your group will have one leader and will participate in 12 rounds of a decision-making task involving a similar group of four participants. The task will be repeated 12 times. In each of these 12 rounds, your group and leader as well as the other group remains the same. At the start of every round each participant will receive an endowment of 20 tokens.

Task

In every round you have to decide how much of your endowment you want to contribute to a group account and how much to leave for your own private account. The number of tokens in your group account compared with the number of tokens contributed to the group account of the other group determines the chance that your group wins a lottery. If no one contributes to the group account in both groups, there will be no lottery. If your group wins the lottery your group members equally divide the tokens in the private accounts of the other group. These tokens will be added to the tokens in your own private account. If your group loses the lottery, you will earn nothing in that round.

More specifically, your earnings in a round will be determined as follows:

Payoffs

Contribution= Number of tokens allocated to the group account.

Your chance of winning the lottery = (Your group's contributions)/(Your group's contributions + The other group's contributions)

And your winnings look like:
Payoff (if you win the auction) = 20 * Your contribution + 1/4 * (80 - Total contributions other group)

If no one in your as well as the other group contributes to the group account however, all participants earn the 20 tokens they put into their private accounts.

Your total earnings (in tokens) for this experiment will be the 20 tokens you earn at the beginning of the experiment plus the sum of the tokens you earn in the 12 rounds of the decision-making task.

In the quiz after the instructions you will see some examples of how this works in practice.
Feedback

During the game the leader of the group first makes his contribution. Then after his or her group members observe this contribution they simultaneously make theirs. After every round, participants get to see whether or not their group won the lottery, the individual contributions of their group members and the total contribution of the other group. Below you can see a screenshot of a feedback screen:

Results period 1:
Your decision was 3, your group contributed in total 11
The other group also contributed 18 You lost the lottery so your payoffs are 0

Feedback period 2:
Member 1 contributed: 3
Member 2 contributed: 7
Member 3 contributed: 1
Member 4 contributed: 0
Your leader is: Member 1

To the next period

Video Clips

Next to being the first one to contribute in the decision making task the leader has another task in this experiment. Before every block of 3 rounds, so before the first round and before the fourth, seventh and tenth round the leader can choose a basic emotion (anger, disgust, fear, happiness and sadness) or neutral emotional state that will be evoked on his or her group members. Tests indicate that showing a video clip that is specially selected to evoke the selected emotion will do this. The leader gets to see neutral videos. The shortest video is about 30 seconds long, while the longest video lasts for about 4 minutes. During the showing of the video clips we need you to carefully watch these clips. If you do watch these clips when the are shown we will warn you first, if it happens a second time though you will not be able to take part in the rest of the experiment and we will not pay you for it. Below you can find a screenshot of the emotion selection menu.
Your choice of emotion

- Anger
- Disgust
- Fear
- Happiness
- Neutral
- Sadness

Trailor

In the trailer below you can watch short segments of some (not all) video clips that can be shown during this experiment to evoke emotions. The trailer shows clips that bring up the five basic emotions and the neutral state in alphabetical order of these emotions. So the first segment is from a video that evokes anger, the second one from one that targets disgust and the following segments are from videos that respectively target fear, happiness, the neutral state and sadness. Remember though that these short segments not necessarily also evoke these emotions but are shown just so you will have an idea of how these clips look and how they could bring up basic emotions.
Leadership Auction

As mentioned before every group has a leader during this experiment. The leader makes the first contribution in all rounds of the decision-making task and chooses the target emotions of the clips shown to the other group members. To determine who will be the leader of your group there will be an auction. The auction works as follows: Every group member can use the twenty tokens received as a starting fee and can bid zero to twenty tokens. The member that sends in the highest bid wins the auction and pays the second highest bid. So, suppose you bid 3 and the other members respectively bid 5, 11 and 0. Then the one bidding 11 will win the auction and will pay 5 for it, while the others pay nothing. The tokens the leader pays are taken from his or her starting fee.

To the Quiz
Appendix 4.C  Rating the videos

The rating of the different movie clips took place in 4 sessions held in the CREED laboratory of the University of Amsterdam in April 2015. In total 87 participants saw 12 movies each, this lasted for around 45 to 50 minutes for which they were compensated with 10 euros. In total 24 movies were rated. All movies were rated either 43 or 44 times.

The procedure followed was based on the procedures described by Gross and Levenson (1995). After the subjects were welcomed in the reception room they were told that they would see emotion evocating movie clips. Furthermore they were asked to rate in a direct manner and not to ‘overthink’ the level of emotion experienced. The subjects were then assigned to a computer. There they all simultaneously watched the clips and rated them immediately thereafter on a 0-8 Likert scale.

We used the intensity and discreteness criteria (Gross and Levenson, 1995) in the following manner: In order to fulfill the intensity criterium a clip had to score at least a 4 on the target emotion (neutral clips should not score higher than 2 on any emotion). To also fulfill the discreteness criterium the score on the target emotion had to be at least 1 point higher for the target emotion than for the next highest emotion recorded. To order the clips we first looked at the intensity score. When these score were within a half point we sometimes changed the order to prevent clips with a very similar motive to be (potentially) played just after each other. If less than four movies satisfied these criteria for a certain basic emotion we used the best performing clip again (as we did for anger and the neutral state). Since only neutral and happiness were chosen more than twice by some leaders (and all leaders saw only neutral clips), only neutral movies were actually seen twice in some occasions. The overall order of clips was:

112
Table 4.C.1: The order of movies per emotion

<table>
<thead>
<tr>
<th>Anger</th>
<th>Disgust</th>
<th>Fear</th>
</tr>
</thead>
<tbody>
<tr>
<td>My Bodyguard</td>
<td>Pink Flamingos</td>
<td>The Silence of the Lambs</td>
</tr>
<tr>
<td>Cry Freedom</td>
<td>The Fly</td>
<td>The Shining</td>
</tr>
<tr>
<td>Crash</td>
<td>Van Wilder</td>
<td>Psycho</td>
</tr>
<tr>
<td>My Bodyguard</td>
<td>Slumdog Millionaire</td>
<td>Mulholland Drive</td>
</tr>
<tr>
<td>Happiness</td>
<td>Neutral</td>
<td>Sadness</td>
</tr>
<tr>
<td>WALL-E</td>
<td>Alaska’s Wild Denali</td>
<td>The Champ</td>
</tr>
<tr>
<td>Remember The Titans</td>
<td>Searching for Bobby Fisher</td>
<td>The Shawshank Redemption</td>
</tr>
<tr>
<td>Love Actually</td>
<td>Alaska’s Wild Denali</td>
<td>My Girl</td>
</tr>
<tr>
<td>When Harry met Sally</td>
<td>Searching for Bobby Fisher</td>
<td>Saving Private Ryan</td>
</tr>
</tbody>
</table>

