

Blanco, Mariana; Engelmann, Dirk; Koch, Alexander K.; Normann, Hans-Theo

**Working Paper**

## Preferences and beliefs in a sequential social dilemma: A within-subjects analysis

DICE Discussion Paper, No. 145

**Provided in Cooperation with:**

Düsseldorf Institute for Competition Economics (DICE), Heinrich Heine University Düsseldorf

*Suggested Citation:* Blanco, Mariana; Engelmann, Dirk; Koch, Alexander K.; Normann, Hans-Theo (2014) : Preferences and beliefs in a sequential social dilemma: A within-subjects analysis, DICE Discussion Paper, No. 145, ISBN 978-3-86304-144-1, Heinrich Heine University Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Düsseldorf

This Version is available at:

<https://hdl.handle.net/10419/96926>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

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

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# DISCUSSION PAPER

No 145

## Preferences and Beliefs in a Sequential Social Dilemma: A Within-Subjects Analysis

Mariana Blanco,  
Dirk Engelmann,  
Alexander K. Koch,  
Hans-Theo Normann

May 2014

## IMPRINT

### DICE DISCUSSION PAPER

Published by

düsseldorf university press (dup) on behalf of  
Heinrich-Heine-Universität Düsseldorf, Faculty of Economics,  
Düsseldorf Institute for Competition Economics (DICE), Universitätsstraße 1,  
40225 Düsseldorf, Germany  
[www.dice.hhu.de](http://www.dice.hhu.de)

### Editor:

Prof. Dr. Hans-Theo Normann  
Düsseldorf Institute for Competition Economics (DICE)  
Phone: +49(0) 211-81-15125, e-mail: [normann@dice.hhu.de](mailto:normann@dice.hhu.de)

### DICE DISCUSSION PAPER

All rights reserved. Düsseldorf, Germany, 2014

ISSN 2190-9938 (online) – ISBN 978-3-86304-144-1

The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors' own opinions and do not necessarily reflect those of the editor.

# Preferences and Beliefs in a Sequential Social Dilemma: A Within-Subjects Analysis\*

Mariana Blanco<sup>a</sup>, Dirk Engelmann<sup>b</sup>, Alexander K. Koch<sup>c</sup>, Hans-Theo Normann<sup>d</sup>

May 2014

## Abstract

In empirical analyses of games, preferences and beliefs are typically treated as independent. However, if beliefs and preferences interact, this may have implications for the interpretation of observed behavior. Our sequential social dilemma experiment allows us to separate different interaction channels. When subjects play both roles in such experiments, a positive correlation between first- and second-mover behavior is frequently reported. We find that the observed correlation primarily originates via an indirect channel, where second-mover decisions influence beliefs through a consensus effect, and the first-mover decision is a best response to these beliefs. Specifically, beliefs about second-mover cooperation are biased toward own second-mover behavior, and most subjects best respond to stated beliefs. However, we also find evidence for a direct, preference-based channel. When first movers know the true probability of second-mover cooperation, subjects' own second moves still have predictive power regarding their first moves.

JEL: C72; C90.

Keywords: Beliefs; consensus effect; social dilemma; experimental economics.

---

\*Financial support from the Nuffield Foundation, grant No. SGS/34070, is gratefully acknowledged. We thank Steffen Altmann, Maria Bigoni, Steve Burks, Ernst Fehr, Simon Gächter, Sebastian Kranz, Louis Levy-Garboua, Michael Naef, Daniele Nosenzo, Matthias Wibral as well as the anonymous advisory editor and referee for helpful comments. Contact: <sup>a</sup> Universidad del Rosario, Economics Department, Calle 14 No. 4-80, Bogotá, Colombia. E-mail: mariana.blanco@urosario.edu.co. <sup>b</sup> Department of Economics, University of Mannheim, L7, 3-5, D-68131 Mannheim, Germany; Centre for Experimental Economics, University of Copenhagen; CERGE-EI, Prague; CESifo, Munich. E-mail: dirk.engelmann@uni-mannheim.de. <sup>c</sup> School of Economics and Management, Aarhus University, Building 1322, 8000 Aarhus C, Denmark. Email: akoch@econ.au.dk. <sup>d</sup> Duesseldorf Institute for Competition Economics (DICE), Duesseldorf University, 40225 Duesseldorf, Germany; and Max-Planck Institute for Research on Collective Goods. E-mail: normann@dice.hhu.de.

# 1 Introduction

Behavioral economic theory offers a wide range of models that predict how actions in social dilemmas will vary for people with different types of (social) preferences and what an individual's best response is for a given set of beliefs.<sup>1</sup> While these models broaden the spectrum of preferences that people may hold, they typically stick to the standard assumption that people hold correct beliefs (in equilibrium).<sup>2</sup> The downside with this approach is to miss a crucial point: how likely a person thinks it is that others will defect in a social dilemma may well depend on her own attitude toward cooperation. As such an interaction of preferences and beliefs is of general importance for decision making in games, the topic appears to be strangely underdeveloped in the economic literature.

The significance of this issue is underlined by recent findings from sequential social dilemma experiments using a within-subjects design.<sup>3</sup> The data show that subjects who defect as first movers are more likely to exploit first-mover cooperation in their second-mover choice than those who cooperate as first movers. Blanco *et al.* (2011) document this for the sequential-move prisoners' dilemma.<sup>4</sup> Altmann *et al.* (2008) and Gächter *et al.* (2012) have a similar result for the trust game and for a sequential voluntary contribution game, respectively.

The observed within-subjects correlation of the first and the second move is provocative in several ways. First, as noted by Altmann *et al.* (2008) and Blanco *et al.* (2011), the finding is at odds with prominent social preference models that are frequently invoked for explaining behavior in social dilemma games. Both *inequality aversion* (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) and *reciprocal preferences* (Dufwenberg and Kirchsteiger 2004) – under standard assumptions, including that beliefs are not correlated with preferences – would predict a negative correlation of first- and second-mover choices, and not the positive correlation observed. While altruism could rationalize

---

<sup>1</sup>Frequently, beliefs are considered to be distributions that merely rationalize revealed preference orders. We follow here the interpretation that beliefs are real, meaning that they are an independent part of decision making, implying that they can be elicited in experiments. See also Costa-Gomes, Huck, and Weizsäcker (2010).

<sup>2</sup>Osborne (2009, p. 379) presents this as the standard approach. Some approaches within behavioral economics relax the assumption of correct beliefs. For example, the level- $k$  literature is explicitly based on assuming very different (non-equilibrium) beliefs. These models have, however, typically not been applied to explaining behavior in social dilemmas.

<sup>3</sup>Earlier experimental analyses of sequential social dilemmas include the sequential-move prisoners' dilemma (Bolle and Ockenfels 1990, Clark and Sefton 2001), the gift-exchange game (Fehr *et al.* 1993), the trust or investment game (Berg *et al.* 1995), the lost wallet game (Dufwenberg and Gneezy 2000), and public-good games with a front runner (Potters *et al.* 2007).

<sup>4</sup>Blanco *et al.* (2011) check for the within-subjects correlation of six different moves in four different games. The correlation of the first and the second move (given first-mover cooperation) in the sequential-move prisoners' dilemma was the strongest among all 15 correlations.

the correlation pattern it would also predict unconditional cooperation, which, however, is at odds with the data (see Bolle and Ockenfels 1990, Clark and Sefton 2001, and Blanco *et al.* 2011).

Second, for simultaneous-move prisoners' dilemma experiments, it has been argued that "fear" and "greed" are main driving forces of behavior (Ahn *et al.* 2001, Simpson 2003). Fear refers to the risk of being exploited by the other player when cooperating. Greed describes a player's willingness to defect if the other player cooperates. The sequential-move prisoners' dilemma separates the two motives: fear applies to the first move and greed to the second move. Thus, the correlation of first and second moves suggests that fear and greed are correlated at the individual level. But it does not seem evident why fearful people should be more greedy.

Third and more fundamentally, following standard game-theoretic arguments, first-mover choices should follow a "best respond to your beliefs" principle<sup>5</sup>, and hence reflect the natural variation in beliefs across subjects in an experiment. Second-mover choices, in contrast, are simple decision problems and should depend on players' preferences only. Thus, one would not expect the choices of a person in the role of first and second mover to be strongly related to each other – unless beliefs and preferences are correlated.

A correlation between preferences and beliefs may, however, be exactly what drives the correlation between first-mover and second-mover decisions. The so-called *consensus effect*, according to which players' beliefs are biased toward their own type, would suggest that those subjects who cooperate as second movers will expect a higher second-mover cooperation rate among others than those subjects who defect as second movers (Mullen *et al.* 1985, Engelmann and Strobel 2000). Second-move cooperators hence will perceive a higher expected payoff from cooperating as first mover than second-move defectors. So, all else equal (that is, if there is no relationship between preferences for cooperation in the role of first and second mover), second-mover cooperators should be more likely than second-mover defectors to cooperate as first mover.

Another response to the above issues raised by the experimental data is to turn to alternative social preference models that are consistent with the observed correlation of choices without assuming systematic differences in beliefs across players. A combination of efficiency concerns with maximin preferences (Charness and Rabin 2002) and reciprocal altruism (Levine 1998, Cox, Friedman, and Sadiraj 2008) are among the alternatives that can explain why first-mover decisions differ between second-mover cooperators and second-mover defectors, even if they hold the same beliefs.

The aforementioned theories presume a *direct, preference-based channel* that influences both

---

<sup>5</sup>For recent experiments investigating this issue see, for example, Dhaene and Bouckaert (2010), Costa-Gomes and Weizsäcker (2008), Rey-Biel (2009), and Koch *et al.* (2009). On the fundamental question whether beliefs are causal for behavior see Costa-Gomes *et al.* (2010). In a trust game they exogenously shift the trustee's repayment and use this shift to instrument the trustor's beliefs. Their results provide evidence of a causal effect of beliefs on actions.

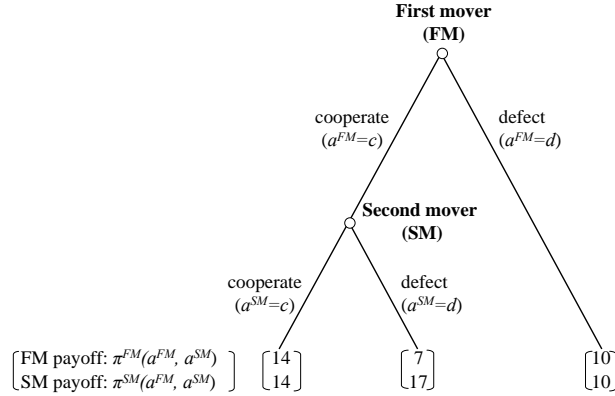


Figure 1: Sequential-move prisoner's dilemma game

first- and second-mover behavior. The consensus effect, in contrast, suggests an *indirect channel* that links preferences (as reflected in a person's second-mover decisions) to the first-mover decision via beliefs. But what is the right approach?

The issue of indirect versus direct channel seems particularly relevant because the consensus effect has emerged already in other settings as a plausible alternative to preference-based explanations in rationalizing certain patterns of behavior. For instance, dictator- and trust-game studies where participants report what they believe their counterpart expects in the game, show significant correlations between these second-order beliefs and actions. An explanation for this pattern is that some people are guilt averse. That is, they experience a utility loss if they believe to let someone down (Charness and Dufwenberg 2006). But Ellingsen *et al.* (2010) conclude from their own experiments that the correlation can almost exclusively be attributed to a consensus effect. When subjects are informed about their counterpart's first-order belief, this belief has almost zero correlation with own behavior. Such a correlation would, however, be required for the guilt-aversion (that is, preference-based) explanation.

We analyze these issues with experimental data from a one-shot sequential-move prisoners' dilemma summarized in Figure 1, referred to as SPD henceforth. Subjects play both the first and the second mover role. If the first mover (FM) cooperates ( $a^{FM} = c$ ), payoffs depend on the action of the second mover (SM). If  $a^{SM} = c$ , payoffs are 14 for both first and second mover; if  $a^{SM} = d$ , the payoff is 7 for the first mover and 17 for second mover. When  $a^{FM} = d$ , the game ends with a payoff of 10 for both first mover and second mover. Unconditional cooperation is precluded by design (this is motivated by the near absence of unconditional cooperation in sequential-move prisoner's dilemma experiments; see Bolle and Ockenfels 1990, Clark and Sefton 2001, and Blanco *et al.* 2011).

The unique subgame-perfect Nash equilibrium of the game in Figure 1 (for rational, selfish players) is  $a^{SM} = d$ ,  $a^{FM} = d$ . The second mover would always defect. Thus, in the first-mover role, if the player knows that the second mover is rational and selfish, she will defect as well. Given the possibility of second-mover cooperation, a selfish first mover will choose  $a^{SM} = c$  if and only if the belief about the frequency of second-mover cooperation is at least  $3/7 \approx 0.43$ .

The following stylized model illustrates the logic of our experimental design, capturing the essential differences in how standard social preference models explain behavior in our sequential prisoners dilemma.<sup>6</sup> Suppose players maximize a von Neumann-Morgenstern utility function with the following components: first, a linear utility from monetary payoffs which depends on player  $i$ 's belief about player  $j$ 's probability of cooperating as second mover ( $b_i^j$ ), and second, the expected utility  $f_i(b_i^j)$  that player  $i$  obtains from cooperating in the role of first mover. This second component is understood to result directly from the act of cooperating, and not from any expected return that cooperation might yield, which the monetary payoff component captures. Note that  $f_i(b_i^j)$  depends on player  $i$ 's belief about second-mover cooperation. We assume that  $f'_i(\cdot) \geq 0$  to capture that the psychological utility one gets from cooperating as first mover may depend on one's expectation that the second mover will reciprocate. The third component  $s_i$  is the utility that player  $i$  gets from cooperating as second mover. Specifically, the first mover chooses between defection, which gives a utility of 10, and cooperation, which yields the following utility:

$$U_i(a^{FM} = c) = 14b_i^j + 7(1 - b_i^j) + f_i(b_i^j) = 7 + 7b_i^j + f_i(b_i^j) \quad (1)$$

The second mover chooses between defection, which gives a utility of 17, and cooperation, which yields:

$$U_j(a^{SM} = c) = 14 + s_j \quad (2)$$

Consider now an experiment where participants play both roles and may be paid for either of them. Inspection of the stylized model shows that second-mover cooperation becomes more attractive the larger  $s_i$ , the second-mover payoff being  $17 + \mathbb{1}_{\{a^{SM}=c\}}(-3 + s_i)$ , where  $\mathbb{1}$  is an indicator function. And first-mover cooperation becomes more attractive the larger  $f_i(b_i^j)$  or  $b_i^j$ , the first-mover payoff being  $10 + \mathbb{1}_{\{a^{FM}=c\}}(-3 + 7b_i^j + f_i(b_i^j))$ . This shows that a possible correlation between first- and second-mover behavior can arise through a direct preference-based channel, exhibited through a positive correlation of  $f_i(b_i^j)$  and  $s_i$  (for any given  $b_i^j$ ), or through an indirect belief-based channel, exhibited through a positive correlation of  $b_i^j$  and  $s_i$ . Our experiment is designed to discriminate between these two channels.

