



# IFRO Working Paper

## Understanding the Nature of Cooperation Variability

*Toke Fosgaard*  
*Lars Gårn Hansen*  
*Erik Wengström*

**IFRO Working Paper 2013 / 4**

Understanding the Nature of Cooperation Variability

Authors: Toke Fosgaard, Lars Gårn Hansen, Erik Wengström

Department of Food and Resource Economics (IFRO)

University of Copenhagen

Rolighedsvej 25

DK 1958 Frederiksberg DENMARK

[www.ifro.ku.dk](http://www.ifro.ku.dk) ( [www.foi.life.ku.dk](http://www.foi.life.ku.dk) )

# Understanding the Nature of Cooperation Variability

Toke Fosgaard<sup>\*</sup>, Lars Gårn Hansen<sup>\*</sup>, and Erik Wengström<sup>†‡</sup>

March 2013

## Abstract:

We investigate framing effects in a large-scale public good experiment. We measure indicators of explanations previously proposed in the literature, which when combined with the large sample, enable us to estimate a structural model of framing effects. The model captures potential causal effects and the behavioral heterogeneity of cooperation variability. We find that framing only has a small effect on the average level of cooperation but a substantial effect on behavioral heterogeneity and we show that this can be explained almost exclusively by a corresponding change in the heterogeneity of beliefs about other subjects' behavior. Preferences are on the other hand stable between frames.

Keywords: Framing, Public Goods, Internet Experiment, Simulation.

JEL codes: C13, C71, C93, H41

---

\* University of Copenhagen, Department of Food and Resource Economics

† University of Copenhagen, Department of Economics

‡ Lund University, Department of Economics

Acknowledgements: We gratefully acknowledge the generous funding provided by The Carlsberg Foundation. Wengström is grateful for financial support from the Wallander-Hedelius Foundation. We are deeply thankful to Jean-Robert Tyran for initiating and administrating the large-scale experiment upon which this paper is built. We thank Steffen Andersen, Giovanna Devetag, Tore Ellingsen, Peter Martinsson, and conference participants at the 6<sup>th</sup> and 7<sup>th</sup> Nordic Conferences on Behavioral and Experimental Economics and the ESA World Meeting in Chicago for their helpful comments.

## 1. Introduction

Cooperation between people is decisive for human welfare, but also a malleable endeavor in which the reasons for success or failure can be elusive. A stylized finding from economic experiments is that cooperation is situational in that people's behavior responds to economically irrelevant changes in the description of the decision situation (see for instance: Andreoni 1995a, Sonnemans et al. 1998, Park 2000, Cubitt et al. 2011a, Dufwenberg et al. 2011, Cubitt et al. 2011b, Fosgaard et al. 2011, Ellingsen et al. 2012, and Cappelen et al. 2013). Such insights might help to explain seemingly inconsistent phenomena such as why the same person can be very cooperative in teamwork at the workplace, but uncooperative when driving home from work.

An important puzzle in this literature is what drives the variability of cooperation. The exploration of this question has implications for several deep-rooted issues in economics. One such issue is whether human preferences are robust to contextual changes (Camerer and Thaler 2005, Levitt and List 2007), while another is the importance of changes in beliefs about other people's behavior (Nyarko and Schotter 2002, Dufwenberg and Battigalli 2009, Fischbacher and Gächter 2010). Finally, it also touches upon the relation between limited cognition and behavior (Köszegi and Rabin 2008a, Köszegi and Rabin 2008b, Bernheim and Rangel 2009, Choe et al. 2009). The connection between these issues and cooperation has been a lively research topic during recent decades. However, one limitation of the previous studies is however that they typically study one determinant at a time. Several studies have investigated to what extent preferences are context dependent (Brewer and Kramer 1986, McCusker and Carnevale 1995, Weber et al. 2004, Goerg and Walkowitz 2010, Iturbe-Ormaetxe et al. 2011). Another strand of the literature has focused on how context influences beliefs about others' behavior (Sonnemans et al. 1998, Dufwenberg et al. 2011 and Ellingsen et al. 2012), and while yet another set of studies have explored context-specific perceptions of the incentive structure (Ferrarro and Volsler 2010, Fosgaard et al. 2011).

Our goal is to evaluate the relative importance of *all of the determinants* previously documented as being important, *within the same study* to determine their relative importance. We do this by conducting a large-scale experiment which measures cooperation in public good games in two distinct, but economically equivalent, contexts. More specifically, we investigate

the effect of framing the cooperation decision as *taking* from a public good vs. *giving* to a public good (Andreoni 1995a). We measure the level of cooperation along with the main determinants: preferences, beliefs, and perception. With this data in hand, we are able to identify and estimate a structural model that allows us to decompose the framing effect into parts which are explained by each of the determinants.

Our estimation results bring new insights about the nature of framing effects. We find that changes in beliefs about others' behavior are the main determinant of framing effects. However, the belief effect is not visible in the summary statistics, which only indicate a small framing effect on mean contributions. Our estimation strategy reveals that the belief effect is counteracted by an equally sized unexplained effect. Hence, one important message of our study is that framing may substantially affect behavior even though it is not visible in the summary statistics of the final outcome. Moreover, we find that changes in cooperation preferences and incentive perceptions have negligible effects.

Another finding of our study is that the relatively small average framing effect masks substantial shifts in the underlying distribution; a finding that has not been reported previously in the literature, presumably because of limited sizes and compositions of the applied subject samples. More specifically, framing has a significant effect on the variance of cooperation levels and we find that essentially all of this effect can be explained by changes in beliefs.

The remainder of this study is organized as follows. The next section introduces our conceptual model. Section 3 then outlines our experimental design. In Section 4 the model is estimated and section 5 outlines how we disentangle the determinants. The decomposition of the determinants is presented in section 6 while section 7 concludes the paper with a discussion of our findings.