The results of all movie clips tested are shown below:
<table>
<thead>
<tr>
<th>Movie</th>
<th>Anger</th>
<th>Disgust</th>
<th>Fear</th>
<th>Happiness</th>
<th>Sadness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska’s Wild Denali (S.E.)</td>
<td>.093</td>
<td>.023</td>
<td>.093</td>
<td>3</td>
<td>.326</td>
</tr>
<tr>
<td>BBC Planeth Earth</td>
<td>0</td>
<td>.931</td>
<td>.614</td>
<td>3.57</td>
<td>.545</td>
</tr>
<tr>
<td>Crash</td>
<td>4.01</td>
<td>3.46</td>
<td>1.81</td>
<td>.326</td>
<td>3.58</td>
</tr>
<tr>
<td>Love Actually</td>
<td>.326</td>
<td>.488</td>
<td>.163</td>
<td>4.88</td>
<td>.419</td>
</tr>
<tr>
<td>My Bodyguard</td>
<td>4.47</td>
<td>2.91</td>
<td>1.86</td>
<td>.116</td>
<td>2.40</td>
</tr>
<tr>
<td>My Girl</td>
<td>.727</td>
<td>.455</td>
<td>.75</td>
<td>.318</td>
<td>6.18</td>
</tr>
<tr>
<td>Pink Flamingos</td>
<td>3.07</td>
<td>6.70</td>
<td>2.16</td>
<td>.744</td>
<td>2.12</td>
</tr>
<tr>
<td>Pride and Prejudice</td>
<td>.232</td>
<td>.163</td>
<td>0</td>
<td>4.23</td>
<td>.442</td>
</tr>
<tr>
<td>Searching for Bobby Fisher</td>
<td>.25</td>
<td>.227</td>
<td>.114</td>
<td>2.48</td>
<td>1.05</td>
</tr>
<tr>
<td>Slumdog Millionaire</td>
<td>1.36</td>
<td>4.80</td>
<td>.568</td>
<td>3.20</td>
<td>1.43</td>
</tr>
<tr>
<td>The Champ</td>
<td>1.49</td>
<td>1.34</td>
<td>1.65</td>
<td>.326</td>
<td>5.93</td>
</tr>
<tr>
<td>The Fly</td>
<td>2.09</td>
<td>6.84</td>
<td>4.35</td>
<td>.163</td>
<td>1.95</td>
</tr>
<tr>
<td>The Shawshank Redemption</td>
<td>1.07</td>
<td>.545</td>
<td>1.52</td>
<td>.568</td>
<td>6.45</td>
</tr>
<tr>
<td>The Shining</td>
<td>.364</td>
<td>.409</td>
<td>4.64</td>
<td>.545</td>
<td>.545</td>
</tr>
<tr>
<td>The Silence of the Lambs</td>
<td>1.26</td>
<td>2.77</td>
<td>4.81</td>
<td>.233</td>
<td>2.05</td>
</tr>
<tr>
<td>Van Wilder</td>
<td>.955</td>
<td>6.66</td>
<td>.477</td>
<td>2.55</td>
<td>.886</td>
</tr>
<tr>
<td>WALL-E</td>
<td>.023</td>
<td>.209</td>
<td>.674</td>
<td>5.77</td>
<td>.488</td>
</tr>
<tr>
<td>When Harry met Sally</td>
<td>.409</td>
<td>.795</td>
<td>.295</td>
<td>4.795</td>
<td>.114</td>
</tr>
</tbody>
</table>

Table 4.C.2: Emotion scores per movie
4.C.1 Instructions rating experiment

Instructions

Welcome to this rating experiment. During this experiment you will rate the emotions you experience while watching video clips. This will earn you in total 10 euros.

You will get to see a total of 12 video clips, all with a duration between 20 seconds and 5 minutes. After every video you will be asked how strong you felt certain emotions. Where an 8 represents the strongest emotions you can imagine while a zero means you didn’t feel this emotion at all. Please report the greatest amount of each emotion you felt during the preceding film. When answering please do not deliberate too much on your choices but answer intuitively.

Before every new video the experimenter will give a signal that indicates that you can now start watching the new video. After this signal please click on the “To the movie” link. Please do so immediately after the signal is given. Make sure though that you wear your headphones as you click on this link, as the video starts immediately after.

During the experiment we would like you to stay attentive, keep your headphones on during the movies and remain focussed on the screen while the videos play. Please also take your time to fill in the questionnaires after carefully.

To the Movie
Chapter 5

Does partner choice foster cooperation?*

5.1 Introduction

In experimental economics it is common to randomly match participants and to keep them from seeing whom they are interacting with. In many everyday interactions, however, people choose their counterpart and, if not, often know whom they interact with. The main reason for not showing with whom people interact is the lack of control of visual effects (like the beauty of a face), reputation, and reciprocity outside the lab, since participants might know each other and will probably interact again either inside or outside the lab. This is a legitimate concern, as has been shown by Andreoni and Petrie (2004), who find that participants are affected—they contribute more—when they can see pictures of their group members in a public good game. With anonymity and no observations of past behavior, however, there is little to base one’s choice on, as participants have no features that can be used to that purpose. Reputation effects are well studied (Camerer, 2003) and are not the subject of this study. Instead, the focus lies on the effect on cooperation of choice based on facial cues. That facial cues are in fact important

*Based on Loerakker (2017). Financial support of the Research Priority Area Behavioral Economics of the University of Amsterdam is gratefully acknowledged.
has been shown already in other settings involving trust (Bonnefon et al., 2013), cooperativeness (Tognetti et al., 2013) and fairness (Van Leeuwen et al., 2017).

To isolate the effect of choice and facial cues, a novel experimental design, using two geographically separated laboratories, is introduced. In essence, I want to find how behavior differs if a person is randomly matched with another individual in an economic game versus a situation in which someone chooses to play the same game with another individual. More specifically, I investigate the effect of partner choice on contribution decisions in a linear public good game and try to find what the determinants are of contributions in such games.

This study confirms that both social preferences as well as beliefs play an important role in social decision making. The roles of facial cues -attractiveness is used here- and partner choice, however, are less clear. Choice has a large, but insignificant effect on the size of the contribution, while attractiveness is important in the choice process but does not seem to affect the contribution.

The organisation of the paper is as follows. In section 2 the relevant literature is discussed. Section 3 presents the experimental design and the hypotheses, while section 4 shows the results. Section 5 concludes.

5.2 Literature

Many studies testing economic theory come from lab environments, although there are some unnatural elements in these environments. One of these elements is that unlike in most everyday (economic) interactions participants in these studies typically do not know or pick their partner. For a person to engage in a cooperative relationship it is important for him or her to know his or her partner. In experimental economics plenty of evidence has been brought up showing that people care about the past behavior of potential (economic) partners and that they take this behavior into account in the new relationship. These effects of reputation are well summarized and described by Camerer (2003), Page et al. (2005), and Palfrey and Prisbrey (1996). A recent
study by Strømland et al. (2016) shows that partner choice, based on reputations and with 
mutual consent, in a repeated prisoner’s dilemma, improves cooperation when compared to 
random matching. 