---

<sup>6</sup>We address in detail in Section 3 how the predictions of various social preference models translate into key correlations in this stylized model.



Next to a standard sequential dilemma, we have a treatment where we elicit beliefs about second mover behavior and, in a third treatment, we give as feedback the actual frequency of second-mover cooperators before subjects decide their first move. This treatment switches off the indirect channel: as we provide subjects with the true number of second-mover cooperators, it should eliminate the correlation between the first- and second-mover choices.

Our main findings are that the observed correlation primarily originates via an indirect channel, where second-mover decisions influence beliefs through a consensus effect, and the first-mover decision is a best response to these beliefs. However, we also find evidence for a more conventional direct, preference-based channel. When first movers know the true probability of second-mover cooperation, subjects’ own second moves still have predictive power regarding their first moves.

The rest of the paper is organized as follows. Design and procedures are discussed in Section 2. A review of behavioral models both inconsistent and consistent with the correlation of moves observed in sequential dilemmas follows in Section 3. Section 4 presents the results and Section 5 concludes.

## 2 Experimental Design and Procedures

### 2.1 Design

Participants play the game in Figure 1 once.<sup>7</sup> All subjects play the game in both the first- and the second-mover role. They first decide as the second mover and then as the first mover. As will become clear below, our design requires this very order of decisions. After participants have made their decisions, they are randomly assigned roles and are randomly matched into pairs, and payoffs are calculated according to the relevant decisions.

The experiments use a neutral frame. We relabelled players and actions as follows: FM=*A player*, SM=*B player*, FM cooperate=*IN*, FM defect=*OUT*, SM cooperate=*LEFT*, SM defect=*RIGHT*. Payoff units were called experimental currency units (ECU).

Table 1 summarizes our treatments. In *Baseline* we neither elicit beliefs about second-mover cooperation nor do we give feedback on the true frequency of second-mover cooperation. In *Elicit\_Beliefs*, participants have to guess how many of the nine other participants in the session cooperate as second movers. This “guess task” is performed between second- and first-mover decisions, and is incentivized. In *True\_Distribution*, before subjects decide in the role of the first mover, they are informed about the actual number of second-mover cooperators among the nine other participants in the session. In order to keep the number and nature of decisions as similar as possible across treatments, we introduced a belief-elicitation task also in *Baseline* and in *True\_Distribution*:

---

<sup>7</sup>The main focus of our experiment is the impact of beliefs on choices in the SPD. With repeated play, beliefs become confounded with experience. In order to keep this apart, our experiment is one-shot.

	<i>Baseline</i>	<i>Elicit_Beliefs</i>	<i>True_Distribution</i>
Task 1	2 <sup>nd</sup> move	2 <sup>nd</sup> move	2 <sup>nd</sup> move
Feedback ( $a_{-i}^{SM}$ )	no	no	yes
Task 2	1 <sup>st</sup> move	beliefs ( $a_{-i}^{SM}$ )	1 <sup>st</sup> move
Task 3	beliefs ( $a_{-i}^{FM}$ )	1 <sup>st</sup> move	beliefs ( $a_{-i}^{FM}$ )
# Participants	40	60	60

Table 1: Treatments.

participants make a guess about the other participants’ *first* move. As beliefs about first-mover choices are not relevant to our research question, we will not analyze these guesses in detail.

The logic of the experimental design in terms of the stylized model is as follows. Treatment *Elicit\_Beliefs* is suitable to identify the indirect link through a potential consensus effect. If players’ beliefs are biased towards their own second-mover choice (that is, if we observe a correlation of  $s_i$  and  $b_i^j$ ), the data would confirm the consensus effect. Treatment *True\_Distribution* disables the indirect channel as it should eliminate the correlation between  $s_i$  and  $b_i^j$ . Hence, it allows us to test whether there is any correlation between  $f_i(b_i^j)$  and  $s_i$  (for any given  $b_i^j$ ) that would be indicative of the direct channel.<sup>8</sup>

For the belief-elicitation task (“guess task”), we use a quadratic scoring rule (Bhattacharya and Pfleiderer 1985, Huck and Weizsäcker 2002). Specifically, we ask subjects how many of the nine other participants in the lab cooperate in the role of second mover, and reward the accuracy of this stated belief using the quadratic scoring rule

$$\text{belief-elicitation task payoff} = 15 \times \left[ 1 - \left( \frac{d_i}{9} \right)^2 \right], \quad (3)$$

where  $d_i$  is the difference between player  $i$ ’s guess and the correct number of second-mover cooperators in the session. An accurate guess of how many of the other nine participants in the session chose to cooperate yields a payoff of 15. Rather than using the above formula, the reward for the accuracy of the guess is presented to the participants in a table (see the instructions in Appendix B).

Three specific design issues deserve further comment. First, pilot sessions of the *True\_Distribution* treatment suggested that strong emphasis of the relevance of the feedback about the other players’

---

<sup>8</sup>Related work by Costa-Gomes *et al.* 2010 investigates the second half of the indirect channel. Specifically, in order to assess the causality of beliefs for first-mover trust, they exogenously change the share that the second mover returns in a trust game by adding a randomly drawn number that is made known to the first but not the second mover.

second-mover choices is warranted. While written instructions in the final design are identical to those in the pilot sessions, the oral summary emphasizes the meaning of the feedback.<sup>9</sup>

Second, in *True\_Distribution*, the design theoretically creates an additional incentive for second-mover cooperation compared to the other treatments. Namely, as participants are informed about the number of second-mover cooperators, cooperation as a second mover could in principle increase the first-mover cooperation rate. If subjects reasoned this way there would be a higher second-mover cooperation rate than in *Baseline* and *Elicit\_Beliefs*, which our data, however, clearly reject (see Section 4.1). (Note that the additional strong (oral) instructions were given only *after* subjects made their second-mover decisions.)

Third, the quadratic scoring rule is incentive compatible for risk-neutral individuals only. A concern could be that subjects in *Elicit\_Beliefs* attempt to reduce payoff risk by reporting guesses closer to 4 or 5, even if these do not correspond to their beliefs. Such behavior would reduce variation in stated beliefs relative to true beliefs, and thus would make it harder to find evidence for any underlying correlations. In our data we, however, find strongly significant correlations of stated beliefs and decisions (see Section 4.3).

## 2.2 Procedures

The experiments were carried out computer based with the experimental software z-Tree (Fischbacher 2007) in the Experimental Laboratory of Royal Holloway, University of London. Participants were students from various disciplines.

We conducted 16 sessions with ten participants each (that is, 160 participants in total). Because the experiment is one-shot, each participant provides an independent observation. There were four sessions for *Baseline*, and six sessions each for *Elicit\_Beliefs* and *True\_Distribution* (see Table 1).

The payment to subjects is usually *either* the payoff from playing the SPD game *or* the payoff from the belief-elicitation task, with the exception of three of the six *Elicit\_Beliefs* sessions where both tasks were paid.<sup>10</sup> A random computer draw decides which of the two tasks are paid, both

---

<sup>9</sup>In the pilot sessions, subjects were told that after they made their second-mover decision “[a]ll participants in the room did the above Decision Task B. Now you will be informed about how many of the nine other participants in the room chose LEFT in Decision Task B.” This summary seemed too brief, as feedback did not have a significant impact on first-mover cooperation rates. In the final design, the oral summary therefore includes additionally: “Note that if you are assigned the role of Person A, one of these nine choices is the choice of the person you will be matched with. This means, for example, that if the information is that nine out of nine chose LEFT, then you know for sure that you will be matched with a person B who chose LEFT ... ” The oral summary continued with other examples (see the Appendix for instructions and oral summaries).

<sup>10</sup>The purpose of varying the payment method was to test whether subjects respond to the presence of hedging opportunities. See Blanco *et al.* (2010) for details. The method of payment causes no significant differences. Indeed,

being equally likely. This procedure removes potential hedging opportunities. To make the possible payoffs from each task approximately equal, we set the scoring factor for the belief-elicitation task to 15 in (3). The final payout in experimental currency units (ECU) was converted into Pounds Sterling at an exchange rate of £ 1 per ECU (£ 0.5 per ECU in the three *Elicit\_Beliefs* sessions where both tasks were paid to keep incentives and average earnings similar across sessions).

In the beginning of each session, participants read through the instructions (see Appendix), followed by a control questionnaire that required them to solve simple examples on how actions determine payoffs. Any questions were answered privately. Prior to each task there was an oral summary. When all participants had finished the control questionnaire, an oral summary for the first task was given; when all had finished the first task, the next task was summarized, and so on.

Participants were informed that, after all tasks were completed, they would be randomly paired with a participant in the room and would randomly be assigned a role (first or second mover). They also knew that, at the moment of making their decisions they would not know their own role or their co-player's decision.

### 3 Theoretical Background

In this section, we review behavioral models both inconsistent and consistent with the correlation of moves observed in sequential social dilemmas. Inevitably, these are non-standard models where players have non-selfish preferences because for standard (selfish) players the prediction is ( $a^{SM} = d$ ,  $a^{FM} = d$ ) throughout.<sup>11</sup>

**Proposition 1** (*Non-selfish preferences inconsistent with positive correlation observed*) *Models with inequality averse players (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) or reciprocal players (Dufwenberg and Kirchsteiger 2004) predict a negative correlation of moves, not the positive correlation observed.*

Consider inequality averse players (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) first. According to Fehr and Schmidt (1999) and in a two-player game, player  $i$  is assumed to maximize

$$U_i(x) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$$

---

the results are virtually identical, so we pool the data from the six *Elicit\_Beliefs* sessions in the analysis below.

<sup>11</sup>Variations in *risk preferences* might also contribute to an explanation of the correlation between first- and second-mover decisions. As second-mover decisions involve no risk, this would require risk tolerance to be positively related to second-mover cooperation. Burks *et al.* (2009) indeed find an indirect relation between these two variables. But in our setting, for typical degrees of risk aversion, risk preferences can only explain variation in first-mover behavior for subjects with a belief that four out of nine second movers cooperate.

where  $x_i$  is the payoff of player  $i$ ,  $x_j$  is the payoff of the other player,  $\alpha_i \geq \beta_i \geq 0$  and  $\beta_i < 1$ .  $\alpha_i$  and  $\beta_i$  capture aversion toward disadvantageous inequality and advantageous inequality, respectively. Second movers who dislike advantageous inequality are *more* inclined to cooperate. That is, players with a larger  $\beta_i$  behave as if they have a larger  $s_i$ . First movers with a larger  $\alpha_i$  are *less* inclined to cooperate than selfish players and behave as if they have a smaller  $f_i(b_i^j)$  for any given  $b_i^j < 1$ . Thus, the postulated positive correlation of  $\alpha_i$  and  $\beta_i$  (Fehr and Schmidt, 1999, p. 864) predicts, for given beliefs  $b_i^j$ , a *negative* correlation of first- and second-mover cooperation (see Appendix A for details).<sup>12</sup> Bolton and Ockenfels (2000, p. 182-3) explicitly prove Proposition 1 for their model so we refrain from discussing this model in detail.

As for *reciprocal preferences*, consider a two-player version of the Dufwenberg and Kirchsteiger (2004) model and drop various belief variables for simplicity. Player  $i$  is assumed to maximize

$$U_i = \pi_i + Y_i k_{ij} \lambda_{iji},$$

where  $\pi_i$  is player  $i$ 's material payoff,  $Y_i$  is  $i$ 's sensitivity to reciprocity concerns toward player  $j$ ,  $k_{ij}$  is  $i$ 's kindness toward  $j$  and  $\lambda_{iji}$  is how kind  $i$  believes  $j$  to be toward  $i$ . Kindness of  $i$  to  $j$  is measured by the deviation of  $j$ 's resulting payoff from the equitable payoff, which is in turn the average of  $j$ 's maximum possible payoff and  $j$ 's minimum possible efficient payoff, and correspondingly for  $j$ 's kindness toward  $i$ . With the payoffs of our SPD, the Dufwenberg and Kirchsteiger model implies that, if there is any correlation between first and second moves at all, it should be negative. The reason is that first-mover behavior coincides with that of a selfish player regardless of the degree of reciprocity and for all beliefs, except for the belief that four out of nine second movers cooperate. In this case, a selfish first mover will cooperate, whereas a sufficiently reciprocal player will defect. Hence a reciprocal player is more likely to defect as first mover, but more likely to cooperate as second mover than a selfish player. See the Appendix A for a proof.