## **2. Our conceptual model**

The conceptual model that we use to guide our study is illustrated in figure 1. Its core is suggested by Fischbacher and Gächter (2010). We extend their model to accommodate misperception of incentives and framing.

When cooperating about the production of a public good, Fischbacher and Gächter argue that subjects formulate a *contribution strategy* based on their conditional *cooperation*

*preferences* (arrow 1 in the figure). The contribution strategy states the subjects' preferred contribution conditional on different levels of contributions made by other subjects<sup>1</sup>. Subjects then determine their actual *contribution* to the production of a public good by combining their *contribution strategy* with their *belief* about other subjects' contributions (arrows 3 and 5 in the figure).

We extend this core model because a number of studies have found that many subjects *misperceive* the incentives to contribute to the production of a public good (Andreoni 1995b and Houser and Kurzban 2002) and a recent study by Fosgaard *et al.* (2011) shows that framing can substantially affect the level of this misperception.

Logically, a subject's *misperception* of incentives may affect both the *contribution strategy* that he thinks reflects his preferences and the *beliefs* he has about other peoples' contributions (arrows 1 and 2 in Figure 1).

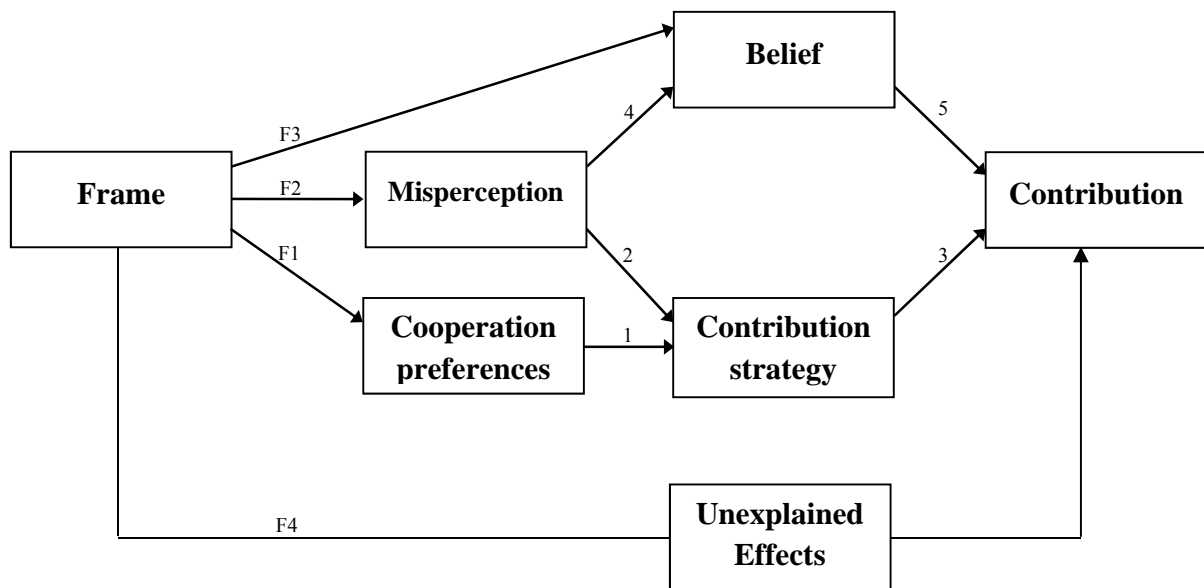


Figure 1: Framing and causal effects in our conceptual model

<sup>1</sup> The strategy indicates the subject's preferred contribution if others on average contribute nothing, if they contribute 1 dollar, etc.

Finally, our conceptual model allows for three possible gates through which framing effects can enter, which correspond to the three types of framing effect that have been suggested in prior experimental literature:

**a) Framing effects through cooperation preferences:** This effect reflects a shift in the subject's underlying preferences for cooperation caused by the change in framing and is captured by arrow F1. McCusker and Carnevale (1995) and Iturbe-Ormaetxe *et al.* (2011) argue that subjects have reference dependent utility and are loss averse, while van Dijk and Wilke (2000) suggest that subjects' 'focus' on personal and group outcomes may shift. If subjects have reference or 'focus' dependent utility functions (e.g. exhibit loss aversion) and these are affected by framing, we should find a significant effect (arrow 1) from framing on contribution strategies when controlling for changes in misperception.

**b) Framing effects through beliefs:** Sonnemans *et al.* (1998), Dufwenberg *et al.* (2011) and Ellingsen *et al.* (2012) have suggested that framing effects on peoples' beliefs are an important mechanism behind framing effects on contribution behavior. In our model, framing can directly affect beliefs (arrow F3) or this affect can go via misperception (arrow 4). Prior studies do not distinguish between these so if they turn out to be significant, it would be consistent with these studies.

**c) Framing effects through misperception:** Fosgaard *et al.* (2011) suggest that it is meaningful to distinguish between subjects' underlying cooperation preferences and subjects' *perception of the game* (e.g. the extent to which subjects correctly understand which contribution strategy maximizes their own income). The authors find that give/take framing has important effects on subjects' perception of the game and that this explains most of what looks like framing effects on underlying contribution preferences. Finding significant effects of misperception which work through subjects' contribution strategies would be consistent with this study (arrow 2). Here we also allow misperception to influence contributions via beliefs (arrow 4).

**d) Unexplained framing effects:** Finally, there could, in principal, be framing effects which work through other mechanisms than those proposed in the literature.