This study does not deal with reputation, but tries to isolate the effects of partner choice 
when only facial cues are present. There is experimental evidence that facial recognition affects 
behavior: Andreoni and Petrie (2004) show that participants in a linear public good game con-
tribute more to the public good if they see with whom they play. In that experiment, though, 
reputation plays an important part as the games were repeated in a partner setting. Furthermore, 
the authors showed that when not only the pictures but also the individual contributions instead 
of the group contribution were shown total contributions increased even further. In a subse-
quent study Andreoni and Petrie (2008) provide evidence that beauty and attractiveness play 
important roles. They show that in a similar linear public good game participants are contribut-
ing more when in a group with more attractive people. Interestingly, they also seem to expect 
more beautiful people to contribute more. This could be of special importance to this study as 
attractiveness could thus be used as a selection criterion. This finding finds further support in 
the work of Mulford et al. (1998), who find that people are more likely to enter play and to 
cooperate with attractive people in an exchange game. Related work by Eckel and Petrie (2011) 
finds that participants in a trust game are willing to spend money in order to find out whom 
they are paired with. This finding indicates that people do believe that they can infer certain 
characteristics from no more than a picture. 

Recently, more studies have shown that people are indeed able to recognize all kinds of 
behavioral characteristics from pictures of others, or at least believe to be able to. Tognetti et al. 
(2013) show that French raters were able to identify from pictures which males contributed 
least to a linear public good game in rural Senegal. Bonnefon et al. (2013) and Van Leeuwen 
et al. (2017) show in a series of experimental studies that participants are, even in the presence 
of cognitive load, able to detect trustworthiness, but become worse at this when the depicted 
others have more distinctive hair and clothing. Oddly enough, when given the choice between a
picture with or without these ‘distractions’ people do prefer them with clothes and hair, unjustly believing this will provide them with more information. Finally, Van Leeuwen et al. (2017) show that people are significantly better than chance at telling from just a picture who will reject unfair offers in an ultimatum game. Together these studies show that people are in fact able to read certain characteristics from a picture of a face, and, perhaps more importantly, already behave accordingly if they only believe they are able to do so.

Next to these effects described above, there might, however, also be an effect of partner choice in itself. Norton et al. (2012), for example, show that people value objects that they have built or put together themselves more than objects they have bought as a finalized piece. In this study this might translate to a difference in attachment towards partners that are selected and partners that are allocated through a random mechanism.

5.3 Design and hypotheses

In this section I will outline the experimental methods used as well as the argumentation for this specific design. This is followed by a disposition of the hypotheses.

5.3.1 Experimental procedures and treatments

The experiment took place simultaneously in two geographically seperated labs, the CREED laboratory in Amsterdam and the CentER laboratory in Tilburg in October en November 2015. In total 312 participants, recruited through either the CentER or the CREED recruitment system, participated and earned on average 13.13 euro. Sessions typically lasted for around 45 minutes to an hour. In the invitation email as well as in the reminder email(s) and during the welcome in the lab participants were informed that they would have to take a picture of themselves using a webcam and that, after this picture was approved as sufficiently neutral, this picture would be used in the other lab, where another participant could see it. After the welcome in the lab participants signed consent forms, that allowed us to send the pictures to the
other lab and to study them afterwards. The experiment consisted of 5 different tasks, followed by a questionnaire. After the instructions (for an example of the instructions see appendix A), in which the participants were fully informed about the entire experiment, the participants first, after a short quiz to test their understanding of the instructions, had to do a variation of the circle test (1), used before by Linde and Sonnemans (2012), with eight different potential partners from the other lab (more on the circle test below). In the next part the participants had to state beliefs (2) about the contribution of the same eight potential partners in the one-shot public good game that they would play in the main task of the experiment. They where asked to predict what these potential partners would contribute if they would be matched with them (more on the task and matching procedures below). Both for the circle test as well as the belief elicitation participants were only paid for the choice they made with their eventual partner. Participants earned one euro if they predicted the contribution of their partner correctly. After the belief elicitation participants were matched with their partner for the main task. The experiment comprised two treatments, Random (N=96) and Choice (N=216), in a between subjects design. These treatments only differed in the way the participants were matched. In the former treatment participants were matched randomly, while in the latter some participants had the opportunity to choose their partner. Then, in the main task, participants played a simple one-shot two-player linear public good game (3). Subsequently, the participants were once more asked to do the circle test (4) and state their beliefs (5), but now regarding their partner from the main task. Finally, participants were asked to rate the attractiveness of their partner in the main task and to fill out a small questionnaire. Participants received feedback only at the end of the experiment and were paid for every task of it.

In the following subsections the different tasks of the experiment as well as the matching procedures will be discussed in greater detail.

**Circle test**

The circle test used in this experiment was identical to the one used by Linde and Sonnemans
(2012) and very similar to the version introduced by Sonnemans et al. (2006). The circle test is a simplification of the ring test introduced by Liebrand (1984), which makes it fit for applications involving multiple counterparts. In general, circle tests are used to measure a participant’s social value orientation or the social tie with a (generalized) other person. In our application participants saw a picture of a potential partner and had to select a point on a circle with a radius of 1500 points, with each 3000 points worth 1 euro. The value on the x-axis represented the payoff to the participant him- or herself and the value on the y-axis the payoff for the potential partner. For all payoff combinations the relationship \( x^2 + y^2 = 1500^2 \) holds. The potential partners were presented in a random order and the circle test was presented without a default choice. Participants could select a point on the circle, observe the outcome and then decide to select another point multiple times, if so desired. Only when a participant clicked on Send a decision was recorded. From this decision one can derive the weight a player attaches to the payoff of his or her potential partner, by calculating the tangent of his or her own payoff and that of his or her partner. This value, denoted by \( \alpha \), is used to see what the effects are of (the changes in) social preferences.

**Main task**

The main task consisted of a one-shot linear public good game with one partner from the other lab. During the task the participants saw a picture of their partner, that was taken at the beginning of the experiment. At the beginning of the task all participants got 8 tokens. These tokens were worth 1 euro each if a participant kept them in his or her private account, while, in the common account, these tokens were worth 0.75 cents to both him- or herself and the partner. This lead to the following payoff function for a participant \( i \) with counterpart \( j \):

\[
\pi_i = 8 - C_i + 0.75(C_i + C_j)
\]  \hspace{1cm} (5.3.1)

where \( C_i \) and \( C_j \) stand for the contributions of, respectively, \( i \) and \( j \). Participants had to decide how many tokens to put into the common account.
Matching procedures

In both treatments the matching procedure was common knowledge. After the belief elicitation task players got to see two randomly selected participants out of the eight participants they had seen before. This was done pairwise so there were always two players in one lab ‘connected’ with two players from the other lab. Subsequently, in the Random treatment, players were randomly matched with one of these two. In the Choice treatment, however, half of the players in one of the two labs got to choose the partner he or she wanted, while the other half of the participants in that lab got to play with the participant that was left over. Participants in the other lab were informed about the fact that some participants in the opposite lab had the opportunity to choose a partner and, later on, that they were either chosen or not.