We now turn to non-selfish preferences and belief-based explanations consistent with observed behavior:

**Proposition 2** (*Explanations consistent with positive correlation observed*) *Models allowing for efficiency concerns combined with maximin preferences (Charness and Rabin 2002), reciprocal altruism (Cox, Friedman, and Sadiraj 2008, Levine 1998), or a consensus effect (Mullen et al. 1985, Engelmann and Strobel 2000) predict the positive correlation of moves observed.*

Conditional cooperation as well as the positive correlation between first- and second-mover behavior can result if efficiency (i.e., total payoff maximization) concerns are combined with maximin

---

<sup>12</sup>Based on the evidence in Blanco *et al.* (2011), who find no significant correlation in their within-subjects estimates of Fehr-Schmidt parameters, the model would predict no correlation of first- and second-mover choices.

preferences. In Charness and Rabin (2002), the utility function of player  $i$  in a two-player game is given by (we use the formulation of their appendix):

$$V_i(\pi_1, \pi_2) = (1 - \lambda_i)\pi_i + \lambda_i[\delta_i \min\{\pi_1, \pi_2\} + (1 - \delta_i)(\pi_1 + \pi_2)].$$

The total payoff,  $\pi_1 + \pi_2$ , increases with both first- and second-mover cooperation. Note that the minimum payoff,  $\min\{\pi_1, \pi_2\}$ , increases with second-mover cooperation after first-mover cooperation, and increases in expectation with first-mover cooperation as long as  $b_i^j > 4/9$ . Thus, a player with  $\delta_i = 1$  behaves as if  $f_i(b_i^j)$  and  $s_i$  are positively correlated for  $b_i^j > 4/9$ , but not otherwise. For smaller  $\delta_i$  the range of beliefs where  $f_i(b_i^j)$  and  $s_i$  are correlated increases.<sup>13</sup>

Cox, Friedman, and Sadiraj (2008) can capture the correlation between first- and second-mover choices through heterogeneity in the degree of altruism, and conditional second-mover cooperation through reciprocity. Altruism in their model is captured via the willingness to pay for the other player's payoff in terms of own payoff. Under the plausible assumption that players who are more altruistic than others in one situation will also be so in other situations, those more altruistic will behave as if both their  $f_i(b_i^j)$  (for any given  $b_i^j$ ) and  $s_i$  are larger. Thus this suggests a positive correlation between  $f_i(b_i^j)$  and  $s$ . The reciprocity axiom in the model by Cox, Friedman, and Sadiraj (2008) says that, if the first mover chooses an opportunity set  $G$  for the second mover that is more generous than opportunity set  $F$ , then the second mover will be more altruistic toward the first mover (her willingness to pay to increase the first mover's payoff is larger at any given allocation). This reciprocity axiom, and the convexity of preferences in both own and other's payoff, imply that the model is consistent with second movers being more likely to cooperate after first-mover cooperation than after first-mover defection. See the Appendix A for details.

In Levine's (1998) model, own *altruism* interacts with a player's estimate of the other's altruism. Specifically, the *adjusted* utility of player  $i$  is assumed to be

$$v_i = u_i + \sum_{j \neq i} \frac{a_i + \lambda a_j}{1 + \lambda} u_j,$$

with  $u_i, u_j$  being the direct utility of players  $i$  and  $j$  from the game (which we can equate with the monetary payoffs in our experiment),  $-1 < a_i < 1$  and  $-1 < a_j < 1$  player  $i$ 's and  $j$ 's degree of spite or altruism, respectively, and  $0 \leq \lambda \leq 1$  reflecting how much player  $i$  cares about how altruistic  $j$  is.

---

<sup>13</sup> Even stronger support for conditional cooperation comes from the more elaborate version of Charness and Rabin's model that includes concern withdrawal – that is, a reduced weight in the utility function on the payoffs of players who “misbehave”. It implies that a first mover who has defected obtains a smaller, possibly negative weight in the utility function. Efficiency concerns without a specific regard for the poorest player (Engelmann and Strobel 2004)—with the notation of Charness and Rabin (2002):  $\lambda_i$  large,  $\delta_i$  small— also predict a positive correlation but would also predict cooperation after first-mover defection. This, however, is very rarely observed. For the same reason, unconditional altruism is an explanation consistent with the positive correlation observed but unlikely to have explanatory power.

Player  $i$  is selfish if  $a_i = 0$ ;  $\lambda = 0$  would correspond to unconditional altruism. Given beliefs  $b_i^j$ , an altruistic first mover is more likely to cooperate than a selfish one, because the larger  $a_i$ , the higher  $i$ 's willingness to pay in terms of  $u_i$  for increasing  $u_j$ . For the same reason, the larger  $a_i$ , the more likely that a player  $i$  cooperates as a second mover. In terms of our stylized model, a first mover with larger  $a_i$  behaves as if  $f_i(b_i^j)$  is larger for given  $b_i^j$  and a second mover with larger  $a_i$  would behave as if  $s_i$  is larger. Levine's model therefore directly implies a positive correlation between  $f_i(b_i^j)$  and  $s_i$ . Furthermore, if  $\lambda_i > 0$ , then a higher estimate of the other player's altruism  $a_j$  yields higher utility from cooperation. This implies that conditional cooperation is more likely than unconditional cooperation, because first-mover defection signals low altruism of the first mover.<sup>14</sup>

As discussed in Section 1, the *consensus effect* (Mullen *et al.* 1985, Engelmann and Strobel 2000) offers a plausible alternative explanation for the positive correlation of first- and second-mover choices. A consensus effect is said to occur when players hold a belief that is biased toward their own preference or choice. If players' beliefs about second-mover behavior are subject to a consensus effect and if their first-mover choices are best responses to their beliefs, this means that they are more likely to cooperate as first movers provided they cooperate as second movers.<sup>15</sup> In terms of our stylized model, the argument based on a consensus effect thus suggests that players are heterogeneous in their  $s_i$  and those with higher  $s_i$  expect others to have a higher  $s_i$ , which implies a higher  $b_i^j$ . This then generates a correlation between first- and second-mover cooperation even when  $s_i$  and  $f_i(b_i^j)$  are uncorrelated for given  $b_i^j$  (including if  $f_i(b_i^j) = 0$  for all  $b_i^j$  so that first movers cooperate only if they expect this to be money maximizing).

## 4 Results

### 4.1 Overview

Overall, 49 percent of the first movers and 54 percent of the second movers cooperate. As Table 2 shows, second-mover cooperation rates are virtually identical across treatments, and pairwise comparisons of second-mover cooperation yield no significant differences either (all two-sided Fisher exact tests yield  $p = 0.999$ ). This indicates that second-mover cooperation in *True\_Distribution* is not increased by strategic considerations as discussed in Section 2.1. First-mover cooperation rates

---

<sup>14</sup>Kranz's (2010) model of *rule consequentialism* also combines concerns for own payoff and efficiency. Some players, so-called compliers, are assumed to care about complying with a moral norm that maximizes social welfare; the other players are selfish payoff maximizers. A norm that prescribes first movers to cooperate and second movers to conditionally cooperate would then in terms of our model amount to  $f_i(b_i^j)$  and  $s_i$  being perfectly correlated.

<sup>15</sup>Obviously, a consensus effect does not explain why some second movers cooperate in the first place. Thus even if the correlation between first- and second-mover choices is best explained by a consensus effect, a complete explanation of the data will require some preference element that rationalizes second-mover cooperation.

	<i>Baseline</i>	<i>Elicit_Beliefs</i>	<i>True_Distribution</i>	Total
first mover ( <i>FM</i> )	27.5%	55.0%	56.7%	48.8%
second mover ( <i>SM</i> )	55.0%	53.3%	55.0%	54.4%

Table 2: Average cooperation rates by treatment.

$a^{FM}$	$a^{SM}$	<i>Baseline</i>	<i>Elicit_Beliefs</i>	<i>True_Distribution</i>	Total
<i>Same choice as first and second mover <math>a^{FM} = a^{SM}</math></i>					
<i>c</i>	<i>c</i>	10 (25.0%)	27 (45.0%)	23 (38.3%)	60 (37.5%)
<i>d</i>	<i>d</i>	17 (42.5%)	22 (36.7%)	16 (26.7%)	55 (34.4%)
Sum		27 (67.5%)	49 (81.7%)	39 (65.0%)	115 (71.9%)
<i>Different choices as first and second mover <math>a^{FM} \neq a^{SM}</math></i>					
<i>c</i>	<i>d</i>	1 (2.5%)	6 (10.0%)	11 (18.3%)	18 (11.3%)
<i>d</i>	<i>c</i>	12 (30.0%)	5 (8.3%)	10 (16.7%)	27 (16.9%)
Sum		13 (32.5%)	11 (18.3%)	21 (35.0%)	45 (28.1%)
Total		40 (100%)	60 (100%)	60 (100%)	160 (100%)

Table 3: Distribution of individual choice pairs by treatment.

are similar, too, with the exception of *Baseline* where fewer subjects cooperate as first movers, and we reject the hypothesis that all three cooperation rates are the same (two-sided Fisher exact test,  $p = 0.008$ ).<sup>16</sup> We will return to this treatment effect in Section 4.3. For now, we remark that it is not the overall cooperation rates that matter for our research question, but the correlation of first- and second-mover decisions.

At the treatment level, first-mover cooperation is a risk-neutral best response, because the second-mover cooperation rate exceeds the threshold of  $3/7 \approx 43$  percent in all treatments. This does not hold in all individual sessions though, and we examine below the individual subjects' best responses.

Crucially for our research question, we find that most subjects make the same choice as first and as second movers, similar to the results in Blanco *et al.* (2011). Table 3 shows that across all treatments, of the 160 subjects, 60 (38%) cooperate in both roles and 55 (34%) defect in both roles. Only 27 subjects (17%) defect as first movers and cooperate as second movers, while the remaining 18 subjects (11%) cooperate as first movers and defect as second movers. We will discuss

<sup>16</sup>We find significant pairwise differences in first-mover cooperation for *Baseline* vs. *Elicit\_Beliefs* (two-sided Fisher exact test,  $p = 0.008$ , Boschloo (1970) test,  $p = 0.007$ ) and *Baseline* vs. *True\_Distribution* ( $p = 0.005$ ;  $p = 0.004$ ); but no significant difference for *Elicit\_Beliefs* vs. *True\_Distribution* ( $p = 0.999$ ;  $p = 0.999$ ).



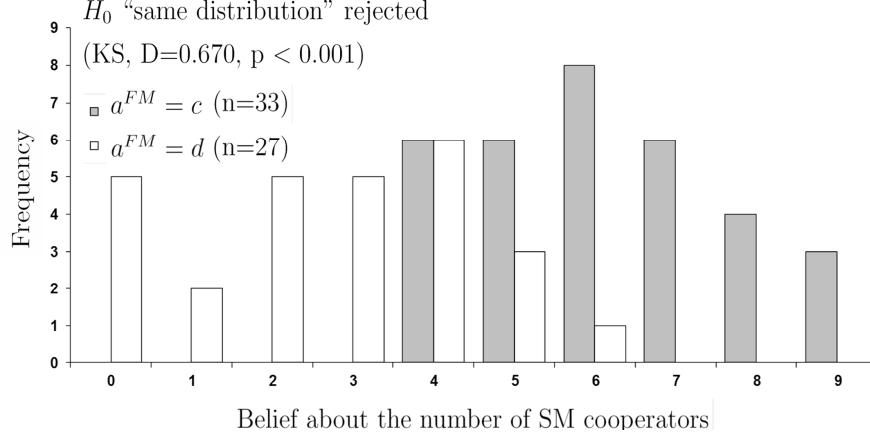


Figure 2: Histogram of stated beliefs (treatment *Elicit\_Beliefs*)

and interpret this finding below, nevertheless, we emphasize at this point that overall 72 percent of our subjects make the same decision in the two situations.

## 4.2 The *Baseline* treatment

Our *Baseline* treatment is the starting point of the analysis and establishes the aforementioned correlation of first- and second-mover choices. We find a significant phi correlation coefficient of  $\phi = 0.388$  ( $\chi^2 = 6.030$ ,  $d.f. = 1$ ,  $p = 0.014$ ). To sum up our findings on *Baseline*:

**Result 1** *In Baseline, the first and second move are positively correlated.*

## 4.3 The *Elicit\_Beliefs* treatment

In *Elicit\_Beliefs*, subjects have to guess how many of the other nine participants are cooperators, before making their first-mover choice. Eliciting beliefs is not necessarily innocuous as it may affect behavior (for example, Croson 2000).<sup>17</sup> Furthermore, the consensus effect may drive subjects to form beliefs that are correlated with their second-mover decision.

Regarding the correlation of behavior in *Elicit\_Beliefs*, 49 of 60 (81.7%) subjects make the same choice as first and second movers. The correlation of choices is significant ( $\phi = 0.598$ ,  $\chi^2 = 21.431$ ,  $d.f. = 1$ ,  $p < 0.001$ ) and stronger than in *Baseline*.

<sup>17</sup>As we find a significant increase in first-mover cooperation relative to *Baseline*, superficially, this looks like contradicting Croson (2000), where cooperation decreases. But in her experiments it is a dominant strategy not to cooperate, whereas in our setting first-mover cooperation is a best response, given the average second-mover cooperation rates in *Baseline* and *Elicit\_Beliefs*. In general, the evidence on the effects of incentivized belief elicitation is mixed. For example, in a public goods game Gächter and Renner (2010) find an increase in contribution rates – in contrast to Croson (2000) – and in a trust game Guerra and Zizzo (2004) find no effect on trust and trustworthiness.

What are the stated beliefs like then? Figure 2 shows the histogram of stated beliefs based on the  $a^{FM}$  choice. It reveals a clear and strong finding: the two belief distributions are significantly different (two-sample Kolmogorov-Smirnov test,  $D = 0.670$ ,  $p < 0.001$ ). Subjects who choose  $a^{FM} = d$  are much more pessimistic about the number of second-mover cooperators (mean belief 2.7) than those who choose  $a^{FM} = c$  (mean belief 6.2). Indeed, the first-mover action and the stated belief are strongly correlated (rank biserial correlation  $r_{rb} = 0.864$ ,  $t = 13.074$ ,  $p < 0.001$ ).

Considering  $a^{SM}$  choices and beliefs, we find a similar correlation ( $r_{rb} = 0.808$ ,  $t = 10.447$ ,  $p < 0.001$ ), and the difference between the distributions of beliefs of  $a^{SM} = c$  and  $a^{SM} = d$  players again is significant ( $D = 0.768$ ,  $p < 0.001$ ). This is consistent with a consensus effect.

Figure 2 also reveals that almost all first movers choose the (risk neutral) selfish best response to their stated belief. Only for ten of 60 subjects, the stated belief is inconsistent with selfish risk-neutral payoff maximization. These subjects believe they are in a session with between four and six cooperators and choose  $a^{FM} = d$  even though they should cooperate.<sup>18</sup> This observation is consistent with a special case of the indirect channel explanation ( $s_i$  and  $f_i(b_i^j)$  are uncorrelated), where  $f_i(b_i^j)$  is (nearly) equal to 0 for most players.<sup>19</sup>

Probit regressions of the  $a^{FM}$  decisions can further add to this point. Using stated beliefs as explanatory variable, specification (E1) in Table 4 shows that the marginal effect of the variable *belief* is 0.35 and significant (at  $p < 0.001$ ) in *Elicit\_Beliefs*. The top panel of Figure 3 illustrates the result: the smooth black line is derived from specification (E1) and superimposed over the actual frequency of  $a^{FM} = c$  choices for a given stated belief about the number of  $a^{SM} = c$  players in the session. The sharp increase in first-mover cooperation rates for a belief of four or larger is consistent with selfish expected utility maximization.