In conclusion, the model presented in Figure 1 incorporates elements which the literature has suggested are the core mechanisms through which framing affects contributions. As such, the model does not add any new elements, but rather it attempts to structure the existing elements and their interactions. One thing that the model does illustrate is the danger of investigating framing effects with an incomplete model. If key variables are omitted from the analysis, an estimated framing effect may become biased, because it may pick up effects which work through the omitted variables, and which are then *not* controlled for in the analysis.

The idea of our experiment (that we report on in the following section) is to generate sound indicators for the key variables in Figure 1, for a large subject pool that we randomly allocate to two different frames of the public good game. With this data, we then estimate the causal effects (the arrows) indicated in Figure 1, including the supplementary unexplained framing effect which captures framing effects that are not explained by our conceptual model.

### **3. Experimental design**

#### *General outline of the experiment*

Our artefactual field experiment was conducted over the Internet in the summer of 2008.<sup>2</sup> Initially, the Danish national bureau of statistics (Statistics Denmark) randomly selected 18,027 potential subjects from the Danish Central Person Register (CPR), which contains all individuals legally residing in Denmark. An invitation to participate in the experiment was sent to each selected potential subject via ordinary mail.<sup>3</sup> In the invitation letter, subjects were given the internet address of the experiment and a personal login code. The experiment was open for one week, and during this week, subjects could log in and out as they wished. After the experiment closed, subjects could log in again to receive feedback on the experimental results and type in their bank account number after which their earnings during the experiment were transferred to the account.

---

<sup>2</sup> See [http://www.econ.ku.dk/cee/iLEE/iLEE\\_home.htm](http://www.econ.ku.dk/cee/iLEE/iLEE_home.htm) for a detailed description of the experiment platform. The platform has been used for numerous studies on different topics; see, e.g. Thöni et al. (2012).

<sup>3</sup> See the Supplementary Information Document for a translation of the invitation letter and screenshots from the experiment.



After logging onto the webpage with their personal log-in code, subjects were randomly allocated to either the ‘give to the Public Good frame’ or to the ‘take from the Public Good frame’ (Andreoni 1995a, Cubitt *et al.* 2011a, 2011b, Dufwenberg *et al.* 2011, Park 2000). Subjects first played a standard one shot public good (PG) game and then played a ‘strategy version’ of the same game (see descriptions below). Subjects were then asked about their perception of game incentives. Finally, they were asked to complete a series of cognitive ability and personality trait tests and background questions. Only the PG games and the game perception questions were framed. The remaining tests were completely identical for all subjects.

A total of 2,042 subjects completed the experiment, with 1,366 subjects in the give treatment and 676 in the take treatment.<sup>4</sup> On average subjects earned 300 Danish kroner (DKK), approximately \$60, during the course of the entire experiment (including the payments from a risk elicitation task that we do not use). Subjects spent on average 50 minutes completing the experiment.

### *Details*

Initially, subjects played a **standard** PG game, in which subjects were randomly divided into groups of four. Each subject was given control of 50 DKK ( $\approx$ \$10) and was allowed to allocate this sum freely between a contribution to a common pool (i.e. the public good) or private income reserved for one self. In the **give** frame, subjects were initially given the 50 DKK as a private endowment, and were then asked what part of this endowment they wanted to contribute to a common pool. In the **take** frame, the 50 DKK was initially allocated to the common pool and subjects were then asked how much they wanted to withdraw from the common pool and instead reserve for themselves as private income. Under both frames, the money allocated to the common pool was doubled and shared equally among all group members. Hence, subjects’ earned the amount they reserved as private income plus a quarter of the final value of the contributions to the common pool. If subjects only care about personal income, we expect everyone to contribute 0, since the marginal private return of contributing 1 DKK to the common

---

<sup>4</sup> Two thirds of the participating subjects were allocated to the give treatment and the remaining one third were allocated to the take treatment. The reason for the uneven (but still random) allocation of subjects between the treatments is that the data from the give treatment have also been used for other papers, and hence these data received a higher priority.

pool is only ½ DKK. After the standard PG game, subjects were asked to state their beliefs about the average contribution of the other three group members. Subjects were rewarded for belief accuracy using the quadratic scoring rule.<sup>5</sup>

After completing the standard PG game, each subject played the **strategy version** of the PG game (strategy game) with the same framing. We applied the game developed by Fischbacher *et al.* (2001), where a profile of PG contributions conditional on different levels of contributions from other group members is elicited from each subject. Initially, subjects were divided into new groups of 4 and informed of this. Each subject was asked to make both an unconditional contribution and a profile of conditional allocations to the PG. First, the unconditional contribution was elicited in exactly the same way as in the previous standard PG. Subjects were then asked to indicate their contribution conditional on values of the three other group members' average contributions which varied from 0 to 50 DKK in steps of 5 DKK. When calculating payoffs, we used the elicited unconditional contributions for three randomly selected group members while the contribution of the fourth subject was based on the elicited conditional contribution profile using the average of the unconditional contributions from the other three group members.

Since contribution profiles are conditional on other group members' contributions, these profiles should not be affected by beliefs about the other group members' contributions. As shown by Fischbacher *et al.* (2001), the strategy method gives incentives to disclose precisely the conditional contribution profile, which underlies the unconditional contribution elicited in the standard PG game.<sup>6</sup>

Right after the strategy game, the subjects were asked incentivized control questions to test for **misperception**. We used the conditional setup introduced in the strategy game, and asked subjects to state the contribution profile they believed would maximize their personal income

---

<sup>5</sup> Subjects received an additional payment in DKK of  $10 - 0.004 d^2 \geq 0$ , where  $d$  is the difference between the belief and the true value.