5.3.2 Hypotheses

This subsection starts with introducing some terminology that will be used throughout this paper. The term ‘choice side’ refers to the pairs in the Choice treatment in the lab where half of the participants could choose their partner (I use the pairs that saw the same pictures as a unit of observation, in that instance). It should be kept in mind, though, that when a reference is made to the Choice side half of the participants that are referred to had no possibility to pick a partner. The ‘chosen side’ similarly refers to the pairs that could be chosen by the ones with the choice option. On the choice side the individuals with this option are named ‘Choosers’, while the participants that are not able to choose on that side are called ‘Passives’. On the chosen side the individuals chosen are referred to as ‘Chosen’, while the ones that are not chosen are named ‘Leftovers’. When discussing the contributions in Choice I will use the scores of both the choice side and the chosen side, effectively take the average of four participants as one observation. For the Random treatment, though, every single contribution is treated as an independent observation.

Contributions are, however, not the only measure of interest. I also want to find out what the drivers are of these contributions and, perhaps more importantly, identify what causes the
potential differences between treatments and roles. Ex-ante there are a multitude of channels through which the treatments and roles could affect the contributions. The possible channels that will be investigated in this paper are: 1) Participants might pick a partner whose payoff they value more; 2) Participants might pick a partner that they believe will contribute more; 3) Participants might pick players that they find more attractive, and consequently contribute more; 4) Choosing somebody might make participants value the payoff of their partner more; 5) Being chosen might make Chosen to like the Chooser more; And, finally, 6) there could be an effect of choice itself, as participants could become more cooperative if they are in an environment with a human choice mechanism compared to an environment with a random choice mechanism. Channels 1), 2), and 3) can only have an effect if their is significant heterogeneity in the population. If Choosers and Passives have identical beliefs, preferences, and attractiveness ratings, it does not matter what the matching will be, while if people have very heterogeneous beliefs and preferences it is likely that the matching procedure described above leads to on average higher beliefs, preferences, and attractiveness ratings on the choice side. In order to test the importance of these channels I need information about the beliefs of the participants, the relative importance of the payoff of the partner for the participants (before and after the partner is allocated in order to see if 4) is of importance), and the attractiveness of their partner. The experiment described above is designed in order to generate exactly these measures.

Next I will present and motivate the hypotheses.

**Hypothesis 5.1.** Contributions in the Choice treatment will be higher than in the Random treatment.

The arguments for this hypothesis are discussed above where the different channels that could lead to this (potential) difference are explained.

**Hypothesis 5.2.** Participants that could choose a partner will believe their partner contributes more than players that could not choose their partner.

Hypothesis 2 is important, as channel 2 can only be of influence if this hypothesis is true. Intuitively it seems sensible as players would simply earn more if they play with someone who
actually contributes more. If players are also conditional cooperators, for which evidence was presented by Fischbacher et al. (2001), than this should also lead to Choosers contributing more than Passives.

**Hypothesis 5.3.** Participants that could choose a partner will end up with a more attractive partner than players that could not choose their partner.

Hypothesis 3 is intuitive as people are attracted to more attractive people. The validity of this hypothesis is essential for channel 3 to have an effect. Furthermore, Andreoni and Petrie (2008) show that people reward beauty, but, perhaps more importantly for this study, do also expect more beautiful people to contribute more to a public good.

**Hypothesis 5.4.** Participants that could choose a partner will increase the valuation of the welfare of their partner more than players that could not choose their partner.

I expect players that make a conscious choice for a partner to start liking this partner more because of this choice. This would lead to a higher $\alpha$-value for Choosers. This effect would be similar to the so-called IKEA-effect (Norton et al., 2012) in psychology. This effect describes how people that build something by themselves feel more attached to this object than people who bought the exact same object. In our setting this could lead to people who selected their partner themselves feel more attached to this partner than players that did not.

**Hypothesis 5.5.** In the Choice treatment Choosers are expected to contribute more than Chosen, who are in turn expected to contribute more than Passives. Finally, Passives are expected to contribute more than Leftovers.

I expect Choosers to make the highest contributions. This as a result of the three previous hypotheses. They can pick a partner that they find attractive, a partner they believe will contribute more, and are likely to become more attached to this partner because of their conscious choice for him or her. All of these effects could lead to Choosers contributing more than others in Choice as well as the participants in Random. Also for the Chosen I expect higher contributions than for the participants in Random as they were favored over someone else by the
Choosers, this could both lead to more attachment to these Choosers. As being chosen as a partner, in a situation that allows for cooperation, by another individual could be seen as an act of kindness and trust. If the Choosers are in fact able to tell from their faces which participants are more cooperative, this would also result in higher contributions in Chosen. For the Passives and the Leftovers it is more difficult to say whether or not they would contribute more or less than participants in Random. The fact that a partnership is due to a decision of another could lead to a stronger bond, but to be either passive while others have a choice or to be rejected might lead to negative emotions and uncooperative behavior. I expect the latter to have an effect and therefore that Passives will contribute more than Leftovers. Lastly, I note that if Choosers are indeed able to see from pictures which participants are cooperative we expect lower contributions for Leftovers as well as for Passives (if they are conditional cooperators), when compared with Random.

5.4 Results

In this section I will first look at the overall behavior on the treatment level, then turn to the role level, and afterwards discuss some estimation results in the same order. The last part of this section is devoted to the partner choice stage. The behavior of the Choosers in this stage gives an insight in what channels the Choosers find ex ante important and about how heterogeneous the participants are in their beliefs and preferences.

5.4.1 Treatment results

In the Choice treatment the average contribution is 4.46 tokens, while in the Random treatment the average contribution is 3.99 tokens. This difference, illustrated in figure 1, is statistically insignificant using a Mann-Whitney Wilcoxon test (MWW, p>0.1). ¹Although in the right di-

¹When determining the significance of this difference I used the average contribution of a group of one Chooser, one Passive, one Chosen, and one Leftover as unit of observation. At the treatment level this method is used for all variables.
rection, this result rejects hypothesis 1. Given the size of the difference between the treatments, however –more than 10% of the contributions in Random and more than 5% of the highest possible contribution– it seems worth investigating further.