Finally, how accurate are stated beliefs? Only seven (12%) of the subjects actually scored a perfect guess (that is, their belief was equal to the correct number of  $a^{SM} = c$  players in their session). Indeed, the large variation in beliefs in Figure 2 is not just noise; to a large part it arises because beliefs are biased toward subjects' own  $a^{SM}$  choices. To sum up our findings on

---

<sup>18</sup>A moderate amount of risk aversion can explain the majority of the deviations from best response. Six of the subjects state a belief of 4/9. For this belief  $a^{FM} = d$  is a best response with CRRA-utility in the empirically relevant range for the risk aversion coefficient of 0.3 to 0.5 (Holt and Laury 2002). As for this belief expected payoffs for  $a^{FM} = c$  exceed those for  $a^{FM} = d$  by only about 1 percent, small decision errors are an alternative explanation. Alternatively, these observations could be explained by these first movers having  $f_i(b_i^j) < 0$ .

<sup>19</sup>Remember that the consensus effect provides an explanation for a correlation between first- and second-mover cooperation if  $f_i(b_i^j)$  and  $s_i$  are uncorrelated, including if  $f_i(b_i^j) = 0$  for all  $b_i^j$ . That most first movers play a selfish best response to their stated beliefs is thus in line with the consensus effect explanation. Alternatively, additional variation in first-mover behavior that is uncorrelated to second-mover behavior could result if  $f_i(b_i^j) \neq 0$  but is not correlated to  $s_i$ .

	<i>Elicit_Beliefs</i>		<i>True_Distribution</i>	
	(E1)	(E2)	(T1)	(T2)
$E [\# a_{-i}^{SM} = c] \text{ ( "belief" )}$	0.35*** (0.09)	0.33*** (0.11)	—	—
$\# a_{-i}^{SM} = c$	— —	—	0.15*** (0.04)	0.14*** (0.04)
$a_i^{SM}$	—	0.07 (0.24)	—	0.26* (0.14)
<i>Observations</i>	60	60	60	60
<i>LR</i> $\chi^2(1)$	45.64	45.72	15.46	18.94
<i>p-value</i>	<0.001	<0.001	<0.001	<0.001
<i>Pseudo R</i> <sup>2</sup>	0.55	0.55	0.19	0.23

**Dependent variable:** first-mover cooperation ( $a_i^{FM} = 1$ , defection  $a_i^{FM} = 0$ ).

**Regressors:**  $a_i^{SM}$ : second-mover cooperation ( $a_i^{SM} = 1$ , defection  $a_i^{SM} = 0$ ).

$E [\# a_{-i}^{SM} = c]$ : stated belief about the number of second-mover cooperators.

$\# a_{-i}^{SM} = c$ : feedback about the true number of second-mover cooperators.

Marginal effects (at sample means). Standard errors in parenthesis.

\* and \*\*\* indicate significance at the 10%- and 1%-level respectively.

Table 4: Probit regressions (marginal effects).

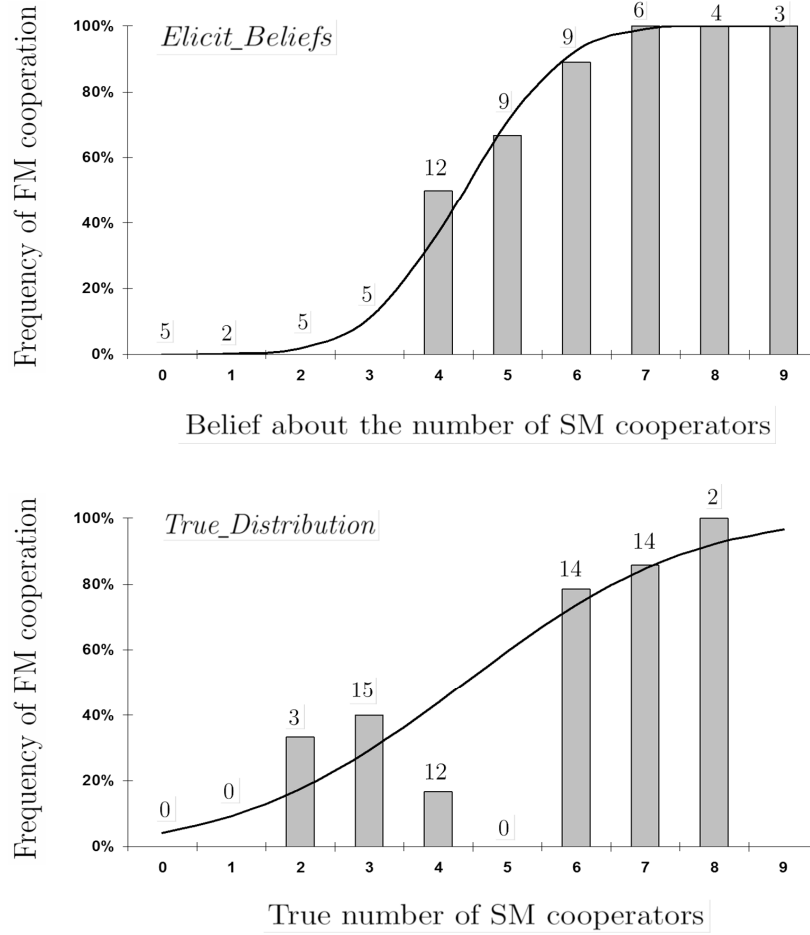


Figure 3: Illustration of probit regressions. Notes: The figure shows how a subject's belief about the number of SM cooperators (top panel), or feedback about the true number of SM cooperators (bottom panel) impact on the frequency of FM cooperation. The gray bars show the actual cooperation rates in the data (above the bars are the number of observations). The smooth black line are the predicted frequencies derived from specifications E1 and T1 in Table 4.

*Elicit\_Beliefs*:

**Result 2** *In Elicit\_Beliefs, the first and second move are positively correlated. First-mover choices are almost always selfish best responses to beliefs, but beliefs are biased toward subjects' own second-mover choices.*

To put it in terms of our stylized model from Section 1, the main finding in this treatment is that  $b_i^j$  and  $s_i$  are positively correlated, and since our subjects choose the selfish best response to their beliefs, we observe a positive correlation of decisions in the first- and second-mover role.

#### 4.4 The *True\_Distribution* treatment

Our final treatment *True\_Distribution* removes the impact of the consensus effect on beliefs. A subject knows the true number of  $a^{SM} = c$  players she faces before making her first-mover choice. Accordingly, this treatment reveals whether first-mover decisions can be explained as selfish best responses to beliefs, or whether the direct channel also operates.

Are subjects best responding to the feedback in *True\_Distribution*? The majority of first movers do: 38 (63.3%) pick the risk neutral best response. Of the remaining 22 subjects, 10 got a feedback that four out of nine cooperated and do not cooperate, which again can be explained by risk aversion.

Note that looking only at the correlation of moves can lead to wrong conclusions in *True\_Distribution*. In particular, the significant correlation of decisions we observe *at the treatment level* ( $p = 0.024$ ) does not necessarily indicate that this is driven by the direct channel. To see this, imagine two experimental sessions. Suppose that every subject defects as second mover in the first session and every subject cooperates in the second session. Now, if all subjects best responded to the feedback they received when making their first-mover choice, the data from both sessions would indicate *all* subjects making the same choice as first and as second movers, even though first-mover choices were completely driven by the feedback. These two hypothetical sessions show that removing the indirect channel in *True\_Distribution* does not preclude a positive correlation of moves, even without the direct channel operating.<sup>20</sup> A better indicator would be the correlation of choices at the session level but then we would have too few observations to make meaningful statements.

To test whether the direct channel operates, we analyze the correlation of first and second moves while controlling for the feedback regarding the second moves—see Table 4. Specification (T1) is an intermediate step which regresses first-mover choices only on the feedback about the exact number

---

<sup>20</sup>Similarly, the nearly identical share of subjects choosing the same action in both moves in *Baseline* and *True\_Distribution* does not imply that the indirect channel does not operate in *Baseline*, because heterogeneity in feedback across sessions drives part of the correlation in *True\_Distribution*.

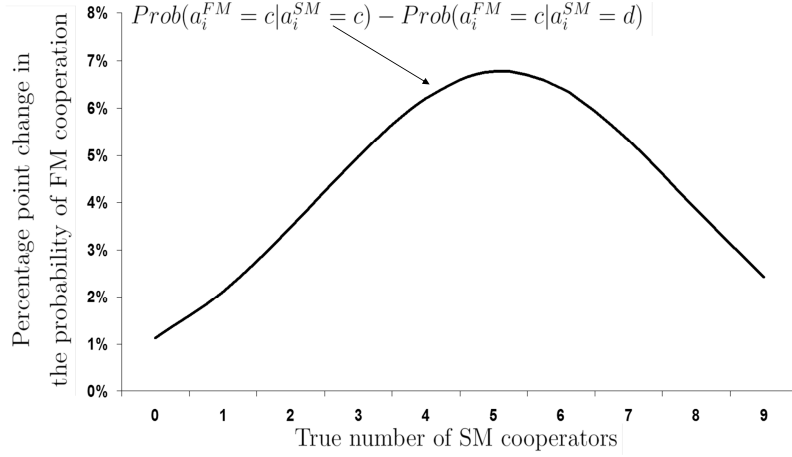


Figure 4: Difference in first-mover cooperation rates between second-mover cooperators and second-mover defectors, conditional on feedback about second-mover cooperation rate

of second-mover cooperators that subject  $i$  faces in her session ( $\# a_{-i}^{SM} = c$ ). In specification (T2), we then add as explanatory variable a dummy for the subject's own second-mover decision (cooperation:  $a_i^{SM} = 1$ , defection:  $a_i^{SM} = 0$ ). The correlation of first and second move prevails even with the feedback given in *True\_Distribution*: the marginal effect of  $a_i^{SM}$  in specification (T2) is 0.26 and it is significant ( $p = 0.056$ ).<sup>21</sup> So, overall, even when we give accurate feedback about second-mover cooperation rates, there still remains a bias toward a player's own type (cooperator or defector).

This bias is also apparent when we consider the 22 subjects who do not play the risk-neutral best response to the feedback. Among these, 16 (73%) choose the same action as first and second mover. In particular, out of seven subjects who choose  $a^{FM} = c$  even though  $a^{FM} = d$  is the best response, five have chosen  $a^{SM} = c$ . Among the 16 players who choose the same action rather than a best response, 11 are second-mover defectors. They fail to cooperate as first mover even though cooperation would maximize their expected payoff. However, 10 among these can again be explained with moderate degrees of risk aversion.

Figure 4 illustrates the difference in predicted cooperation rates of  $a^{SM} = c$  and  $a^{SM} = d$  players,

<sup>21</sup>In robustness checks, we replaced the linear control for " $\# a_{-i}^{SM} = c$ ". First, we included in the regression a set of dummies to distinguish those with "optimistic" induced beliefs ( $\# a_{-i}^{SM} = c \geq 4$ , where the risk-neutral best response is to cooperate) from those with "pessimistic" induced beliefs ( $\# a_{-i}^{SM} = c < 4$ ). Second, we took the signed quadratic distance of the received feedback from the threshold-belief, above which the risk-neutral best response is to cooperate ( $threshold = 9 \times 3/7 \approx 3.9$ ), as an explanatory variable. This captures if those further below (above) the threshold become increasingly more reluctant (more willing) to cooperate relative to those who hold an induced belief close to the threshold. Both specifications yield a positive significant coefficient on  $a_i^{SM}$  ( $p$ -values are 0.014 and 0.049, respectively), suggesting that our results regarding *True\_Distribution* are robust in this respect.

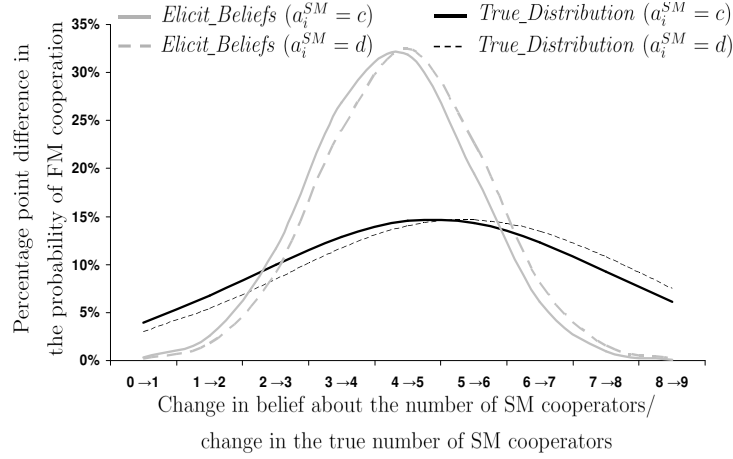


Figure 5: Impact of beliefs/feedback on first-mover cooperation rates

respectively, based on specification (T2). The differences are quantitatively substantial: in the range of feedback of two to eight that we observe in the data, a second-mover cooperator is between three and seven percentage points more likely to cooperate as first mover than a second-mover defector.

**Result 3** *In True\_Distribution, although most subjects best respond to the feedback, we find a significant correlation between the first and second move, despite controlling for feedback.*

In terms of our stylized model from Section 1, the observed positive correlation of first- and second-mover choices suggests that  $f_i(b_i^j)$  and  $s_i$  are positively correlated, for given beliefs  $b_i^j$ . Note that in this treatment, a correlation between preferences to cooperate as first and second mover cannot be driven by a correlation of  $b_i^j$  and  $s_i$  because we fixed  $b_i^j$  by providing our subjects with the number of second-mover cooperators that they were facing in their session.

## 4.5 Discussion

Comparing the *Elicit\_Beliefs* and the *True\_Distribution* data, we note two findings that are relevant for our research question. First, in *True\_Distribution*, a positive correlation between first- and second-mover decisions remains even after conditioning on feedback. This suggests that the correlations found in previous experiments (where such feedback was not given) are not driven exclusively by a consensus effect. This is also consistent with the positive marginal effect of the subject's own second-mover choice in the specification (E2) for *Elicit\_Beliefs* in Table 4, although this effect is small and insignificant. Second, even though best-response behavior in *True\_Distribution* is frequent (63.3 percent of first-mover choices, 80 percent if we allow for a moderate degree of risk aversion), the rate is somewhat below that in *Elicit\_Beliefs* (where 83.3 percent or 93.3 percent best respond, respectively). The differences are significant (if we only consider risk-neutral best

responses: two-sided Fisher exact test,  $p = 0.022$ , Boschloo test,  $p = 0.014$ ; otherwise:  $p = 0.058$ ,  $p = 0.038$ ).