<sup>6</sup> This assumes that the subjects assign non-degenerate probabilities to each of the given average contributions of the others. In case this is not true and the subjects assign a zero probability to one of the proposed contribution levels, the subject is indifferent with regards to the amount to state. In principle it could also be that contribution preferences do not depend on the average contributions, but rather on the distribution of contributions in the group or the strategy profiles of the others.

and the contribution profile they believed would maximize other group members' income. The test consisted of six questions in all. It was emphasized that each question only had one correct answer and that subjects earned 5 DKK for each correct answer. The first three questions asked subjects what public good contribution they should make if they wanted to maximize their own income, when the others, on average, contributed 0 DKK (question 1), 25 DKK (question 2) and 50 DKK (question 3). In the last three questions, subjects were asked what contribution they should make if they wanted to maximize the income of the group, when the others, on average, contributed 0 DKK (question 4), 25 DKK (question 5) and 50 DKK (question 6). We interpret incorrect answers to these questions as misperceptions about how to implement the specified goals in the public good game.

#### **4. Estimating a model for decomposing framing effects**

In this section, we specify and estimate the structural model of framing, developed in Section 2, using our experimental data. In Section 6, we then use this model to decompose the total framing effect on contributions measured in our experiment into parts working through misperception, beliefs and preferences. This allows us to quantify the relative importance of these three different possible mechanisms through which framing could work and to quantify the importance of any remaining unexplained framing effect on contributions.

In the experiment, we measure the exogenous framing variable as well as 4 out of the 5 endogenous variables (misperception, contribution strategies, beliefs and contributions) introduced in the conceptual model (See the document with Supplementary Information for a detailed data description). We are not able to elicit cooperation preferences directly and so we are not able to estimate the framing effect on cooperation preferences directly.

Although we do not observe cooperation preferences, we can nevertheless estimate the framing effect which works through them (arrow F1 in Figure 1). Our conceptual model implies that any framing effect on contribution strategies, which does not go via misperception (which we observe), must go through cooperation preferences. By including a direct framing effect on contribution strategies in our estimations, we can therefore (under the assumptions of our conceptual model) interpret this as the effect of framing which works through cooperation preferences (and hence arrow F1 is solid in Figure 1).

Finally, we do not measure an indicator of unexplained framing effects directly. However, the theoretically blocked direct causal links between framing and contributions provide us with a test of our conceptual model. If we do find a significant direct effect of framing on contributions (effect F4 in Figure 1) it will indicate that the conceptual model is incomplete and that there are important unexplained mechanisms which transmit the effect of framing into contributions. We therefore include this (conceptual blocked) direct effect in the empirical model.

We also include a number of exogenous controls including cognitive ability, big five personality traits, gender, and age as explanatory variables for the estimation of all four endogenous variables.<sup>7</sup>

Formally, the empirical specification of our model is the following recursive system:

$$\begin{aligned}
 \text{Misperception (pe):} & \quad pe = g1(fr_{F2}, x, e1) \\
 \text{Contribution strategy (cs):} & \quad cs = g2(pe_2, fr_{F1}, x, e2) \\
 \text{Belief (bl):} & \quad bl = g3(pe_4, fr_{F3}, x, e3) \\
 \text{Contribution(cn):} & \quad cn = g4(bl_4, cs_3, fr_{F4}, x, e4)
 \end{aligned}$$

Where  $fr$  is a frame dummy,  $x$  is a vector of exogenous control variables and  $e1, \dots, e4$  are stochastic variables which capture the effects of unobserved exogenous variables. The subscripts to the framing indicator and other endogenous explanatory variables indicate the corresponding arrow in Figure 1. Our extensive battery of control variables includes potentially highly important causes of correlation between equations such as cognitive abilities and personality traits. When estimating the system, we assume that the stochastic variables ( $e1$ ,  $e2$ ,  $e3$  and  $e4$ ) are independent.

In the estimations, misperception ( $g1$ ) is modeled as a binary probit. Beliefs ( $g3$ ) are modeled as a multi-nominal probit.<sup>8</sup> Contribution strategies ( $g2$ ) are modeled as a two-step

---

<sup>7</sup> See the Supplementary Information Document for a full description of the exogenous variables.

process where a subject's strategy type follows a multi-nominal probit and the specific profile is selected randomly among the set of observed profiles for subjects of this strategy type and with the subject's set of explanatory variables. Finally, contributions ( $g4$ ) are modeled as a multi-nominal probit<sup>9</sup> where we followed Fischbacher and Gächter (2010) and first combined belief ( $bl$ ) and contribution strategy ( $cs$ ) to generate the subjects' preferred contribution ( $pc$ ), and then used both preferred contribution and beliefs as explanatory variables in the estimation of the actual contributions. That is, by including both preferred contributions (which depend on beliefs) and beliefs in the final estimation, we allow beliefs to affect contributions beyond their role in generating the preferred contribution. Fischbacher and Gächter (2010) find clear evidence for this, and we want our empirical model to follow theirs. Thus, this equation becomes:

$$\text{Contribution: } \quad cn = g4(pc, bl_4, fr_{F4}, x, e4) \quad \text{where } pc=f(bl_4, cs_3)$$

The preferred contribution,  $pc=f(bl_4, cs_3)$ , is found as the contribution indicated in the subject's contribution strategy that corresponds to his stated belief about the contributions of others.

We include first-order effects of controls and both first and second-order effects of endogenous explanatory variables (i.e. we include squared endogenous variables as regressors in all probit models). These specifications are quite flexible and do not impose restrictive functional relationships on the key endogenous variables. Further, in addition to the linear functional form assumed by Fischbacher and Gächter (2010) for the contribution equation, a large class of other functional forms is accommodated. A detailed description of the empirical specification of the four equations is found in the Supplementary Information Document, Section 3.