Below, in table 1, descriptive statistics for the treatment level are displayed. In this table, \( \Delta \alpha \) is the change in \( \alpha \)-value (measured before and after the main task), while \( \Delta \text{Belief} \) is the change in belief. I find that only \( \Delta \alpha \) is (weakly) significantly higher in Choice and that although for other variables such as the Belief, Attractiveness and the contribution the level is higher, but that these differences are not significant. This does neither support the second hypothesis, that participants will choose partners that they expect to contribute more, nor hypotheses 3 and 4 that Choosers end up with more attractive partners or that the active choice of a partner makes a participant value a partner’s welfare more (hypotheses 3 and 4). Later on I will try to confirm this by first looking at the different roles and then by estimating what determines the size of the contribution using a regression. Finally, it is interesting to note that (changes in) beliefs do not seem to differ much between the treatments. This could be due to the fact that beliefs are rather
homogeneous, for example, because people pick their partner based on other characteristics (attractiveness?), or that beliefs do not play much of a role at all (we will later find out that this is certainly not the case).

Table 5.1: Descriptive statistics per treatment

<table>
<thead>
<tr>
<th>Average per treatment</th>
<th>Random (N=96)</th>
<th>Choice (54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution</td>
<td>3.99 (2.57)</td>
<td>4.46 (2.43)</td>
</tr>
<tr>
<td>$\alpha$ (before game)</td>
<td>35.7 (31.7)</td>
<td>33.3 (30.6)</td>
</tr>
<tr>
<td>Belief (before game)</td>
<td>3.80 (2.01)</td>
<td>3.97 (2.38)</td>
</tr>
<tr>
<td>$\Delta \alpha$</td>
<td>-0.55 (31.4)</td>
<td>3.13 (28.7)*</td>
</tr>
<tr>
<td>$\Delta $Belief</td>
<td>0.30 (2.10)</td>
<td>0.17 (2.13)</td>
</tr>
<tr>
<td>Attractiveness</td>
<td>2.89 (2.41)</td>
<td>3.26 (2.32)</td>
</tr>
</tbody>
</table>

Note: standard deviations between brackets; * indicates statistical significance between treatments at the 10% level, using a MWW-test.

5.4.2 Results per role

In table 2, the same statistics that were shown and discussed above on the treatment level are now shown for the different roles.

Table 5.2: Descriptive statistics per role in choice

<table>
<thead>
<tr>
<th>Avg per role</th>
<th>Random(N=96)</th>
<th>Choosers (54)</th>
<th>Passives (54)</th>
<th>Chosen (54)</th>
<th>Leftovers (54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution</td>
<td>3.99 (2.57)</td>
<td>4.94 (2.47)**</td>
<td>4.5 (2.33)</td>
<td>4.07(2.41)</td>
<td>4.33 (2.50)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>35.7 (31.7)</td>
<td>37.9 (24.1)</td>
<td>34.4 (33.2)</td>
<td>34.8 (34.0)</td>
<td>29.0 (29.8)*</td>
</tr>
<tr>
<td>Belief</td>
<td>3.80 (2.01)</td>
<td>4.24 (2.35)*</td>
<td>3.69 (2.59)</td>
<td>3.76 (2.65)</td>
<td>4.20 (2.38)</td>
</tr>
<tr>
<td>$\Delta \alpha$</td>
<td>-0.55 (31.4)</td>
<td>3.45 (20.7)</td>
<td>3.25 (29.0)</td>
<td>0.97 (30.6)</td>
<td>4.90 (33.7)</td>
</tr>
<tr>
<td>$\Delta $Belief</td>
<td>0.30 (2.10)</td>
<td>0.39 (2.31)</td>
<td>0.39 (1.85)</td>
<td>0.02 (2.10)</td>
<td>-0.11 (2.27)</td>
</tr>
<tr>
<td>Attractiveness</td>
<td>2.89 (2.41)</td>
<td>3.65 (2.16)*</td>
<td>3.55 (2.59)*</td>
<td>2.80 (2.41)</td>
<td>3.04 (2.07)</td>
</tr>
</tbody>
</table>

Note: standard deviations using average errors per individual between brackets; *, ** indicates significance between treatments at the 10% and 5% level, respectively, MWW-test.

The first statistic that catches the eye is the contribution of the Choosers. Choosers contribute on average almost a full token more (4.94 vs. 3.99) than participants in Random. When looking ahead for potential channels that could cause this difference, two other statistics stand out. Firstly, I find that Choosers find their partners (weakly) significantly more attractive than participants in Random, showing only weak support for the third hypothesis. This could indi-
cate that attractiveness is an important factor in the choice for a partner. Secondly, it seems that Choosers also believe that their partners will contribute more in the game. As this difference is only weakly significant this neither confirms nor rejects the second hypothesis. These two results could have separate causes, but could also have a common cause if attractiveness is seen as a proxy for cooperativeness, as found by Andreoni and Petrie (2008). In the next subsections I will try to answer this question. Choosers however do not value the welfare of their partners higher than others, leaving no support for the fourth hypothesis. Although Choosers contribute most, I do not find support for the differences in contributions predicted in the fifth hypothesis, as Passives and even Leftovers contribute more than Chosen. A potential caveat, however, is that all differences between these roles are insignificant for the non-Choosers. Another finding is that Passives also seem to find their partners to be significantly more attractive. This could imply that there is quite some heterogeneity in the population.

5.4.3 Estimation results at treatment level

In this subsection I will evaluate the key determinants for the contributions. When looking at the treatment level, I use again the average of the statistics per group in Choice as unit of observation. I first estimate\(^2\) the effect of each variable ($\alpha$, Belief, $\Delta \alpha$, $\Delta$Belief, Attractiveness), in combination with the treatment dummy, on the average contribution, using an ordinary least squares regression for each separate variable, I find all variables but Attractiveness ($p \approx 0.07$) to be significant at the 1% level, with the treatment dummy typically being insignificant. This is an interesting finding in itself, as it indicates that participants might choose attractive partners but may not behave more or less cooperative when matched with an attractive partner.

To see what the total effect of these variables are on the contribution we add them all together to the regression. The results of this regression are shown in table 3 below.

\(^2\)For these as well as all other regressions I also used an Interval Regression (as the dependent variable is censored on both sides). However, because these estimates are extremely similar to the OLS estimates, I only report the latter.
The estimation on the treatment level makes clear that the ‘usual suspects’, beliefs and preferences do indeed play a significant part in determining the contribution. What is perhaps more surprising is that attractiveness, when correcting for beliefs and preferences, does not seem to be important. Beliefs, however, seems to mediate attractiveness, as attractiveness is a significant predictor (p<0.04 using OLS, result is robust against controlling for role) of Belief, while Belief has a significant effect on the contribution. Another finding of importance is that, when controlling for (the change in) beliefs, preferences, and attractiveness, the treatment effect is insignificant.