The second finding is consistent with the marginal effects reported in Table 4. Note that the marginal effect of the reported belief in the *Elicit\_Beliefs* treatment is more than twice as large as that of the feedback given in the *True\_Distribution* treatment, suggesting that subjects respond more strongly to their own belief in *Elicit\_Belief* than to the feedback given in *True\_Distribution*. This pattern is illustrated in Figure 5, which plots the marginal effects from specifications (E2) and (T2) in Table 4. As one can see, over most of the range the marginal effect of *belief* in *Elicit\_Beliefs* is much larger than that of feedback about the number of second-mover cooperators in *True\_Distribution* (recall, that we actually only observe feedback values between two and eight). Similarly, the comparison of the top and bottom panels in Figure 3 illustrates this stronger reaction to beliefs in *Elicit\_Belief*. While the above regression-based results need to be taken with a grain of salt because of potential collinearity, the bottom panel in Figure 3 does reveal that in *True\_Distribution* a share of subjects cooperate as first movers even when defection is the risk-neutral best response (that is, when feedback is less than four). This never happens in *Elicit\_Belief*, as the top panel shows.

We can make sense of the above findings as follows. If the direct channel operates, knowledge of how strong the preference for second-mover cooperation  $s_i$  is would help predict the first-mover action. Beliefs correlate with second-mover choices and might actually tell us more about  $s_i$  than the binary second-mover action because they are on a finer grid. As a result, in *Elicit\_Beliefs*, an effect of preferences on the first-mover decision via the direct channel cannot be distinguished from an effect through the indirect channel. Specifically, in *Elicit\_Beliefs* the direct channel can dominate the indirect channel only for subjects with very high (or very low)  $s_i$ . In that case, they might deviate from their selfish best response because they also have a very high (very low)  $f_i$ . But because of the consensus effect these subjects will also have a very high (very low)  $b_i^j$ . This means that the prediction via the direct channel will typically agree with that of the indirect channel.

This suggests an important caveat when interpreting data from social dilemma experiments. Even if regression results seem to attribute the correlation of first- and second-mover choices completely to a consensus effect, this may in fact not be the right conclusion. The direct link between choices and preferences may just be hidden because the constrained set of choices does not fully reflect the intensity of preferences.

From treatment *True\_Distribution* we further infer that combining the inequality aversion models of Bolton and Ockenfels (2000) or Fehr and Schmidt (1999), or the reciprocity model of Dufwenberg and Kirchsteiger (2004) with a consensus effect cannot provide a rationalization for all our results. Such combined models could rationalize the results in *Baseline* and *Elicit\_Beliefs*, as the preference



element can capture the second-mover cooperation and the consensus effect the correlation between first- and second-mover cooperation. But all these models predict a negative correlation, or no correlation between first- and second-mover cooperation when beliefs are exogenously imposed as in *True\_Distribution* – contrary to the positive correlation we find.

As our final point, if subjects are prone to a consensus effect, this should also show up in the beliefs about first-mover choices that we elicit at the end of *True\_Distribution* and *Baseline*. Indeed, we find significant correlations of own first-mover choice and the belief about the other subjects’ first-mover choices in *Baseline* (rank biserial correlation  $r_{rb} = 0.414$ ,  $t = 2.802$ ,  $p < 0.001$ ) and *True\_Distribution* (rank biserial correlation  $r_{rb} = 0.531$ ,  $t = 4.767$ ,  $p < 0.001$ ).

## 5 Conclusion

In spite of its importance for decision making in games, the interaction of preferences and beliefs is rather unexplored in the economics literature. We present an experiment specifically designed to shed light on this interdependence. Recent findings in sequential social dilemma experiments employing within-subjects designs show that subjects who defect as first movers are more likely to exploit first-mover cooperation in their second-mover choice than those who cooperate as first movers. Possible explanations for the positive correlation of first- and second-mover decisions fall into two camps. One predicts an indirect link between preferences and first-mover decisions based on a consensus effect, according to which people think others behave similarly as they do and best respond to these beliefs. The other predicts a direct link between decisions based on some underlying (social) preference – a channel that should operate even if beliefs are held fixed.

To explore whether the direct or indirect channel, or both, are driving the correlation between first- and second-mover decisions, we run three treatments of a sequential-move prisoner’s dilemma experiment. In our baseline treatment, subjects choose in both roles. In a second treatment, we additionally elicit first-order beliefs about second-mover cooperation. In line with previous experiments, we observe a strong correlation of the two moves, no matter whether we elicit beliefs or not. Elicited beliefs, too, are strongly correlated with both moves. This supports the view that the relationship between first- and second-mover decisions operates through the indirect channel. While this result is in line with a number of recent studies on similar games, it is in conflict with traditional views that (at least implicitly) consider beliefs and preferences as independent.

Our third treatment, where we give as feedback the actual frequency of second-mover cooperators before subjects decide their first move, eliminates the indirect channel. Nevertheless, the correlation of the first- and second-mover decisions prevails in this treatment. This suggests that the correlation found in the other treatments and previous experiments is not exclusively driven by a consensus

effect, but that there also is an underlying non-belief based motive affecting both second- and first-mover choices. We discuss a number of social preference theories that would provide a preference-based explanation for the correlation of first- and second-mover cooperation, such as a mixture of total surplus and maximin preferences with concern withdrawal for defecting first movers (Charness and Rabin 2002), or reciprocal altruism (Levine 1998, Cox, Friedman, and Sadiraj 2008).

It is plausible that, when we provide subjects with accurate feedback, this is often different from their originally held beliefs about second-mover behavior. They might thus experience a tension between their natural inclination to defect or cooperate and the information they receive about what is the selfish best response. We cannot precisely tell how subjects resolve this conflict. Our above discussion assumes that they rationally decide given their preferences, or, in terms of our model, player  $i$  decides to cooperate if  $f_i(b_i^j)$  is large enough even if  $b_i^j$  is small (or defect if  $f_i(b_i^j)$  is small enough even if  $b_i^j$  is large). Alternatively, the tension might trigger a psychological mechanism that makes subjects partly discard the feedback we give them and thus lead to insufficient updating. Our design does not allow us to test whether such mechanisms matter here. In a way, though, the precise mechanism might not be that important, because the main message we can derive is that preferences can override feedback, whether by consciously going against selfish payoff maximization or whether by subconsciously suppressing new information. Recent evidence on consensus effects (Engelmann and Strobel 2012), however, does actually not find any support for insufficient updating if experimental subjects obtain information that is in contrast to their prior beliefs.

Our experiment suggests that the consensus effect plays a major role for the observed behavior in social dilemmas. The empirical relevance of behavioral economic theory could thus be increased if it paid closer attention to this effect. Nevertheless, in our findings the direct channel also has a role to play. Hence it is worth to incorporate this channel into models and to further investigate the precise forces at work.

Indeed, the relationship between first- and second-mover behavior as well as beliefs is complex, as has become clear from recent studies using a variety of approaches, including classical laboratory experiments, survey studies, field experiments and physiological studies. In line with our results, studies on trust games, which are structurally similar to our sequential prisoner's dilemma, have shown that the decision to trust is not only determined by beliefs and risk attitudes. See Fehr (2009) for an extensive review and discussion of this issue.

Our paper also showcases a method for dealing with the problem that beliefs may not be randomly distributed in the population studied. The difficulty here is to distinguish best response behavior to beliefs from underlying factors that may influence both beliefs and behavior. By switching off the belief channel, one can identify the direct effect of such underlying factors on behavior. The consensus effect studied in the paper provides one example. Another example is that social

background may influence a person's beliefs (for example, through different life experiences and peer group exposure) as well as preferences for particular actions. With our method, the direct effect of preferences on actions can be disentangled from the indirect effect social background has via beliefs. Our method works similarly to letting subjects play against a programmed strategy, but avoids the disadvantage that the latter method also eliminates the effects of social preferences, which themselves may be important for the research questions to be studied.

## References

- AHN, T. K., E. OSTRÖM, D. SCHMIDT, R. SHUPP, AND J. WALKER (2001): "Cooperation in PD Games: Fear, Greed, and History of Play," *Public Choice*, 106(1-2), 137–55.
- ALLEN, F. (1987): "Discovering Personal Probabilities When Utility Functions Are Unknown," *Management Science*, 33(4), 542–544.
- ALTMANN, S., T. DOHMEN, AND M. WIBRAL (2008): "Do the Reciprocal Trust Less?" *Economics Letters*, 99(3), 454–457.
- ASHRAF, N., I. BOHNET, AND N. PIANKOV (2006): "Decomposing Trust and Trustworthiness," *Experimental Economics*, 9(3), 193–208.
- BERG, J., J. DICKHAUT, AND K. MCCABE (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10(1), 122–142.
- BHATTACHARYA, S., AND P. PFLEIDERER (1985): "Delegated Portfolio Management," *Journal of Economic Theory*, 36(1), 1–25.
- BLANCO, M., D. ENGELMANN, A. K. KOCH, AND H.-T. NORMANN (2010): "Belief Elicitation in Experiments: Is There a Hedging Problem?" *Experimental Economics* 13, 412-438 .
- BLANCO, M., D. ENGELMANN, AND H.-T. NORMANN (2011): "A Within-Subject Analysis of Other-Regarding Preferences," *Games and Economics Behavior* 72(2), 321-338.
- BOLLE, F., AND P. OCKENFELS (1990): "Prisoners' Dilemma as a Game with Incomplete Information," *Journal of Economic Psychology*, 11(1), 69 – 84.
- BOLTON, G. E., AND A. OCKENFELS (2000): "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 90(1), 166–193.
- BOSCHLOO, R. D. (1970): "Raised Conditional Level of Significance for the 2x2-Table when Testing the Equality of Two Probabilities," *Statistica Neerlandica*, 24, 1–35.

- BURKS, S. V., J. P. CARPENTER, L. GOETTE, AND A. RUSTICHINI (2009): “Cognitive Skills Explain Economic Preferences, Strategic Behavior, and Job Attachment,” *Proceedings of the National Academy of Science (USA)*, 106(19), 7745–7750.
- CHARNESS, G., AND M. DUFWENBERG (2006): “Promises and Partnership,” *Econometrica*, 74(6), 1579–1601.
- CHARNESS, G., AND M. RABIN (2002): “Understanding Social Preferences With Simple Tests,” *Quarterly Journal of Economics*, 117(3), 817–869.
- CLARK, K., AND M. SEFTON (2001): “The Sequential Prisoner’s Dilemma: Evidence on Reciprocity,” *Economic Journal*, 111(468), 51–68.
- COSTA-GOMES, M. A., HUCK, S., AND G. WEIZSÄCKER (2010): “Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect,” Discussion paper, University College London.
- COSTA-GOMES, M. A., AND G. WEIZSÄCKER (2008): “Stated Beliefs and Play in Normal-Form Games,” *Review of Economic Studies*, 75(3), 729 – 762.
- COX, J.C. COX, D. FRIEDMAN , AND V. SADIRAJ (2008): “Revealed Altruism,” *Econometrica*, 76(1), 31 – 69.
- CROSON, R. T. A. (2000): “Thinking Like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play,” *Journal of Economic Behavior and Organization*, 41(3), 299–314.
- DAWES, R. M. (1989): “Statistical Criteria for Establishing a Truly False Consensus Effect,” *Journal of Experimental Social Psychology*, 25, 1–17.
- DHAENE, G., AND J. BOUCKAERT (2010): “Sequential Reciprocity in Two-player Two-stage Games: An Experimental Analysis,” *Games and Economic Behavior* 70, 289–303.
- DUFWENBERG, M., AND U. GNEEZY (2000): “Measuring Beliefs in an Experimental Lost Wallet Game,” *Games and Economic Behavior*, 30, 163–82.
- DUFWENBERG, M., AND G. KIRCHSTEIGER (2004): “A Theory of Sequential Reciprocity,” *Games and Economic Behavior*, 47(2), 268–298.
- ELLINGSEN, T., M. JOHANNESSON, G. TORSVIK, AND S. TJØTTA (2010): “Testing Guilt Aversion,” *Games and Economic Behavior*, 68(1), 95–107.
- ENGELMANN, D., AND M. STROBEL (2000): “The False Consensus Effect Disappears if Representative Information and Monetary Incentives are Given,” *Experimental Economics*, 3(3), 241–260.

- ENGELMANN, D., AND M. STROBEL (2004): “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments,” *American Economic Review*, 94(4), 857–869.
- ENGELMANN, D., AND M. STROBEL (2012): “Deconstruction and Reconstruction of an Anomaly,” *Games and Economic Behavior*, 76(2), 678–689.
- FEHR, E. (2009): “On the Economics and Biology of Trust,” *Journal of the European Economic Association*, 7(2-3), 235–266.
- FEHR, E., G. KIRCHSTEIGER, AND A. RIEDL (1993): “Does Fairness Prevent Market Clearing? An Experimental Investigation,” *Quarterly Journal of Economics*, 108(2), 437–460.
- FEHR, E., AND K. M. SCHMIDT (1999): “A Theory Of Fairness, Competition, And Cooperation,” *Quarterly Journal of Economics*, 114(3), 817–868.
- FISCHBACHER, U. (2007): “Z-Tree - Zurich Toolbox for Ready-Made Economic Experiments,” *Experimental Economics*, 10(2), 171–178.
- GÄCHTER, S., D. NOSENZO, E. RENNER, AND M. SEFTON (2012): “Who Makes a Good Leader? Cooperativeness, Optimism, and Leading-by-Example,” *Economic Inquiry*, 50(4), 953–967.
- GÄCHTER, S., AND E. RENNER (2010): “The Effects of (Incentivized) Belief Elicitation in Public Good Experiments,” *Experimental Economics* 13, 364–377.
- GUERRA, G., AND D. J. ZIZZO (2004): “Trust Responsiveness and Beliefs,” *Journal of Economic Behavior and Organization*, 55(1), 25–30.
- HOLT, C. A. (1986): “Scoring-Rule Procedures for Eliciting Subjective Probability and Utility Functions,” in *Bayesian Inference and Decision Techniques*, ed. by P. Goel, and A. Zellner, pp. 279–290. Elsevier, Amsterdam.
- HOLT, C. A., AND S. K. LAURY (2002): “Risk Aversion and Incentive Effects,” *American Economic Review*, 92(5), 1644–1655.
- HUCK, S., AND G. WEIZSÄCKER (2002): “Do Players Correctly Estimate what Others Do? Evidence of Conservatism in Beliefs,” *Journal of Economic Behavior and Organization*, 47(1), 71–85.
- KARNI, E. (2009): “A Mechanism for Eliciting Probabilities,” *Econometrica*, 77(2), 603–606.
- KOCH, A. K., A. MORGENSTERN, AND P. RAAB (2009): “Career Concerns Incentives: An Experimental Test,” *Journal of Economic Behavior and Organization* 72, 571–588.