## 5. Simulating framing effects using the estimated model

Because we have access to a large dataset, we are able to estimate our structural model of the underlying relationships without restrictive assumptions about the functional form. This makes it possible for us to use the model to simulate behavior with reasonable precision.

---

<sup>8</sup> Beliefs are in practice categorized since almost all subjects report beliefs that are divisible by 5 and with some values attracting a large proportion of reported beliefs. The few observations that were not a multiple of 5 (around 2% of the observations) were rounded to the nearest 5 kroner.

<sup>9</sup> Like beliefs, contributions were in practice rounded to the nearest multiple of 5.

When simulating, we first use the estimated parameters for the misperception equation  $g1(\cdot)$  to calculate the probability of misperceiving, when a subject with a specific set of personality traits, cognitive abilities, gender, and age is placed in a given frame. For our simulations, we then select the specific realization of the misperception variable through a random draw between the two possible realizations which reflect their calculated probabilities. Using the simulated realization of the subject's misperception variable, we then simulate the subject's belief about others' contributions. Using the estimated parameters for the belief equation  $g3(\cdot)$ , we calculate the probability of each possible value of this belief for a subject with the given combination of personal characteristics, frame and the simulated value of the misperception variable. The realized belief is again chosen through a random draw among the possible realizations which reflect the calculated probabilities. The subject's contribution strategy is simulated in basically the same way. First, the subject's strategy type is simulated using the estimated multinomial probit parameters and a random draw in the same way that beliefs are simulated. Then the specific profile is selected randomly among the set of observed profiles for subjects of this strategy type and with the same frame and the simulated value of the misperception variable. Finally, the simulated contribution strategy and belief are combined to find the preferred contribution, which together with the framing variable, beliefs and controls are used to simulate the subject's contribution (for more details on the simulation procedure see the Supplementary Information Document, Section 4).

Before using the estimated structural model to simulate how subjects react to a change in frame (in the following section), we tested the estimated model's ability to reproduce the original experimental data in a number of ways. We did this by comparing the distributions of contributions in the experiment, with the distribution simulated by our model. A comprehensive description of the tests and results can be found in the Supplementary Information Document, Section 5. We do not find significant differences between simulated and original distributions in any of our tests and we conclude that our model captures the systematic variation in our data well.

## **6. Disentangling framing effects**

In this section, we decompose the observed framing effect on contributions into framing effects which work through misperception, preferences, beliefs and remaining unexplained framing

effects. We do this by simulating the change in contributions for subjects who are moved from the give frame into the take frame. By changing the frame variable for one equation at a time, we are able to simulate the marginal effect on contributions for each of the suggested paths in Figure 1.

Initially we simulate the *total framing effect* on contributions for our sample. Using the estimated coefficients of our system of equations, we first simulate the contributions for all subjects in the give frame (i.e. all frame variables set to give and all endogenous variables which enter equations are those simulated for the give frame). We then simulate the corresponding distribution of contributions when all subjects are ‘placed’ in the take frame (i.e. all frame variables set to take and all endogenous variables which enter equations are simulated for the take frame). The total simulated framing effect on contributions is then the difference between the simulated distributions for the two frames (note that both simulated distributions use all subjects). In the following, we focus on how moving from the take to the give frame affects the mean and the standard deviation of the distribution of contributions.<sup>10</sup> This simulated total framing effect on the mean of contributions is shown in Figure 6 (right bar), while the simulated total framing effect on the standard deviation of contributions is shown in Figure 7 (right bar). The simulated effects correspond almost exactly to the observed framing effects in our data (see the Supplementary Information Document, Section 1). The indicated 5% confidence intervals around the simulated effects are larger than for the observed effects (see the Supplementary Information Document, Section 1) because these take the estimation uncertainty of our model parameters into account.<sup>11</sup>

We now *decompose the total framing effect* into its four parts using the simulation model to generate marginal effects. Starting with all subjects in the take frame, we only let the framing affect one of the underlying mechanisms at a time. The four partial effects that we simulate in this way are:

---

<sup>10</sup> Framing also significantly effects higher order distribution moments but the quantitatively most important effects are on the first two ditribution moments.

<sup>11</sup> Standard errors for simulated effects illustrated in figures 6 and 7 are found by the standard bootstrapping procedure (randomly drawing subjects from the original subject pool with replacement and re-estimating and resimulating 250 times. See the Supplementary Information Document for details) .

Misperception ( $F2 \rightarrow 4 \rightarrow 5$ ,  $F2 \rightarrow 2 \rightarrow 3$ )

Beliefs ( $F3 \rightarrow 5$ )

Preferences ( $F1 \rightarrow 3$ )

Unexplained framing effect ( $F4$ )

The *misperception* mechanism is made up of two sub-effects. First, misperception can affect contribution strategies, which then carry over to contributions ( $F2 \rightarrow 2 \rightarrow 3$ ). Second, misperception can affect beliefs ( $F2 \rightarrow 4 \rightarrow 5$ ). The *belief* mechanism explains the part of the total framing effect on the distribution of contributions which works directly through beliefs ( $F3 \rightarrow 5$ ), but excludes the part of the change in beliefs that is caused by a framing effect on misperception ( $F2 \rightarrow 4 \rightarrow 5$ ). The *preference* mechanism explains the part of the total framing effect which is caused by a framing induced change in basic preferences for cooperation ( $F1 \rightarrow 3$ ). Finally, the *unexplained* effect captures the part of the total framing effect that is not explained by the three modeled causal mechanisms. In Figure 2, we show the simulated partial framing effects on mean distributions and in Figure 3 the corresponding partial effects on the contribution distributions standard deviation.