### 5.4.4 Estimation results at role level

When introducing the different roles into the estimation, the results do not change much as can be seen in table 4.
Remarkably, even though not significant, the difference in contribution between the participants in the different roles in Choice and Random still remains quite large (about 10%), even when controlling for preferences, beliefs and attractiveness. When estimating the effects of the different variables for the separate roles the picture is very similar to that of table 4, with all variables but attractiveness being significant at, at least, the 5% level.

5.4.5 Behavior of Choosers

The next issue to sort out is to see what determines the choice of Choosers. Unfortunately, I only have an attractiveness measure for the actual partners and not for the ‘would-be’ partners of the Choosers. The fact that the attractiveness is highest for the Choosers and significantly higher than that of the participants in Random is a strong indication though that attractiveness might very well matter in the partner choice (see table 2). Therefore I can only estimate the effect of beliefs and preferences on the choice of the Choosers. This is done using a logit model, where the choice for the potential partner on the left of the choice screen is explained by the difference in the beliefs and preferences of the Chooser regarding, respectively, the contribution and welfare of that potential partner and the potential partner on the right. In table 5 the results of this estimation are displayed, where ΔBelief-P and Δα-P, respectively, denote the differences.
in beliefs and preferences between potential partners.

Table 5.5: Estimation of the effect of beliefs and preferences on the partner choice

<table>
<thead>
<tr>
<th></th>
<th>Choice for potential partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔBelief-P</td>
<td>0.54 (0.28)*</td>
</tr>
<tr>
<td>Δα-P</td>
<td>0.03 (0.02)*</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.11 (0.30)</td>
</tr>
<tr>
<td>Observations</td>
<td>54</td>
</tr>
</tbody>
</table>

Note: standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1

Both beliefs and preferences seem to be important, but the effects are only weakly significant. For the beliefs, this seems largely due to the fact that for the majority (39 out of 54) of the Choosers there is no difference in belief for their potential partners. When I estimate the same equation, using their actual partner as the dependent variable, for the other roles neither the difference in beliefs nor that in preferences is significant at the 10% level. This is to be expected as all the others could not choose their partner.

5.5 Conclusion

The experimental evidence presented here gives little support for the effect of choice itself. People who get to choose their partners do contribute more than others (table 2), but this works not via the the direct effect of choice, but through beliefs and preferences (with attractiveness as a mediator). Beliefs, and to a lesser extend preferences, however, do play an important role in the decision to contribute as well as the choice of a partner, as can be seen from tables 3, 4, and 5. Furthermore, beliefs seem to mediate the effects of attractiveness, since attractiveness is an important factor in choosing a partner, but not in determining the contribution. Besides these results I find no evidence of a change in the valuation of the welfare of the partner due to a conscious choice for this partner (hypothesis 4). A potential cause for the relative unimportance of preferences in the choice decision is the fact that if a participant chooses to play with a more likeable partner it is not only an opportunity to be cooperative and help someone you like more,
but it also increases the price of being selfish, as being selfish is more costly if it comes at
the cost of someone you like or find attractive. Furthermore it is interesting to note that both
Passives and Leftovers do not seem to be negatively affected by the fact that, respectively, the
choice of their partner is in the hand of another person or that they are not chosen while they
could have been. As both contribute more than participants in Random, and, perhaps more
surprisingly, even more than Chosen. All these differences are however insignificant. The
strongest indication that there might be a direct effect of having the ability to choose a partner
comes from the finding that when controlled for attractiveness, beliefs, and preferences, as well
as the change in beliefs and preferences Choosers on average still contribute 10% more than
participants in Random (table 4). Further investigation into the robustness of this effect seems
of interest, therefore, but is left for future research.

Appendix 5.A Instructions

Below an example of the instructions of the experiment are shown. This example comes from
the Choice treatment.
Instructions

Welcome to this experiment on decision-making. In this experiment your decisions will influence the money you will earn as well as the earnings of others. All your decisions will be recorded anonymously and you will be paid confidentially. As mentioned before the pictures of your (potential) partner will be shown during this experiment, just as your picture will be shown to potential partners in the other lab.

This experiment is centered around one main task. In this task you will be paired with a partner, in the main task you will see a picture of the player you will be matched with. The second and third task are tasks that involve your potential partners for the main task. The fourth and fifth task also involve your partner.

Note: All the pictures you will see are from participants who are currently sitting in the other lab. Your picture will only be shown to participants in the other lab and you will only interact with participants in the other lab.

First task

In this task you will sequentially be shown 8 potential partners, among them is your partner for the main task. For each of them you will have to choose a point on a circle. This point allocates points to both you as well as your potential partner. On the horizontal axis you find the points allocated to you, while the vertical axis shows the number of points for the other (both axis range from -1500 till 1500). Below you find an example of such a circle task:

![Task Example](image)

At the end of the experiment we will pay you 1 euro per 3000 points that you allocated to yourself when paired with your future partner in the main task. Your partner will also be paid 1 euro per 3000 points you allocated to him or her. Similarly you will be paid 1 euro per 3000 points your partner allocated to you in the same task. The points you allocate to yourself and your potential partners that turn out to be not your actual partner will not earn you or your potential partners any money. Remember though that at this point it is not yet determined who your partner in the main task will be.

Second Task

In the second task you again see the same 8 potential partners as in the first task. Now though, you will be asked what contribution you expect your potential partner to make if he or she would be matched with you in the main task. If your expectation is correct, this can obviously only be tested for your actual partner, this will earn you one euro.

Partner

In this experiment you do not have the opportunity to select your partner. Half of the participants in your lab however do. After the second task you as well as another participant in this lab will get to see a random set of two out of the eight potential partners you saw in the first two tasks. This other participant gets to choose which one will be their partner in the main task. The participant that is not picked will be matched with you.

Main Task

In the main game both you as well as your partner will get 8 tokens. These tokens are each worth 1 euro. During the task you and your partner have to decide how many of these 8 tokens you want to contribute to a common account. The tokens in the common account are multiplied by 1.5 and afterwards equally split between you and your partner. So if you contribute 4 tokens and your partner contributes 2 there are in total 6 tokens contributed to the common account. They will be multiplied by 1.5, making 9, now these 9 tokens are divided by 2 earning both players 4.5 euros. In total you will earn 8.5 euros while your partner earns 10.5 euros in this example.

Note: You will get to know the results from this game as well as your earnings from the other tasks only at the end of the experiment and so will your partner.
Tasks After

After the main task you will again do the circle task with your partner from the main task as the other, with a similar payment structure as in the first task. Thereafter you will be asked what you expect your partner contributed in the main task. If your expectation is correct you will earn one euro.

