American Economic Review 2010, 100:1, 541–556
http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.1.541

Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments

By Urs Fischbacher and Simon Gächter*

In this paper we investigate the role of social preferences and beliefs for voluntary cooperation. Numerous public goods experiments have shown that many people contribute more to the public good than pure self-interest can easily explain. An equally important observation, however, is that free riding increases in repeatedly played public goods experiments across various parameters and participant pools.[1] The facts are clear, but their explanation is not.

An obvious candidate for explaining the decline of cooperation was learning the free-rider strategy. However, Andreoni (1988) showed that cooperation resumed after a restart, which is inconsistent with a pure learning argument. In a subsequent paper he interpreted the decline in cooperation "to be due to frustrated attempts at kindness, rather than learning the free-riding incentives" (Andreoni 1995, 900). Several subsequent papers argue that contributions that are not due to confusion might possibly be explained by other-regarding preferences (e.g., Thomas R. Palfrey and Jeffrey E. Prisbrey 1997; Jordi Brandts and Schram 2001; Jacob K. Goeree, Charles A. Holt, and Laury 2002; Daniel Houser and Robert Kurzban 2002; Paul J. Ferraro and Christian A. Vossler 2005; Ralph C. Bayer, Renner, and Rupert Sausgruber 2009). One type of social preference—long argued by social psychologists (e.g., Harold Kelley and Anthony Stahelski 1970)—is many people's propensity to cooperate (in lab and field environments) provided others cooperate as well (e.g., Robert Sugden 1984; Joel Guttman 1986; Keser and van Winden 2000; Fischbacher, Gächter, and Fehr 2001; Bruno S. Frey and Stephan Meier 2004; Croson, Fatas, and Neugebauer 2005; Croson 2007; Richard Ashley, Cheryl Ball, and Catherine Eckel forthcoming; Croson and Jen Shang 2008; see Gächter 2007 for an overview). Such "conditional cooperation"

* Fischbacher: University of Konstanz, PO Box D 131, D-78457 Konstanz, Germany (e-mail: urs.fischbacher@uni-konstanz.de), and Thurgau Institute of Economics, Hauptstrasse 90, CH-8280, Kreuzlingen (e-mail: fischbacher@twi-kreuzlingen.ch); Gächter: CESifo, IZA, and University of Nottingham, School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, United Kingdom (e-mail: simon.gaechter@nottingham.ac.uk). This paper is part of the MacArthur Foundation Network on Economic Environments and the Evolution of Individual Preferences and Social Norms and the Research Priority Program of the University of Zurich on "The Foundations of Human Social Behavior: Altruism versus Egoism." Support from the EU-TMR Research Network ENDEAR (FMRX-CT98-0238) is gratefully acknowledged. Sybille Gübeli, Eva Poen, and Beatrice Zanella provided very able research assistance. We also thank very constructive referees and Nick Bardsley, Rachel Croson, Armin Falk, Ernst Fehr, Justina Fischer, Martin Kocher, Michael Kosfeld, Andreas Kuhn, Jo Morgan, Charles Noussair, Eva Poen, Daniel Schunk, Martin Sefton, Ran Spiegler, and various conference, workshop, and seminar participants for their helpful comments. Gächter thanks the University of St. Gallen, CES Munich, CMPO Bristol, the University of Copenhagen, and Bar-Ilan University for the hospitality he enjoyed while working on this paper.

[1] See, for instance, R. Mark Isaac, James M. Walker, and Susan H. Thomas (1984); James Andreoni (1988, 1995); Joachim Weimann (1994); Susan K. Laury, Walker, and Arlington Williams (1995); Rachel Croson (1996); Roberto Burlando and John Hey (1997); Gächter and Ernst Fehr (1999); Axel Ockenfels and Weimann (1999); Joep Sonnemans, Arthur Schram, and Theo Offerman (1999); Claudia Keser and Frans van Winden (2000); Fehr and Gächter (2000); Eun-Soo Park (2000); David Masclet et al. (2003); Croson, Enrique Fatas, and Tibor Neugebauer (2005); Jeffrey P. Carpenter (2007); Martin Sefton, Robert Shupp, and Walker (2007); Martijn Egas and Arno Riedl (2008); Gächter, Elke Renner, and Sefton (2008); Benedikt Herrmann, Christian Thöni, and Gächter (2008); Nikos Nikiforakis and Hans Normann (2008); and Neugebauer et al. (2009). The decline of cooperation has also been observed in prisoner's dilemma experiments. See, e.g., Reinhard Selten and Rolf Stoecker (1986); Andreoni and John H. Miller (1993) and Russell Cooper et al. (1996). The breakdown of cooperation is also a frequent outcome in naturally occurring situations, like in the overdepletion of common resources and in environmental degradation.