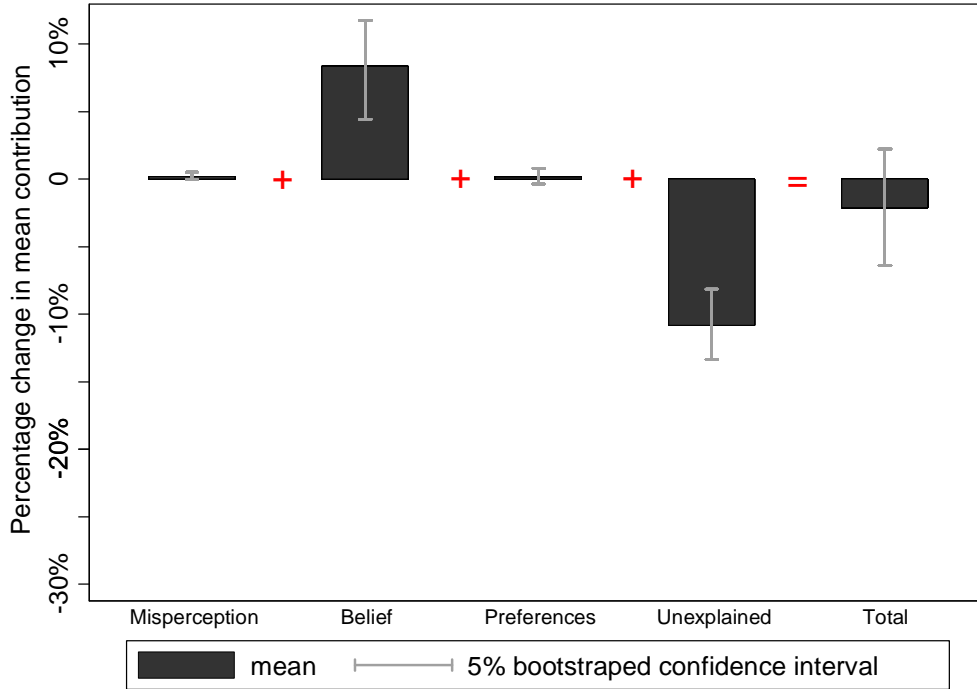


Figure 2: Simulated effects on mean contribution when moving all subjects from the take frame to the give frame

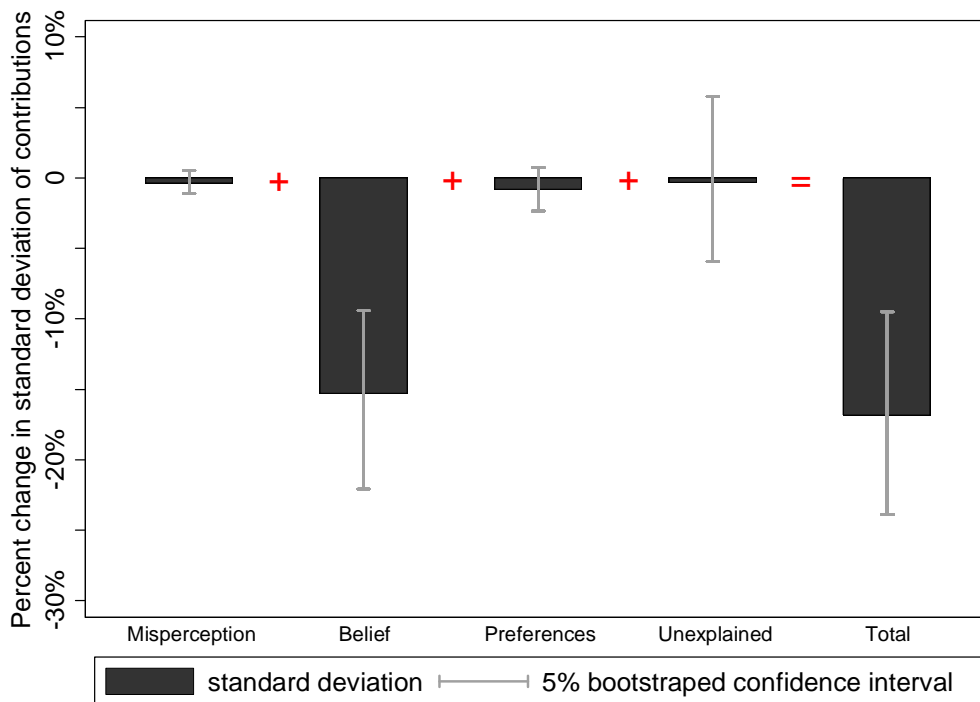


Figure 3: Simulated effects on the standard deviation of contributions when moving all subjects from the take frame to the give frame

Looking at the total framing effects in Figure 2 and 3, the first interesting observation is that a relatively small framing effect on mean contributions masks a relatively large framing effect on the standard deviation of the distributions. Thus, even though mean contributions do not change much; the subjects' contribution behavior is in fact affected substantially by framing.

Looking at the decomposed effects in Figure 2 and 3, we see that essentially no framing effects work through misperception or preferences. The estimated effects are in both cases close to zero and standard errors are quite tight around this. On the other hand, beliefs explain a substantial and significant part of the total framing effects. Thus, it seems clear that there are substantial framing effects on subjects working through beliefs, while the effects which work through misperception and preferences are negligible in comparison. This is not obvious from the outset since there are substantial framing effects on misperception and on elicited contribution strategies (see the Supplementary Information Document, section 1). In addition, Fosgaard *et al.* (2011) show that without controlling for misperceptions, more subjects give low

contributions in the take frame, which is consistent with more subjects understanding how to maximize own income in the take frame. Furthermore, they show that, when controlling for misperceptions, this framing effect is removed. Our study reveals that the importance of this mechanism for explaining the change in average contributions is small compared to the importance of the belief mechanism.