- KOSFELD, M., M. HEINRICHS, P. J. ZAK, U. FISCHBACHER, AND E. FEHR (2005): “Oxytocin Increases Trust in Humans,” *Nature*, 435, 673–676.
- KRANZ, S. (2010): “Moral Norms in a Partly Compliant Society,” *Games and Economic Behavior* 68, 255–274.
- LEVINE, D. K. (1998): “Modeling Altruism and Spitefulness in Experiment,” *Review of Economic Dynamics*, 1(3), 593–622.
- MULLEN, B., J. L. ATKINS, D. S. CHAMPION, C. EDWARDS, S. HARDLY, D., J. E., AND M. VANDERKLOK (1985): “The False Consensus Effect: A Meta-Analysis of 155 Hypothesis Tests,” *Journal of Experimental Social Psychology*, 21, 262–283.
- MURPHY, A. H., AND R. L. WINKLER (1970): “Scoring Rules in Probability Assessment and Evaluation,” *Acta Psychologica*, 34, 273–286.
- NAEF, M., E. FEHR, U. FISCHBACHER, J. SCHUPP, AND G. WAGNER (2008): “Decomposing Trust: Explaining National and Ethnical Trust Differences,” Working Paper, Institute for Empirical Research in Economics, University of Zurich.
- OFFERMAN, T., J. SONNEMANS, G. VAN DE KUILEN, AND P. P. WAKKER (2009): “A Truth-Serum for Non-Bayesians: Correcting Proper Scoring Rules for Risk Attitudes,” *Review of Economic Studies*, 76(4), 1461–1489.
- OSBORNE, M.J. (2009): *An Introduction to Game Theory*. Oxford University Press, Oxford, U.K.
- POTTERS, J., M. SEFTON, AND L. VESTERLUND (2007): “Leading-by-Example and Signaling in Voluntary Contribution Games: An Experimental Study,” *Economic Theory*, 33(1), 169–182.
- REY-BIEL, P. (2009): “Equilibrium Play and Best Response to (Stated) Beliefs in Normal Form Games,” *Games and Economic Behavior*, 65(2), 572–585.
- SAVAGE, L. J. (1971): “Elicitation of Personal Probabilities and Expectations,” *Journal of the American Statistical Association*, 66(336), 783–801.
- SCHLAG, K. H., AND J. VAN DER WEELE (2013): “Eliciting Probabilities, Means, Medians, Variances and Covariances without Assuming Risk Neutrality?” *Theoretical Economics Letters*, 3(1), 38–42.
- SELTEN, R. (1998): “Axiomatic Characterization of the Quadratic Scoring Rule,” *Experimental Economics*, 1(1), 43–61.

SIMPSON, B. T. (2003): "Sex, Fear, and Greed: A Social Dilemma Analysis of Gender and Cooperation," *Social Forces*, 82(1), 35–52.

# Appendix

## A Proofs

In this appendix, we prove statements made in Section 3 regarding Propositions 1 and 2.

### A.1 Fehr-Schmidt inequality-averse preferences imply a negative correlation between first- and second-mover behavior

Consider *inequality averse* players with Fehr and Schmidt (1999) preferences as described in the main text. Second movers who dislike advantageous inequality will be reluctant to exploit first-mover cooperation. More precisely, a second mover  $i$  will choose  $a^{SM} = c$  if and only if (ignoring the probability zero event of indifference)

$$\begin{aligned} 14 - \alpha_i \max\{14 - 14, 0\} - \beta_i \max\{14 - 14, 0\} &> 17 - \alpha_i \max\{7 - 17, 0\} - \beta_i \max\{17 - 7, 0\} \\ \Leftrightarrow 14 &> 17 - 10\beta_i \\ \Leftrightarrow \beta_i &> 0.3 \end{aligned}$$

Therefore, players with  $\beta_i > 0.3$  will behave as if they have a large  $s_i$  in our stylized model. For the first move, note that the utility of  $U_i(a^{FM} = d) = 10$  is independent of  $\alpha_i$  and  $\beta_i$ . However,

$$U_i(a^{FM} = c) = 14b_i^j + (1 - b_i^j)(7 - 10\alpha_i) = 7 + 7b_i^j - 10(1 - b_i^j)\alpha_i$$

is decreasing in  $\alpha_i$  and therefore for any  $b_i^j < 1$ ,  $i$  is less likely to cooperate the larger  $\alpha_i$ . Hence players with a larger  $\alpha_i$  are less likely to cooperate as first mover whereas players with sufficiently large  $\beta_i$  will cooperate as second mover. Therefore, if beliefs  $b_i^j$  are independent of preferences, a positive correlation between  $\alpha_i$  and  $\beta_i$  implies a negative correlation between first- and second-mover cooperation.

### A.2 Dufwenberg-Kirchsteiger reciprocity preferences imply a negative correlation between first- and second-mover behavior

In the Dufwenberg and Kirchsteiger (2004) model,  $i$ 's kindness toward  $j$ ,  $k_{ij}$ , is measured by comparison of the expected resulting payoff of  $j$  and the equitable payoff, which is in turn the average of  $j$ 's maximum possible payoff and  $j$ 's minimum possible efficient payoff. Since first-mover cooperation yields more than the equitable payoff whatever the second mover chooses, first-mover cooperation is unambiguously kind and first-mover defection is unambiguously unkind. Player  $i$ 's expectation about  $j$ 's kindness as second mover is

$$\lambda_{iji} = 17b_i^j + 7(1 - b_i^j) - \frac{17 + 7}{2} = 10b_i^j - 5.$$



Player  $i$  thus expects  $j$  to be kind if  $i$  expects  $j$  to be more likely to cooperate than to defect. The reciprocity term  $Y_i k_{ij} \lambda_{iji}$  is thus positive whenever  $i$  cooperates as first mover and  $b_i^j > 1/2$  or  $i$  defects and  $b_i^j < 1/2$  and (weakly) negative otherwise. In terms of our stylized model, this translates as  $f_i(b_i^j)$  is increasing in  $Y_i$  for  $b_i^j > 1/2$ , but decreasing in  $Y_i$  if  $b_i^j < 1/2$ . Hence a reciprocal ( $Y_i > 0$ ) player is *more* inclined to cooperate than a selfish ( $Y_i = 0$ ) player if her belief about the second-mover cooperation probability is greater than  $1/2$  and *less* inclined to cooperate if her belief is smaller than  $1/2$ . Given that only second-movers who are sufficiently reciprocal will cooperate, this means that players are predicted to behave as if  $f_i(b_i^j)$  and  $s_i$  are negatively correlated for  $b_i^j < 1/2$  and positively correlated for  $b_i^j > 1/2$ .

Considering the implications for the SPD with the parameters of our experiment, remember that expected material payoffs  $\pi_i$  are maximized by  $a^{FM} = d$  for  $b_i^j < 3/7$  and by  $a^{FM} = c$  for  $b_i^j > 3/7$ . Therefore if  $b_i^j < 3/7$ , both the material payoff term as well as the reciprocity term in  $U_i = \pi_i + Y_i k_{ij} \lambda_{iji}$  are maximized by  $a^{FM} = d$  whatever  $i$ 's sensitivity to reciprocity. Similarly, for  $b_i^j > 1/2$ , both the material payoff term as well as the reciprocity term are maximized by  $a^{FM} = c$  whatever  $i$ 's sensitivity to reciprocity. In contrast for  $3/7 < b_i^j < 1/2$ , the material payoff term  $\pi_i$  will be maximized by first-mover cooperation, but the reciprocity term  $Y_i k_{ij} \lambda_{iji}$  will be negative for  $a^{FM} = c$  and positive for  $a^{FM} = d$ . A relatively selfish individual would therefore choose  $a^{FM} = c$ , whereas if  $i$  is sufficiently strongly concerned with reciprocity, she would choose  $a^{FM} = d$ . In our experiment the only possible belief with  $3/7 < b_i^j < 1/2$  is that four out of nine second movers cooperate. Therefore, only for this belief, the behavior of reciprocal first movers will differ from that of selfish first movers. In this case, a selfish first mover will cooperate, whereas a sufficiently reciprocal player will defect. Sufficiently reciprocal players will also cooperate as second movers. Therefore, if players who have preferences as in Dufwenberg and Kirchsteiger (2004) but differ in their concerns for reciprocity, it is those who might be cooperating as second movers who are more likely to defect as first movers, but first-mover behavior will only be affected for a belief that four out of nine second movers cooperate. Hence, if there is any correlation between first and second moves at all, it should be negative, assuming again that beliefs are independent of preferences.

### A.3 Cox-Friedman-Sadiraj reciprocity preferences imply a positive correlation between first- and second-mover behavior

In Cox, Friedman, and Sadiraj (2008) a preference ordering  $\mathcal{A}$  over “my” and “your” payoff is considered more altruistic than another preference ordering  $\mathcal{B}$  if we get for the willingness to pay to increase the other's payoff in terms of own payoff  $WTP_{\mathcal{A}}(m, y) \geq WTP_{\mathcal{B}}(m, y)$  for all  $(m, y)$  in a given domain  $D$ . It is reasonable to assume that participants whose preferences are more altruistic

than others' on some domain tend to also be more altruistic on other domains.<sup>22</sup> This means that a participant who is more altruistic and hence more likely to cooperate when he is a first mover in the SPD tends to also be more altruistic and hence more likely to cooperate if he is a second mover in the SPD. In terms of our stylized model, if a participant is (generally) more altruistic, he will behave as if both his  $f_i(b_i^j)$  (for any given  $b_i^j$ ) and his  $s_i$  are larger. Heterogeneity in the degree of altruism would thus imply a correlation between first- and second-mover choices. A key component of the Cox, Friedman, and Sadiraj (2008) model is their reciprocity axiom. Opportunity set  $G$  for the second mover is more generous than opportunity set  $F$  if for maximum achievable own ( $m^*$ ) and other's ( $y^*$ ) payoffs we have  $m_G^* \geq m_F^*$  and  $m_G^* - m_F^* \geq y_G^* - y_F^*$ . The reciprocity axiom then says that if a first mover chooses a more generous opportunity set, then the second mover's preferences become more altruistic. This reciprocity axiom, and the convexity of preferences in both own and other's payoff, imply that the model is consistent with second movers being more likely to cooperate after first-mover cooperation than after first-mover defection because the second-mover's opportunity set after first-mover cooperation is more generous than after first-mover defection.

## B Experimental Instructions

### B.1 Instructions for *Elicit\_Beliefs*

*As described in Section 2.1, we conducted two variants of Elicit\_Beliefs to test for hedging confounds. In Variant 1 (three sessions) we pay for both the decisions and the beliefs, whereas in Variant 2 (the remaining three sessions), we pay for either the belief or the decision (both with equal probability).*

You are now taking part in an experiment. If you read the following instructions carefully, you can, depending on your and other participants' decisions, earn a considerable amount of money. It is therefore important that you take your time to understand the instructions. Please do not communicate with the other participants during the experiment. If you have any questions, please raise your hand and ask us. All the information you provide will be treated anonymously.

At the end of the experiment your earnings will be converted from Experimental Currency Units (ECU) to Pounds Sterling at a rate of [*Variant 1*: ECU 2 = £ 1/ *Variant 2*: ECU 1 = £ 1], and paid to you in cash. Your earnings will also be treated confidentially.

**SITUATION UNDERLYING THE EXPERIMENT:** We start by explaining the situation underlying the experiment, which is represented in Figure 1 [corresponds to Figure 1 in this paper, but with the neutral frame labels]. There are two people involved, Person A and Person B. Person A can choose

---

<sup>22</sup>It is in general not obvious that one can always order participants according to their altruism. For example, A may be more altruistic than B if others have a higher payoff while B may be more altruistic than A if others have a lower payoff.

between two options: IN or OUT. If Person A picks OUT, Person B has no choice to make, and both Person A and B get ECU 10 each. If Person A picks IN, Person B then has a choice between two options: LEFT or RIGHT. If LEFT is chosen, both Person A and B get ECU 14 each. If RIGHT is chosen, Person A gets ECU 7 and Person B gets ECU 17.

OVERVIEW OF THE EXPERIMENT: The experiment consists of three parts. You and the other participants will each make decisions both in the role of Person A and of Person B. Additionally, we will ask you to make a guess how the other participants in the room decided. At the end of the experiment, the computer will randomly assign you either the role of Person A or the role of Person B, and will also randomly match you and the other participants in pairs. Note that you will have to make your decisions without knowing the role that you will ultimately be assigned. Also, at the time when you make your decisions, you will not know the decision made by the participant matched to you. Below, we will explain how your payment from the experiment is determined. But let us first have a closer look at your tasks in the order that they will appear.

1. *Decision Task B*: You have the role of Person B in the situation described in Figure 1. Given that Person A chose IN: Do you choose LEFT or RIGHT?
2. *Guess Task*: There are 10 participants in the room, you and 9 other participants. All of them also did the above Decision Task B. How many of the 9 other participants do you think chose LEFT?
3. *Decision Task A*: Now you have the role of Person A. Do you choose IN or OUT?

PAYMENTS: [*Variant 1*: At the end of the experiment you will be paid both for the Decision Tasks and for the Guess Task. Your overall payoff will be converted at a rate of  $\text{ECU } 2 = \text{£ } 1$ . Payoffs for the individual tasks are determined as follows.

*Payoff for the Decision Tasks*: As mentioned,...]

[*Variant 2*: At the end of the experiment, the computer will randomly decide whether your payment will be based on the Decision Tasks or the Guess Task. Each type of tasks is equally likely to be the one determining your payoff, and will be the same for all subjects. (This means whenever you are paid based on the Decision Tasks, also all other participants are paid based on the Decision Tasks; and whenever you are paid for the Guess Task, this is also the case for all other participants.) Your overall payoff will be converted at a rate of  $\text{ECU } 1 = \text{£ } 1$ . Depending on the random draw of the computer, payoffs are determined as follows.