is an interesting candidate for explaining the fragility of cooperation, because this motivation depends directly on how others behave or are believed to behave. Conditional cooperators who observe (or believe) that others free ride will reduce their contributions and thus contribute to the decline of cooperation. It is unknown, however, to what extent conditional cooperation, and inter-individual differences in this regard, can explain (the decline of) cooperation. Our paper aims to shed empirical light on this question. For a related theoretical approach see Attila Ambrus and Parag A. Pathak (2009).

The novelty of our paper is to combine two observations from previous research: beliefs about other people's contributions matter, and people differ in their cooperative preferences (some are free rider types, whereas others are conditional cooperators).[2] In our approach we measure people's cooperation preferences in a specially designed public goods game played in the strategy method (called the "P-experiment") and then observe the same people in a sequence of ten one-shot games (labeled the "C-experiment"), in which we also elicit their beliefs about others' contributions. This allows us to quantify how preference heterogeneity and beliefs interact in voluntary cooperation. Specifically, we can disentangle whether contributions decline because of cooperation preferences and/or because of the way people form (and change) their beliefs about how others will behave.

Our data from the P-experiment show that people differ strongly in their contribution preferences. This is consistent with previous evidence (see footnote 2). Using the classification proposed by Fischbacher, Gächter, and Fehr (2001), we have (i) 55 percent conditional cooperators who cooperate if others cooperate, (ii) 23 percent free riders who never contribute anything, irrespective of how much others contribute, (iii) 12 percent "triangle contributors" who increase their contributions with the contribution of others up to a point and then decrease their own contributions the more others contribute, and (iv) 10 percent unclassifiable. We push beyond this observation of preference heterogeneity by investigating how measured preferences and beliefs are related to observed contribution behavior. We have therefore designed our experiments such that we can use the P-experiment to make a point prediction for each participant about his or her contribution in the C-experiment, given his or her beliefs. Our approach allows us to answer our main research question—how do beliefs and preference heterogeneity affect the decline of cooperation?

Our main result, which we detail in Section II, is that contributions decline because, on average, people are "imperfect conditional cooperators" who match others' contributions only partly. The presence of free-rider types is not necessary for this result; contributions also decline if everyone is an imperfect conditional cooperator. We further show that belief formation can be described as a partial adjustment of one's belief into the direction of the observed contribution of others in the previous period. More specifically, beliefs in a given period are a weighted average of what others contributed in the previous period and one's own belief in the previous period. As we will show with the help of simulation methods, our estimated belief formation process implies that beliefs decline only if contributions decline, but not vice versa. Furthermore, we find that contributions are significantly positively correlated with predicted contributions, that is, the elicited preferences. In addition to their preferences, people's contributions depend directly on their beliefs about others' contributions. This implies that the P-experiment understates the extent of conditional cooperation that occurs in the C-experiment.

[2] See, e.g., Fischbacher, Gächter, and Fehr (2001); Burlando and Francesco Guala (2005); Kurzban and Houser (2005); Nicholas Bardsley and Peter Moffatt (2007); Martin G. Kocher et al. (2008); Laurent Muller et al. (2008); John Duffy and Jack Ochs (2009); and Herrmann and Thöni (2009).

I. Design and Procedures

Our basic decision situation is a standard linear public goods game. The participants are randomly assigned to groups of four persons. Each participant is endowed with 20 tokens, which he or she can either keep or contribute to a "project," the public good. The payoff function is given as

$$\pi_i = 20 - g_i + 0.4 \sum_{j=1}^{4} g_j, \tag{1}$$

where the public good is equal to the sum of the contributions of all group members. Contributing a token to the public good yields a private marginal return of 0.4 and the social marginal benefit is 1.6. Standard assumptions, therefore, predict that all group members free ride completely, that is, $g_j = 0$ for all j. This leads to a socially inefficient outcome.

The instructions (see the Web Appendix, available at http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.1.541) explained the public good problem to the participants. Since we want to measure subjects' preferences as accurately as possible, we also took great care to ensure that the participants understood both the rules of the game and the incentives. Therefore, after participants had read the instructions, they had to answer ten control questions. The questions tested the subjects' understanding of the comparative statics properties of (1) to ensure that participants were aware of their money-maximizing incentives and the dilemma situation. We did not proceed until all participants had answered all questions correctly. We can thus safely assume that the participants understood the game.