Questionnaire and Results

After these last task we kindly ask you to fill out a short questionnaire. After you did this you will get to see your final results and if everybody (in Amsterdam as well as in Tilburg) is done with their questionnaires we will pay you confidentially and one by one.

Quiz

Before we start we like you to do a short quiz to see if you fully understood the instructions.

To the quiz
Bibliography


Samenvatting (Summary in Dutch)

Menselijk gedrag is moeilijk te vatten in eenvoudige modellen. In eenzelfde omgeving vertonen verschillende mensen vaak tegenovergesteld gedrag en doen kleine veranderingen in deze omgeving dit gedrag soms radicaal veranderen, terwijl grote veranderingen soms volledig worden genegeerd. De verklaring voor deze bevindingen worden traditioneel op twee manieren benaderd. In de (gedrags-)economie wordt een individu gereduceerd tot een set overtuigingen, preferenties en een set mogelijke acties, waarbij de, niet altijd rationale, hersenen het lichaam volledig controleren. Aan de andere kant staan de psychologie en de neurowetenschappen, waar hormonen, emoties en andere fysieke prikkels de mens drijven en rationaliteit een kleinere rol speelt. Dit proefschrift probeert deze twee benaderingen dichter bij elkaar te brengen door te laten zien rationele en emotionele elementen samen gemodelleerd kunnen worden. Verder onderzoekt deze thesis de rol van emotioneel leiderschap in conflict situaties en het effect van partnerkeuze op het coperatie niveau als men de partner enkel op basis van fysieke kenmerken kan kiezen.

Een van die, moeilijk door orthodoxe economische theorie verklaarbare, waarnemingen is herhaaldelijk destructief gedrag dat niet alleen ten koste gaat van de omgeving, maar ook de actor zelf schaadt. Hoewel we dit gedrag waarnemen bij de verspreiding van computervirussen, vandalisme en langlopende vetes, is dit gedrag Weinig bestudeerd in de economische wetenschap. Recent zijn er echter studies uitgevoerd waaruit blijkt dat mensen preferenties kunnen ontwikkelen voor gedrag dat anderen schaadt. Voorbeelden van deze studies zijn Zizzo (2003), Nikoforakis en Engelmann (2011) en Bolle et al. (2014). In veel van deze studies is de meest
côöperatieve actie die een proefpersoon kan ondernemen de actie uit het (standaard) Nash evenwicht (denk hierbij bijvoorbeeld aan het accepteren van een aanbod in een *Ultimatum game*). Om proefpersonen de kans te geven om zowel coöperatief als destructief gedrag te tonen introduceeren we in hoofdstukken 2 en 3 een *Fragile Public Good* (FPG) spel. Het FPG spel is een tweepersoons niet-lineair *public goods game* met een gelijkblijvende stijging van de marginale kosten. In dit spel kunnen deelnemers echter niet alleen bijdragen aan een gemeenschappelijke rekening, maar hier ook tokens (geld) uithalen. Dit heeft hetzelfde negatieve effect op de speler zelf als bijdragen, maar heeft precies het omgekeerde effect van bijdragen op de andere speler, het schaadt de ander evenveel als bijdragen de ander oplevert. Deze symmetrie maakt het de spelers mogelijk om zowel negatieve of destructieve voorkeuren te uiten als positieve of coöperatieve voorkeuren.

In hoofdstuk 2 worden het ontstaan en de ontwikkeling van wederzijdse destructieve relaties onderzocht. Verder is er aandacht voor *framing*, waarbij we kijken of (min of meer) hetzelfde spel in een andere context of met andere bewoordingen ander gedrag oplevert. Hierbij vinden we dat als we het spel negatief *framen*, door het Nash evenwicht te verplaatsen naar een situatie waarbij beide spelers een token van de gemeenschappelijke rekening halen, we bijna twee keer zoveel destructieve keuzes (ongeveer 20% van de totale keuzes) zien als in de symmetrische variant, waar spelers in het Nash evenwicht niets doen. Verder vinden we steun voor het *Ties*-model van Van Dijk en Van Winden (1997). Dit model, dat centraal staat in hoofdstuk 3 en wordt gebruikt in hoofdstuk 4, beschrijft de ontwikkeling van interpersoonlijke preferenties tussen twee individuen. In het kort beschrijft het model dat wanneer de een aardiger is voor de ander dan de ander verwacht (in een neutrale omgeving, deze verwachting moet gezien worden als een referentiewaarde en niet als de daadwerkelijke verwachte actie op dat moment), de ander in het vervolg meer rekening gaat houden, of een groter gewicht gaat toekennen aan de uitkomsten van de een. Het model staat ook negatieve *ties* toe. Mocht een speler een negatieve *tie*-waarde ($\alpha$) hebben dan is hij of zij dus bereid zijn opbrengst te verlagen als dat leidt tot een voldoende grote afname in de opbrengst van de ander. In het experiment zien we dat paren die
herhaaldelijk negatieve bijdrages doen in een negatieve spiraal geraken en daar steeds moeilijker uitkomen, hetgeen in lijn is met het *ties*-model.

Hoofdstuk 3 bouwt voort op de studie van Bault et al. (2017) en analyseert het voorspelend vermogen van het *ties* model. Verder worden vooruitkijkende spelers geïntroduceerd en het model uitgebreid door een verschil in de kracht van positieve en negatieve impulsen toe te staan. De belangrijkste bevindingen zijn: dat het *ties*-model het gedrag van onze proefpersonen beter voorspelt dan zowel een standaard model met gelijkblijvende sociale preferenties als een gevestigd leer model, dat de introductie van vooruitkijkende spelers de prestaties van het model niet significant verbetert (iets dat mogelijk wordt verklaard door de manier waarop we vooruitkijkende spelers modelleren) en dat, tegen onze hypothese in, negatieve impulsen een groter absoluut effect hebben op preferenties dan positieve impulsen. Tot slot, wordt in hoofdstuk 3 ook geïllustreerd hoe een eenvoudig *ties*-model de ad-hoc veranderingen van strategie, zoals gevonden door Dal Bó and Fréchette (2011) en Fudenberg et al. (2012), kan verklaren.