*Payoff if the random draw of the computer selects the Decision Tasks*: As mentioned,] the computer will randomly and anonymously pair you with another participant in the room. One of you will randomly be assigned the role of Person A, and the other one will be assigned the role of Person

B. The computer will then take your and the other participant’s relevant Decision Task choices to compute your payoffs as shown in Figure 1.

[*Variant 1: Payoff for the Guess Task:* In addition to the payoff for the Decision Tasks, you receive a payoff for the Guess Task, which depends...]

[*Variant 2: Payoff if the random draw of the computer selects the Guess Task:* The payoff for the Guess Task depends] on the accuracy of your guess. The better your guess, the higher will be your payoff. Take a look at Table 1 (on a separate page) [corresponds to Table 5]. The table shows how your guess and the actual choices of the other participants determine your payoff.

- You can see that a perfect guess earns you ECU 15. For example, if your guess was 6, and if there are actually 6 people who chose LEFT in Decision Task B, you get ECU 15.
- If your guess is completely off the mark you earn nothing. This occurs if you guess that 9 other participants chose LEFT, while none of them did so; or if you guess that none of the other participants chose LEFT, while all of them did so.
- Otherwise, your payoff depends on how close to accurate your guess was. For example, if 6 out of the other 9 participants chose LEFT, and your guess was that 3 participants would do so, you earn ECU 13.30.

Before starting with the actual experiment, we will ask you to answer a few control questions. Then we will go through the three parts of the experiment. There will be plenty of time before each decision to ask questions. At the end of the experiment we ask you to answer a few questions. These answers will not affect your final payment.

Are there any questions? If so, please raise your hand.

## B.2 Instructions for *Baseline*

*The instructions are the same as in Elicit\_Beliefs (Variant 2), except that the Guess Task (included only to achieve balanced designs) now comes last, and asks about the first-mover decisions of the other participants:*

1. *Decision Task B:* You have the role of Person B in the situation described in Figure 1. Given that Person A chose IN: Do you choose LEFT or RIGHT?
2. *Decision Task A:* Now you have the role of Person A. Do you choose IN or OUT?
3. *Guess Task:* There are 10 participants in the room, you and 9 other participants. All of them also did the above Decision Task A. How many of the 9 other participants do you think chose IN?

		Actual number of other participants choosing LEFT									
		9	8	7	6	5	4	3	2	1	0
<b>Your guess: number of other participants choosing LEFT</b>	<b>9</b>	15.0	14.8	14.3	13.3	12.0	10.4	8.3	5.9	3.1	0.0
	<b>8</b>	14.8	15.0	14.8	14.3	13.3	12.0	10.4	8.3	5.9	3.1
	<b>7</b>	14.3	14.8	15.0	14.8	14.3	13.3	12.0	10.4	8.3	5.9
	<b>6</b>	13.3	14.3	14.8	15.0	14.8	14.3	13.3	12.0	10.4	8.3
	<b>5</b>	12.0	13.3	14.3	14.8	15.0	14.8	14.3	13.3	12.0	10.4
	<b>4</b>	10.4	12.0	13.3	14.3	14.8	15.0	14.8	14.3	13.3	12.0
	<b>3</b>	8.3	10.4	12.0	13.3	14.3	14.8	15.0	14.8	14.3	13.3
	<b>2</b>	5.9	8.3	10.4	12.0	13.3	14.3	14.8	15.0	14.8	14.3
	<b>1</b>	3.1	5.9	8.3	10.4	12.0	13.3	14.3	14.8	15.0	14.8
	<b>0</b>	0.0	3.1	5.9	8.3	10.4	12.0	13.3	14.3	14.8	15.0

Table 5: Payoff for the Guess Task(in ECU)

### B.3 Instructions for *True\_Distribution*

*The instructions are the same as in Baseline, except that after Decision Task B there is feedback about the second-mover decisions of the other participants:*

1. *Decision Task B:* You have the role of Person B in the situation described in Figure 1. Given that Person A chose IN: Do you choose LEFT or RIGHT?

The decision task is followed by a feedback stage. There are 10 participants in the room, you and 9 other participants. All of them also did the above Decision Task B. The feedback stage informs you about how many of the 9 other participants chose LEFT in Decision Task B.

2. *Decision Task A:* ...

## C Oral Summaries

### C.1 Oral Summary for *Elicit\_Beliefs*

Welcome to the experiment. Please do not communicate with the other participants during the experiment. If you have any questions, please raise your hand and ask us. At your seat you will find a set of instructions. Read them carefully now. Please answer the questions you find on a separate page and raise your hand if you are finished. Before the experiment starts we will give a brief summary.

*After instructions were read, before Decision Task B:* To summarize: Please look at Figure 1 in the instructions. The experiment starts with Decision Task B. Next will be the Guess Task, and finally we come to Decision Task A. You will have to do each task only once. We will briefly summarize the tasks when we get to them.

[*Variant 2:* At the end of the experiment the computer will randomly decide for all participants whether the Decision Tasks are going to be the basis for payments, or the Guess task.]

At the end of the experiment, the computer randomly matches you with one of the other participants in the room. One of you will be assigned the role of Person A and the other that of Person B — both roles are equally likely. [*Variant 1:* The payoffs for the Decision Tasks will then be computed based on your and the other participant's choices in the relevant Decision Tasks. In addition, the Guess task will be paid. The whole amount will then be converted to Pounds Sterling at a rate of  $2 \text{ ECU} = \text{£ } 1$ / *Variant 2:* If the Decision tasks will be the basis for payments, the payoffs will then be computed based on your and the other participant's choices in the relevant Decision tasks. Otherwise, the Guess task will be paid. The payoff amount will then be converted to Pounds Sterling at a rate of  $1 \text{ ECU} = \text{£ } 1$ .]

We start with Decision Task B. You are asked to make a choice between LEFT and RIGHT for the case that you are assigned the role of Person B and Person A chose IN before. If you choose LEFT, both you and the other participant matched to you will get 14 ECU. If you choose RIGHT, you get 17 ECU and the other participant 7 ECU. Note that you will learn your actual role only at the end of the experiment. Also, if you actually are assigned the role of Person B you will learn Person A's actual choice only at the end of the experiment. [*Variant 2:* You will also learn only at the end of the experiment whether the Decision Tasks or the Guess Task will determine the payoffs.]

Are there any questions?

*Between Decision Task B and Guess Task:* We now come to the Guess Task. You are asked to guess how many of the 9 other participants in the room chose LEFT in the Decision Task B. Have a look at Table 1. It shows how your guess and the actual choices of the other participants determine your payoff. Also, go through the examples given in the instructions (p.3 bottom).

Are there any questions?

*Between Guess Task and Decision Task A:* We now come to Decision Task A. You are asked to make a choice between IN and OUT for the case that you are assigned the role of Person A. If you choose IN, your payoff and that of the other participant matched to you in the role of Person B depend on the choice between LEFT and RIGHT of that participant, as described in Figure 1. If you choose OUT, both of you receive 10 ECU, and the choice of the other participant matched to you is irrelevant for payoffs. Again, you learn your actual role only at the end of the experiment.

Are there any questions?

## C.2 Oral Summary for *Baseline*

*The summary is the same as in Elicit\_Beliefs (Variant 2), except that the Guess Task now comes last, and asks about the first-mover decisions of the other participants.*

## C.3 Oral Summary for *True\_Distribution*

*The summary is the same as in Baseline, except that there is an additional Feedback Stage after Decision Task B:*

*Between Decision Task B and Feedback Stage:*

All participants in the room did the above Decision Task B. Now you are informed about how many of the 9 other participants in the room chose LEFT in Decision Task B.

Note that if you are assigned the role of Person A, one of these 9 choices is the choice of the person you will be matched with. This means, for example, that if the information is that 9 out of 9 chose LEFT, then you know for sure that you would be matched with a Person B who chose LEFT. So if you chose IN, you would get 14 ECU. If the information is that 0 out of 9 chose LEFT, then you know for sure that you would be matched with a Person B who chose RIGHT. In this case, if you chose IN, you would get 7 ECU. As a third example, suppose that the information is that 6 out of 9 chose LEFT. This means that there is a 2 in 3 chance that you would be matched with a Person B who chose LEFT (and you would get 14 ECU if you chose IN), while there is a 1 in 3 chance that you would be matched with a Person B who chose RIGHT (and you would get 7 ECU if you chose IN).

Are there any questions?

*Between Feedback Stage and Decision Task A: ...*

## PREVIOUS DISCUSSION PAPERS

- 146 Alipranti, Maria, Milliou, Chrysovalantou and Petrakis, Emmanuel, Price vs. Quantity Competition in a Vertically Related Market, May 2014.  
Forthcoming in: Economics Letters.
- 145 Blanco, Mariana, Engelmann, Dirk, Koch, Alexander K., and Normann, Hans-Theo, Preferences and Beliefs in a Sequential Social Dilemma: A Within-Subjects Analysis, May 2014.  
Forthcoming in: Games and Economic Behavior.
- 144 Jeitschko, Thomas D., Jung, Yeonjei and Kim, Jaesoo, Bundling and Joint Marketing by Rival Firms, May 2014.
- 143 Benndorf, Volker and Normann, Hans-Theo, The Willingness to Sell Private Data, May 2014.
- 142 Dauth, Wolfgang and Suedekum, Jens, Globalization and Local Profiles of Economic Growth and Industrial Change, April 2014.
- 141 Nowak, Verena, Schwarz, Christian and Suedekum, Jens, Asymmetric Spiders: Supplier Heterogeneity and the Organization of Firms, April 2014.
- 140 Hasnas, Irina, A Note on Consumer Flexibility, Data Quality and Collusion, April 2014.
- 139 Baye, Irina and Hasnas, Irina, Consumer Flexibility, Data Quality and Location Choice, April 2014.
- 138 Aghadadashli, Hamid and Wey, Christian, Multi-Union Bargaining: Tariff Plurality and Tariff Competition, April 2014.
- 137 Duso, Tomaso, Herr, Annika and Suppliet, Moritz, The Welfare Impact of Parallel Imports: A Structural Approach Applied to the German Market for Oral Anti-diabetics, April 2014.  
Forthcoming in: Health Economics.
- 136 Haucap, Justus and Müller, Andrea, Why are Economists so Different? Nature, Nurture and Gender Effects in a Simple Trust Game, March 2014.
- 135 Normann, Hans-Theo and Rau, Holger A., Simultaneous and Sequential Contributions to Step-Level Public Goods: One vs. Two Provision Levels, March 2014.  
Forthcoming in: Journal of Conflict Resolution.
- 134 Bucher, Monika, Hauck, Achim and Neyer, Ulrike, Frictions in the Interbank Market and Uncertain Liquidity Needs: Implications for Monetary Policy Implementation, March 2014.
- 133 Czarnitzki, Dirk, Hall, Bronwyn, H. and Hottenrott, Hanna, Patents as Quality Signals? The Implications for Financing Constraints on R&D?, February 2014.
- 132 Dewenter, Ralf and Heimeshoff, Ulrich, Media Bias and Advertising: Evidence from a German Car Magazine, February 2014.  
Forthcoming in: Review of Economics.
- 131 Baye, Irina and Sapi, Geza, Targeted Pricing, Consumer Myopia and Investment in Customer-Tracking Technology, February 2014.



- 130 Clemens, Georg and Rau, Holger A., Do Leniency Policies Facilitate Collusion? Experimental Evidence, January 2014.
- 129 Hottenrott, Hanna and Lawson, Cornelia, Fishing for Complementarities: Competitive Research Funding and Research Productivity, December 2013.
- 128 Hottenrott, Hanna and Rexhäuser, Sascha, Policy-Induced Environmental Technology and Inventive Efforts: Is There a Crowding Out?, December 2013.
- 127 Dauth, Wolfgang, Findeisen, Sebastian and Suedekum, Jens, The Rise of the East and the Far East: German Labor Markets and Trade Integration, December 2013. Forthcoming in: Journal of European Economic Association.
- 126 Wenzel, Tobias, Consumer Myopia, Competition and the Incentives to Unshroud Add-on Information, December 2013. Published in: Journal of Economic Behavior and Organization, 98 (2014), pp. 89-96.
- 125 Schwarz, Christian and Suedekum, Jens, Global Sourcing of Complex Production Processes, December 2013. Forthcoming in: Journal of International Economics.
- 124 Defever, Fabrice and Suedekum, Jens, Financial Liberalization and the Relationship-Specificity of Exports, December 2013. Published in: Economics Letters, 122 (2014), pp. 375-379.
- 123 Bauernschuster, Stefan, Falck, Oliver, Heblich, Stephan and Suedekum, Jens, Why Are Educated and Risk-Loving Persons More Mobile Across Regions?, December 2013. Published in: Journal of Economic Behavior and Organization, 98 (2014), pp. 56-69.
- 122 Hottenrott, Hanna and Lopes-Bento, Cindy, Quantity or Quality? Knowledge Alliances and their Effects on Patenting, December 2013.
- 121 Hottenrott, Hanna and Lopes-Bento, Cindy, (International) R&D Collaboration and SMEs: The Effectiveness of Targeted Public R&D Support Schemes, December 2013. Forthcoming in: Research Policy.
- 120 Giesen, Kristian and Suedekum, Jens, City Age and City Size, November 2013.
- 119 Trax, Michaela, Brunow, Stephan and Suedekum, Jens, Cultural Diversity and Plant-Level Productivity, November 2013.
- 118 Manasakis, Constantine and Vlassis, Minas, Downstream Mode of Competition With Upstream Market Power, November 2013. Published in: Research in Economics, 68 (2014), pp. 84-93.
- 117 Sapi, Geza and Suleymanova, Irina, Consumer Flexibility, Data Quality and Targeted Pricing, November 2013.
- 116 Hinlopen, Jeroen, Müller, Wieland and Normann, Hans-Theo, Output Commitment Through Product Bundling: Experimental Evidence, November 2013. Published in: European Economic Review 65 (2014), pp. 164-180.
- 115 Baumann, Florian, Denter, Philipp and Friehe Tim, Hide or Show? Endogenous Observability of Private Precautions Against Crime When Property Value is Private Information, November 2013.
- 114 Fan, Ying, Kühn, Kai-Uwe and Lafontaine, Francine, Financial Constraints and Moral Hazard: The Case of Franchising, November 2013.