From Figure 3 it seems clear that framing effects on beliefs can explain virtually the entire framing effect on the standard deviation of the contribution distribution. The remaining unexplained effect is small and insignificant (though the confidence interval around this estimate is relatively large). Thus, the substantial reduction in belief variances in our data can explain the reduction in contribution variance when we control for other possible explanations.

Finally, looking at Figure 2, we see that though the net framing effect is small, there is a substantial framing effect through beliefs, which drives mean contributions up and a substantial unexplained framing effect, which drives mean contributions down. Thus, the increase in mean beliefs that we see as we move to the give frame has the expected effect of increasing contributions. However, a stronger *unexplained* framing effect counteracts this and ends up causing a (small) net reduction in mean contributions.

Comparing with earlier literature, our results are in line with recent studies that find important framing effects which work through beliefs (Dufwenberg *et al.* (2011), Ellingsen *et al.* (2012), and Sonnemans *et al.* (1998)). We add to this in several ways. First, we provide sound evidence that indicates that the two other proposed explanations (framing induced changes in preferences and misperception) are negligible in comparison. Our results also show that framing which works through beliefs has an important effect on the variance of the contribution distribution in addition to the effect mean contributions (which has been the sole focus of earlier investigations). Finally, we find that, in addition to beliefs, there are important unexplained framing effects on mean contributions.

These unexplained effects could be caused by incorrect specifications of the mechanisms we have modeled. Though our functional forms are flexible, in general, it is possible that, for example, the process of generating preferred contributions by combining beliefs with the strategy profile is too simple. We assume that subjects hold point beliefs about others' contributions and

choose the contribution in their profile that corresponds to this belief. In reality, few subjects are certain about what others will do and, hence, it may be more appropriate to ask subjects how they believe the contributions of others will be distributed. Using belief distributions, rather than point beliefs, would open up new ways of understanding our results. For example, if subjects feel worse about giving too little than giving too much compared to others, increased belief uncertainty could make them contribute more (to insure against giving less than others). In our case, subjects may become more certain about their beliefs as they are moved to the give frame (the reduction in contribution variance suggests this). If so, this would reduce the expected cost of giving too little and subjects would reduce their contributions. This effect could outweigh the effect of the increase in expected beliefs. Thus, the unexplained negative effect on mean contributions that we identify could be caused by subjects being more certain about what others contribute in the give frame, and thereby which action is appropriate. However, this is just speculation.<sup>12</sup> The unexplained framing effect could also be caused by mechanisms outside our model transferring framing effects to contributions. At any rate, the fact that we find two strong underlying framing effects with opposite signs fits well with the mixed results on the direction of the net framing effect in the literature. When a relatively small net effect is caused by two large opposite underlying effects of the same magnitude, small changes in setting, subject pool, etc. can cause the sign of the net effect to change.

## 7. Conclusions

In our study, we investigate the effect of changing the default state in a game of public good production. We measure indicators of the three main explanations proposed in the literature and we use a substantially larger subject pool for our experiment than prior studies. This makes it possible for us to estimate a structural model which captures all three causal effects and the behavioral heterogeneity of our subject pool. We find that framing has only a small effect on the average level of cooperation. However, we find a substantial effect of this change on behavioral

---

<sup>12</sup> One way to view this is that, in the give treatment, there is less uncertainty regarding the social norms governing the choice situation. This is consistent with people spending more effort on understanding the game in the take frame which was argued in Fosgard et. al (2011). Regarding the effects of norm uncertainty and framing, Dreber *et al.* (2012) argue that higher norm uncertainty in the generalized ultimatum game (in which the responder can only reduce the proposer's payoff by a small fraction) compared to the dictator game (in which the responder cannot affect the proposer's payoff at all) may explain why framing matters in the former game but not in the latter. Their reasoning is that framing is less influential in situations with a clearly established norm such as in the dictator game.

heterogeneity across our subjects: cooperative behavior is much more *heterogeneous* in some settings than in others and we show that this can be explained almost exclusively by a corresponding change in the heterogeneity of beliefs about other subjects' cooperative behavior. Our results suggest that it is easier for subjects to identify the behavioral norm in some settings than in others and that *this* is the main explanation for the variability in cooperation that we see in prior studies and in the real world. Preferences are on the other hand stable across contextual changes in our experiment.

Our results suggest that cooperation preferences are stable over a wider range of presentation and context changes which implies that underlying incentives are of primary importance for the long-term success of cooperation. The stability of preferences across different settings should also imply a corresponding stability of behavioral norms. However, presentation and context appear to be very important for how easy it is for people to *identify* the behavioral norm in a given setting.

Thus, our results suggest a new understanding of how context and presentation affect the success or failure of cooperation. Even if the mean level of cooperation is the same across different settings, our results imply that the *variance* in outcomes for initial rounds of the same basic cooperative venture may be substantially larger for some settings than for others. Because of the dynamics of cooperative processes that depend on conditional preferences, this can be critical. Fundamentally, sound ventures may fail in the initial rounds of cooperation, or never get started, if many (potential) participants in a given venture have overly pessimistic initial expectations of others' behavior. A greater variance in expectations will cause a larger proportion of cooperative ventures in this setting to fail in the critical upstart phase because of overly pessimistic initial expectations. Initially, this is compensated by high cooperation levels in other ventures where participants have overly optimistic expectations. But as the venture continues and expectations are updated, cooperation will stabilize at a level which reflects the underlying preferences. In the long run, those ventures that have survived the critical initial rounds will end up cooperating at the same level in both settings (reflecting cooperation preferences and the actual behavior of others). However, fewer ventures will survive in the setting in which people initially found it more difficult to identify the behavioral norm.