Hoofdstuk 4 bestudeert de effecten van emotie en leiderschap in een *conflict game*. In het door ons ontworpen spel spelen twee groepen van vier, twaalf rondes tegen elkaar. Elke speler moet in elke ronde kiezen hoeveel van zijn 20 tokens hij of zij voor zichzelf houdt en hoeveel hij of zij aan het conflict besteedt. Na elke ronde bepaalt een loterij welke groep het conflict won. De kans dat een groep de loterij wint is gelijk aan de totale bijdrage aan het conflict van die groep gedeeld door de totale bijdrage aan het conflict van de beide groepen samen. De leden van de winnende groep behouden bij winst de tokens die zij voor zichzelf hebben gehouden en krijgen ieder een kwart van de tokens die de andere groep buiten het conflict heeft gehouden. De groepen hebben in het experiment een leider die als eerste een, alleen voor zijn groepsgenoten zichtbare, bijdrage aan het conflict levert. Deze leiderschapspositie wordt in het experiment via een veiling verkocht. Deelnemers aan dit experiment kregen voor de eerste en daarna na elke drie rondes een filmpje te zien, dat voor de volgers (de spelers die niet de leider zijn) in twee van de drie condities emotioneel geladen is. In een van de twee ‘emotionele condities kon de leider niet alleen als eerste bijdragen, maar ook kiezen welke emotie er werd opgeroepen in de filmpjes
van zijn volgers. Anders dan voorspeld door standaard Nash modellen, vinden wij dat leiders in alle condities meer bijdragen dan volgers (het standaard model voorspelt zelfs dat leiders helemaal niet bijdragen), dat volgers hun leiders in hoge mate volgen en dat, hoewel emoties wel degelijk een effect lijken te hebben op het contributieniveau van volgers, leiders hier geen strategisch gebruik van maken. Verder tonen we aan dat de door ons gevonden resultaten door andere modellen uit de gedragseconomie, zoals bijvoorbeeld het eerder genoemde ties-model, verklaard kunnen worden.

In hoofdstuk 5 bestudeer ik het effect van partnerkeuze op het coöperatieniveau in een lineair public goods game. Dit is in het bijzonder van belang omdat proefpersonen in de meeste economische experimenten met anonieme partners interacteren, terwijl dit in de wereld buiten het lab een uitzondering is. Om de potentiële problemen van deze keuzevrijheid, reciprociteit buiten het lab en het niet kunnen uitsluiten van reeds voor het experiment bestaande relaties tussen partners, vond dit experiment simultaan in twee verschillende laboratoria in twee verschillende steden plaats. Hoewel proefpersonen in de conditie waarin zij hun partner in het spel kunnen kiezen bijna 10% meer bijdragen dan in een conditie zonder partnerkeuze, is dit verschil niet significant. Verder worden de bijdrages vooral verklaard door de verwachte bijdrage van de ander (wat aangeeft dat mensen enkel conditioneel willen bijdragen) en de initiële tie (hoe meer gewicht iemand aan de opbrengst van de ander toekent, hoe meer iemand bereid is om bij te dragen). Daarnaast blijkt iemands (fysieke) aantrekkelijkheid wel van belang voor de keuze van de partner, maar niet voor de hoogte van de bijdrage.
The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

657 M. KOUDSTAAL, Common Wisdom versus Facts; How Entrepreneurs Differ in Their Behavioral Traits from Others

658 N. PETER, Essays in Empirical Microeconomics

659 Z. WANG, People on the Move: Barriers of Culture, Networks, and Language

660 Z. HUANG, Decision Making under Uncertainty-An Investigation from Economic and Psychological Perspective

661 J. CIZEL, Essays in Credit Risk, Banking, and Financial Regulation

662 I. MIKOLAJUN, Empirical Essays in International Economics

663 J. BAKENS, Economic Impacts of Immigrants and Ethnic Diversity on Cities

664 I. BARRA, Bayesian Analysis of Latent Variable Models in Finance

665 S. OZTURK, Price Discovery and Liquidity in the High Frequency World

666 J. JI, Three Essays in Empirical Finance

667 H. SCHMITTDIEL, Paid to Quit, Cheat, and Confess

668 A. DIMITROPOULOS, Low Emission Vehicles: Consumer Demand and Fiscal Policy

669 G.H. VAN HEUVELEN, Export Prices, Trade Dynamics and Economic Development

670 A. RUSECKAITE, New Flexible Models and Design Construction Algorithms for Mixtures and Binary Dependent Variables

671 Y. LIU, Time-varying Correlation and Common Structures in Volatility

672 S. HE, Cooperation, Coordination and Competition: Theory and Experiment

673 C.G.F. VAN DER KWAAK, The Macroeconomics of Banking

674 D.H.J. CHEN, Essays on Collective Funded Pension Schemes
675 F.J.T. SNIEKERS, On the Functioning of Markets with Frictions

676 F. GOMEZ MARTINEZ, Essays in Experimental Industrial Organization: How Information and Communication affect Market Outcomes

677 J.A. ATTEY, Causes and Macroeconomic Consequences of Time Variations in Wage Indexation

678 T. BOOT, Macroeconomic Forecasting under Regime Switching, Structural Breaks and High-dimensional Data

679 I. TIKOUDIS, Urban Second-best Road Pricing: Spatial General Equilibrium Perspectives

680 F.A. FELS, Empirical Studies of Consumer and Government Purchase Decisions

681 Y. GAO, Stability and Adaptivity: Preferences over Time and under Risk

682 M.J. ZAMOJSKI, Panta Rhei, Measurement and Discovery of Change in Financial Markets

683 P.R. DENDERSKI, Essays on Information and Heterogeneity in Macroeconomics

684 U. TURMUNKH, Ambiguity in Social Dilemmas

685 U. KESKIN, Essays on Decision Making: Intertemporal Choice and Uncertainty

686 M. LAMMERS, Financial Incentives and Job Choice

687 Z. ZHANG, Topics in Forecasting Macroeconomic Time Series

688 X. XIAO, Options and Higher Order Risk Premiums

689 D.C. SMERDON, Everybodys doing it: Essays on Trust, Norms and Integration

690 S. SINGH, Three Essays on the Insurance of Income Risk and Monetary Policy

691 E. SILDE, The Econometrics of Financial Comovement

692 G. DE OLIVEIRA, Coercion and Integration

693 S. CHAN, Wake Me up before you CoCo: Implications of Contingent Convertible Capital for Financial Regulation

694 P. GAL, Essays on the role of frictions for firms, sectors and the macroeconomy

695 Z. FAN, Essays on International Portfolio Choice and Asset Pricing under Financial
Contagion

696 H. ZHANG, Dealing with Health and Health Care System Challenges in China: Assessing Health Determinants and Health Care Reforms

697 M. VAN LENT, Essays on Intrinsic Motivation of Students and Workers

698 R.W. POLDERMANS, Accuracy of Method of Moments Based Inference

699 J.E. LUSTENHOUWER, Monetary and Fiscal Policy under Bounded Rationality and Heterogeneous Expectations

700 W. HUANG, Trading and Clearing in Modern Times

701 N. DE GROOT, Evaluating Labor Market Policy in the Netherlands

702 R.E.F. VAN MAURIK, The Economics of Pension Reforms

703 I. AYDOGAN, Decisions from Experience and from Description: Beliefs and Probability Weighting

704 T.B. CHILD, Political Economy of Development, Conflict, and Business Networks

705 O. HERLEM, Three Stories on Influence

706 J.D. ZHENG, Social Identity and Social Preferences: An Empirical Exploration