- 113 Aguzzoni, Luca, Argentesi, Elena, Buccirossi, Paolo, Ciari, Lorenzo, Duso, Tomaso, Tognoni, Massimo and Vitale, Cristiana, They Played the Merger Game: A Retrospective Analysis in the UK Videogames Market, October 2013.  
Forthcoming in: Journal of Competition and Economics under the title: "A Retrospective Merger Analysis in the UK Videogames Market".
- 112 Myrseth, Kristian Ove R., Riener, Gerhard and Wollbrant, Conny, Tangible Temptation in the Social Dilemma: Cash, Cooperation, and Self-Control, October 2013.
- 111 Hasnas, Irina, Lambertini, Luca and Palestini, Arsen, Open Innovation in a Dynamic Cournot Duopoly, October 2013.  
Published in: Economic Modelling, 36 (2014), pp. 79-87.
- 110 Baumann, Florian and Friehe, Tim, Competitive Pressure and Corporate Crime, September 2013.
- 109 Böckers, Veit, Haucap, Justus and Heimeshoff, Ulrich, Benefits of an Integrated European Electricity Market, September 2013.
- 108 Normann, Hans-Theo and Tan, Elaine S., Effects of Different Cartel Policies: Evidence from the German Power-Cable Industry, September 2013.  
Forthcoming in: Industrial and Corporate Change.
- 107 Haucap, Justus, Heimeshoff, Ulrich, Klein, Gordon J., Rickert, Dennis and Wey, Christian, Bargaining Power in Manufacturer-Retailer Relationships, September 2013.
- 106 Baumann, Florian and Friehe, Tim, Design Standards and Technology Adoption: Welfare Effects of Increasing Environmental Fines when the Number of Firms is Endogenous, September 2013.
- 105 Jeitschko, Thomas D., NYSE Changing Hands: Antitrust and Attempted Acquisitions of an Erstwhile Monopoly, August 2013.  
Published in: Journal of Stock and Forex Trading, 2 (2) (2013), pp. 1-6.
- 104 Böckers, Veit, Giessing, Leonie and Rösch, Jürgen, The Green Game Changer: An Empirical Assessment of the Effects of Wind and Solar Power on the Merit Order, August 2013.
- 103 Haucap, Justus and Muck, Johannes, What Drives the Relevance and Reputation of Economics Journals? An Update from a Survey among Economists, August 2013.
- 102 Jovanovic, Dragan and Wey, Christian, Passive Partial Ownership, Sneaky Takeovers, and Merger Control, August 2013.
- 101 Haucap, Justus, Heimeshoff, Ulrich, Klein, Gordon J., Rickert, Dennis and Wey, Christian, Inter-Format Competition Among Retailers – The Role of Private Label Products in Market Delineation, August 2013.
- 100 Normann, Hans-Theo, Requate, Till and Waichman, Israel, Do Short-Term Laboratory Experiments Provide Valid Descriptions of Long-Term Economic Interactions? A Study of Cournot Markets, July 2013.  
Forthcoming in: Experimental Economics.
- 99 Dertwinkel-Kalt, Markus, Haucap, Justus and Wey, Christian, Input Price Discrimination (Bans), Entry and Welfare, June 2013.
- 98 Aguzzoni, Luca, Argentesi, Elena, Ciari, Lorenzo, Duso, Tomaso and Tognoni, Massimo, Ex-post Merger Evaluation in the UK Retail Market for Books, June 2013.

- 97 Caprice, Stéphane and von Schlippenbach, Vanessa, One-Stop Shopping as a Cause of Slotting Fees: A Rent-Shifting Mechanism, May 2012.  
Published in: Journal of Economics and Management Strategy, 22 (2013), pp. 468-487.
- 96 Wenzel, Tobias, Independent Service Operators in ATM Markets, June 2013.  
Published in: Scottish Journal of Political Economy, 61 (2014), pp. 26-47.
- 95 Coublucq, Daniel, Econometric Analysis of Productivity with Measurement Error: Empirical Application to the US Railroad Industry, June 2013.
- 94 Coublucq, Daniel, Demand Estimation with Selection Bias: A Dynamic Game Approach with an Application to the US Railroad Industry, June 2013.
- 93 Baumann, Florian and Friehe, Tim, Status Concerns as a Motive for Crime?, April 2013.
- 92 Jeitschko, Thomas D. and Zhang, Nanyun, Adverse Effects of Patent Pooling on Product Development and Commercialization, April 2013.  
Published in: The B. E. Journal of Theoretical Economics, 14 (1) (2014), Art. No. 2013-0038.
- 91 Baumann, Florian and Friehe, Tim, Private Protection Against Crime when Property Value is Private Information, April 2013.  
Published in: International Review of Law and Economics, 35 (2013), pp. 73-79.
- 90 Baumann, Florian and Friehe, Tim, Cheap Talk About the Detection Probability, April 2013.  
Published in: International Game Theory Review, 15 (2013), Art. No. 1350003.
- 89 Pagel, Beatrice and Wey, Christian, How to Counter Union Power? Equilibrium Mergers in International Oligopoly, April 2013.
- 88 Jovanovic, Dragan, Mergers, Managerial Incentives, and Efficiencies, April 2014 (First Version April 2013).
- 87 Heimeshoff, Ulrich and Klein Gordon J., Bargaining Power and Local Heroes, March 2013.
- 86 Bertschek, Irene, Cerquera, Daniel and Klein, Gordon J., More Bits – More Bucks? Measuring the Impact of Broadband Internet on Firm Performance, February 2013.  
Published in: Information Economics and Policy, 25 (2013), pp. 190-203.
- 85 Rasch, Alexander and Wenzel, Tobias, Piracy in a Two-Sided Software Market, February 2013.  
Published in: Journal of Economic Behavior & Organization, 88 (2013), pp. 78-89.
- 84 Bataille, Marc and Steinmetz, Alexander, Intermodal Competition on Some Routes in Transportation Networks: The Case of Inter Urban Buses and Railways, January 2013.
- 83 Haucap, Justus and Heimeshoff, Ulrich, Google, Facebook, Amazon, eBay: Is the Internet Driving Competition or Market Monopolization?, January 2013.  
Published in: International Economics and Economic Policy, 11 (2014), pp. 49-61.
- 82 Regner, Tobias and Riener, Gerhard, Voluntary Payments, Privacy and Social Pressure on the Internet: A Natural Field Experiment, December 2012.
- 81 Dertwinkel-Kalt, Markus and Wey, Christian, The Effects of Remedies on Merger Activity in Oligopoly, December 2012.

- 80 Baumann, Florian and Friehe, Tim, Optimal Damages Multipliers in Oligopolistic Markets, December 2012.
- 79 Duso, Tomaso, Röller, Lars-Hendrik and Seldeslachts, Jo, Collusion through Joint R&D: An Empirical Assessment, December 2012.  
Forthcoming in: The Review of Economics and Statistics.
- 78 Baumann, Florian and Heine, Klaus, Innovation, Tort Law, and Competition, December 2012.  
Published in: Journal of Institutional and Theoretical Economics, 169 (2013), pp. 703-719.
- 77 Coenen, Michael and Jovanovic, Dragan, Investment Behavior in a Constrained Dictator Game, November 2012.
- 76 Gu, Yiquan and Wenzel, Tobias, Strategic Obfuscation and Consumer Protection Policy in Financial Markets: Theory and Experimental Evidence, November 2012.  
Forthcoming in: Journal of Industrial Economics under the title "Strategic Obfuscation and Consumer Protection Policy".
- 75 Haucap, Justus, Heimeshoff, Ulrich and Jovanovic, Dragan, Competition in Germany's Minute Reserve Power Market: An Econometric Analysis, November 2012.  
Published in: The Energy Journal, 35 (2014), pp. 139-158.
- 74 Normann, Hans-Theo, Rösch, Jürgen and Schultz, Luis Manuel, Do Buyer Groups Facilitate Collusion?, November 2012.
- 73 Riener, Gerhard and Wiederhold, Simon, Heterogeneous Treatment Effects in Groups, November 2012.  
Published in: Economics Letters, 120 (2013), pp 408-412.
- 72 Berlemann, Michael and Haucap, Justus, Which Factors Drive the Decision to Boycott and Opt Out of Research Rankings? A Note, November 2012.
- 71 Muck, Johannes and Heimeshoff, Ulrich, First Mover Advantages in Mobile Telecommunications: Evidence from OECD Countries, October 2012.
- 70 Karaçuka, Mehmet, Çatik, A. Nazif and Haucap, Justus, Consumer Choice and Local Network Effects in Mobile Telecommunications in Turkey, October 2012.  
Published in: Telecommunications Policy, 37 (2013), pp. 334-344.
- 69 Clemens, Georg and Rau, Holger A., Rebels without a Clue? Experimental Evidence on Partial Cartels, April 2013 (First Version October 2012).
- 68 Regner, Tobias and Riener, Gerhard, Motivational Cherry Picking, September 2012.
- 67 Fonseca, Miguel A. and Normann, Hans-Theo, Excess Capacity and Pricing in Bertrand-Edgeworth Markets: Experimental Evidence, September 2012.  
Published in: Journal of Institutional and Theoretical Economics, 169 (2013), pp. 199-228.
- 66 Riener, Gerhard and Wiederhold, Simon, Team Building and Hidden Costs of Control, September 2012.
- 65 Fonseca, Miguel A. and Normann, Hans-Theo, Explicit vs. Tacit Collusion – The Impact of Communication in Oligopoly Experiments, August 2012.  
Published in: European Economic Review, 56 (2012), pp. 1759-1772.
- 64 Jovanovic, Dragan and Wey, Christian, An Equilibrium Analysis of Efficiency Gains from Mergers, July 2012.

- 63 Dewenter, Ralf, Jaschinski, Thomas and Kuchinke, Björn A., Hospital Market Concentration and Discrimination of Patients, July 2012 .  
Published in: Schmollers Jahrbuch, 133 (2013), pp. 345-374.
- 62 Von Schlippenbach, Vanessa and Teichmann, Isabel, The Strategic Use of Private Quality Standards in Food Supply Chains, May 2012.  
Published in: American Journal of Agricultural Economics, 94 (2012), pp. 1189-1201.
- 61 Sapi, Geza, Bargaining, Vertical Mergers and Entry, July 2012.
- 60 Jentzsch, Nicola, Sapi, Geza and Suleymanova, Irina, Targeted Pricing and Customer Data Sharing Among Rivals, July 2012.  
Published in: International Journal of Industrial Organization, 31 (2013), pp. 131-144.
- 59 Lambarraa, Fatima and Riener, Gerhard, On the Norms of Charitable Giving in Islam: A Field Experiment, June 2012.
- 58 Duso, Tomaso, Gugler, Klaus and Szücs, Florian, An Empirical Assessment of the 2004 EU Merger Policy Reform, June 2012.  
Published in: Economic Journal, 123 (2013), F596-F619.
- 57 Dewenter, Ralf and Heimeshoff, Ulrich, More Ads, More Revs? Is there a Media Bias in the Likelihood to be Reviewed?, June 2012.
- 56 Böckers, Veit, Heimeshoff, Ulrich and Müller Andrea, Pull-Forward Effects in the German Car Scrappage Scheme: A Time Series Approach, June 2012.
- 55 Kellner, Christian and Riener, Gerhard, The Effect of Ambiguity Aversion on Reward Scheme Choice, June 2012.
- 54 De Silva, Dakshina G., Kosmopoulou, Georgia, Pagel, Beatrice and Peeters, Ronald, The Impact of Timing on Bidding Behavior in Procurement Auctions of Contracts with Private Costs, June 2012.  
Published in: Review of Industrial Organization, 41 (2013), pp.321-343.
- 53 Benndorf, Volker and Rau, Holger A., Competition in the Workplace: An Experimental Investigation, May 2012.
- 52 Haucap, Justus and Klein, Gordon J., How Regulation Affects Network and Service Quality in Related Markets, May 2012.  
Published in: Economics Letters, 117 (2012), pp. 521-524.
- 51 Dewenter, Ralf and Heimeshoff, Ulrich, Less Pain at the Pump? The Effects of Regulatory Interventions in Retail Gasoline Markets, May 2012.
- 50 Böckers, Veit and Heimeshoff, Ulrich, The Extent of European Power Markets, April 2012.
- 49 Barth, Anne-Kathrin and Heimeshoff, Ulrich, How Large is the Magnitude of Fixed-Mobile Call Substitution? - Empirical Evidence from 16 European Countries, April 2012.  
Forthcoming in: Telecommunications Policy.
- 48 Herr, Annika and Suppliet, Moritz, Pharmaceutical Prices under Regulation: Tiered Co-payments and Reference Pricing in Germany, April 2012.
- 47 Haucap, Justus and Müller, Hans Christian, The Effects of Gasoline Price Regulations: Experimental Evidence, April 2012.

- 46 Stühmeier, Torben, Roaming and Investments in the Mobile Internet Market, March 2012.  
Published in: Telecommunications Policy, 36 (2012), pp. 595-607.
- 45 Graf, Julia, The Effects of Rebate Contracts on the Health Care System, March 2012,  
Forthcoming in: The European Journal of Health Economics.
- 44 Pagel, Beatrice and Wey, Christian, Unionization Structures in International Oligopoly,  
February 2012.  
Published in: Labour: Review of Labour Economics and Industrial Relations, 27 (2013),  
pp. 1-17.
- 43 Gu, Yiquan and Wenzel, Tobias, Price-Dependent Demand in Spatial Models,  
January 2012.  
Published in: B. E. Journal of Economic Analysis & Policy, 12 (2012), Article 6.
- 42 Barth, Anne-Kathrin and Heimeshoff, Ulrich, Does the Growth of Mobile Markets  
Cause the Demise of Fixed Networks? – Evidence from the European Union,  
January 2012.  
Forthcoming in: Telecommunications Policy.
- 41 Stühmeier, Torben and Wenzel, Tobias, Regulating Advertising in the Presence of  
Public Service Broadcasting, January 2012.  
Published in: Review of Network Economics, 11/2 (2012), Article 1.

Older discussion papers can be found online at:

<http://ideas.repec.org/s/zbw/dicedp.html>

**Heinrich-Heine-University of Düsseldorf**

**Düsseldorf Institute for  
Competition Economics (DICE)**

Universitätsstraße 1 \_ 40225 Düsseldorf  
[www.dice.hhu.de](http://www.dice.hhu.de)