## 8. References

- Andreoni, J. (1995a). 'Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments', *Quarterly Journal of Economics*, vol. 110, pp. 1-21
- Andreoni, James. (1995b). 'Cooperation in Public-Goods Experiments: Kindness or Confusion?' *American Economic Review*, 85(4): 891-904.
- Battigalli, P. and Dufwenberg M. (2009) 'Dynamic Psychological Games', *Journal of Economic Theory* 144 (2009), pp. 1-35.
- Bernheim, B.D., Rangel, A. (2009) 'Beyond Revealed Preference: Choice Theoretic Foundations for Behavioral Welfare Economics'. *Quarterly Journal of Economics*, vol. 124(1), pp. 51-104.
- Brewer, M. B., and Kramer, R. M. (1986) 'Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing', *Journal of Personality and Social Psychology*, vol. 50(3), pp. 543-549.
- Camerer C, and Thaler R.H. (1995) Anomalies: Ultimatums, Dictators and Manners. *The Journal of Economic Perspectives*. 9(2), pp.209-219
- Cappelen, A. W., Nielsen, U. H., Sørensen, E.Ø., Tungodden, B., and Tyran, J. (2013). 'Give and take in dictator games. *Economics Letters*, vol. 118(2), pp. 280–283
- Chou, E., McConnell, M., Nagel, R. and Plott, C. (2009). 'The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person "guessing" game'. *Experimental Economics*, vol 12, pp.159–179.
- Cubitt, R., Drouvelis, M. and Gächter, S (2011a). 'Framing and Free Riding: Emotional Responses and Punishment in Social Dilemma Games', *Experimental Economics*, vol. 14, pp. 254- 272.
- Cubitt, R., Drouvelis, M., Gächter, S and Kabalin, R. (2011b). 'Moral Judgments in Social Dilemmas: How Bad is Free Riding?' *Journal of Public Economics*, vol. 95, pp. 253-264.
- Dreber, A., Ellingsen, T., Johannesson, M. and Rand, D (2012). 'Do people care about social context? Framing effects in dictator games', *Experimental Economics*, forthcoming
- Dufwenberg, M, Gächter, S. and Hennig-Schmidt, H. (2011), 'The Framing of Games and the Psychology of Play', *Games and Economic Behavior* 73, 459-78.
- Ellingsen, T., Johannesson, M., Mollerstrom, J. and Munkhammar, S. (2012). 'Social framing effects: preferences or beliefs?' *Games and Economic Behavior*, in press.

- Ferraro, P. J., and Vossler, C. A. (2010). 'The Source and Significance of Confusion in Public Goods Experiments', *The B.E. Journal of Economic Analysis & Policy*, 10(1): 1-40.
- Fischbacher, U., Gächter, S. and Fehr, E. (2001). 'Are People Conditionally Cooperative? Evidence from Public Goods Experiment', *Economics Letters*, vol. 71, pp. 397-404.
- Fischbacher, U. and Gächter, S. (2010). 'Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments', *American Economic Review*, vol. 100, pp. 541-56.
- Fosgaard, T., Hansen, L. G., and Wengström, E. (2011) 'Framing and misperception in a public good experiment', Institute for Food and Resource Economics, Working paper no. 2011/11.
- Goerg, S.J. and Walkowitz, G (2010) 'On the prevalence of framing effects across subject-pools in a two-person cooperation game'. *Journal of Economic Psychology*, vol. 31 (6), pp. 849–859.
- Houser, D. and Kurzban, R.. (2002). 'Revisiting Kindness and Confusion in Public goods Experiments', *American Economic Review*, 92(4): 1062-69.
- Iturbe-Ormaetxe , I., Ponti, G., Tomás, J. and Ubeda, L. (2011). 'Framing Effects in Public Goods: Prospect Theory and Experimental Evidence', *Games and Economic Behavior*, vol 72, pp.439-447.
- Köszegi, B., Rabin, M. (2008a). 'Choices, Situations, and Happiness', *Journal of Public Economics*, 92(8-9), pp. 1821-32.
- Köszegi and Rabin. (2008b). Revealed mistakes and revealed preferences. In *The foundations of positive and normative economics* (eds) Caplin, A. & Schotter, A. Oxford university press. New York.
- Levitt, S. and List, J. (2007). 'What do laboratory experiments measuring social preferences reveal about the real world?' *Journal of Economic Perspectives*, 21(2), pp. 151-174.
- McCusker, C. and Carnevale, P. J. (1995). 'Framing in Resource Dilemmas: Loss Aversion and the moderating Effect of Sanctions', *Organizational Behavior and Human Decision Processes*, vol. 61, pp. 190-201.
- Nyarko, Y., and Schotter, A. (2002). An experimental of belief learning using elicited beliefs, *Econometrica*, Vol. 70 (3), pp. 971–1005
- Park, E. (2000). 'Warm-glow versus Cold-prickle: a further Experimental Study of Framing Effects on Free-riding', *Journal of Economic Behavior and Organization*, vol. 43, pp. 405-21.
- Sonnemans, J., Schram, A. and Offerman, T. (1998). 'Public Good Provision and Public Bad Prevention: The Effect of Framing', *Journal of Economic Behaviour & Organization*, vol. 34, pp. 143-61.

Thöni C., Tyran, J.-R. and Wengström E. (2012). 'Microfoundations of social capital', *Journal of Public Economics*, Vol. 96 (7–8), pp. 635-643.

Van Dijk, E. and Wilke, H. (2000). 'Decision-Induced Focusing in Social Dilemmas: Give-Some, Keep-Some, Take-Some, and Leave-Some Dilemmas', *Journal of Personality and Social Psychology*, vol. 78, pp. 92-104.

Weber, J.M., Kopelman, J., and Messick, D. M. (2004). 'A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness' *Personality and Social Psychology Review*, vol. 8(3), pp. 281-307.