Within this basic setup we conducted two types of experiments. The first type of experiment (the P-experiment) elicits people's contribution *preferences* in a linear one-shot public goods game. In the second type of experiment participants make *contribution choices* in a repeatedly played linear public goods environment (labeled C-experiment). The C-experiment consists of ten rounds in the random matching mode. We chose a random matching protocol to minimize strategic effects from repeated play. All participants play both types of experiments. For example, participants first go through the preference elicitation experiment in the P-C sessions before making their contribution choices in the C-experiment. Our C-P sessions counterbalance the order of experiments to control for possible sequence effects. The C-P sequence allows for a particularly strong test of measured preferences because people experience ten rounds of decisions in the C-experiment before their cooperation preferences are elicited in the P-experiment.

The rationale of the P-experiment is to elicit people's cooperation preferences: to what degree are people willing to cooperate given other people's degrees of cooperation?[3] Being able to observe contributions as a function of other group members' contributions without using deception requires the observation of contributions that can be contingent on others' contributions. Fischbacher, Gächter, and Fehr (2001) (henceforth FGF) introduced an experimental design that accomplishes this task.[4] Since we use exactly the same method as FGF, we refer the reader to FGF for all details.

The central idea of the P-experiment is to apply a variant of the so-called "strategy method" (Selten 1967). The participants' main task in the experiment is to make two decisions, an "unconditional contribution" and a "conditional contribution." In the conditional contribution a subject has to indicate—in an incentive-compatible way—how much he or she wants to contribute to the public good for each rounded average contribution level of other group members. Specifically, participants were shown a "contribution table" of the 21 possible values of the average contribution

[3] Our approach does not require eliciting a utility function since we do not need a complete preference order for our purposes. It is sufficient to know subjects' best replies conditional on others' contributions.

[4] Ockenfels (1999) developed a similar design independently of FGF.

of the other group members (from 0 to 20) and were asked to state their corresponding contribution for *each* of the 21 possibilities. Since the FGF method elicits the contribution schedules in an incentive-compatible way, free-rider types have an incentive to enter a zero contribution for each of the 21 possible average contributions of other group members; conditional cooperators have an incentive to enter increasing contributions; and unconditional cooperators have an incentive to enter their preferred contribution level.[5] The experiment was played only once, and the participants knew this. We wanted to elicit subjects' preferences, without intermingling preferences with strategic considerations.

Participants in the P-C sessions (C-P sessions) were informed only after finishing the P-experiment (C-experiment) that they would participate in another experiment. When we explained the C-experiment, we emphasized that the groups of four would be randomly reshuffled in each period.[6] After each period, participants were informed about the sum of contributions in their group in that period. In addition to their contribution decisions, subjects also had to indicate their *beliefs* about the average contribution of the other three group members in the current period. In addition to their earnings from the public good experiment, we paid participants based on the accuracy of their estimates.[7]

We elicited beliefs for two reasons. First, we can assess the correlation between beliefs and contributions, which we expect to differ between types of players and which helps us to check on the player type as elicited in the P-experiment. Second, by evaluating an elicited schedule at the elicited belief in a given period we can make a point prediction about an individual's contribution in the C-experiment: if an individual in the P-experiment indicates that he or she will contribute y tokens if the other group members contribute x tokens on average, then the prediction for this individual in the C-experiment is to contribute y tokens if he or she believes that others contribute x tokens on average. We will later refer to this as the "predicted contribution."

The sequence of experiments was reversed in the C-P sessions. The comparison of results from the P-experiments in the C-P sequence with those of the P-C sequence allows us to assess the relevance of experience with the public goods game for elicited cooperation preferences.

All experiments were computerized, using the software z-Tree (Fischbacher 2007). The experiments were conducted in the computer lab of the University of Zurich. Our participants were undergraduates from various disciplines (except economics) from the University of Zurich and the Swiss Federal Institute of Technology (ETH) in Zurich. We conducted six sessions (three in the P-C sequence and three in the C-P sequence). In each of five sessions we had 24 participants and in another session we had 20 participants. A postexperiment questionnaire confirmed that participants were largely unacquainted with one another. Our 140 participants were randomly assigned to cubicles in each session, where they took their decisions in complete anonymity from

[5] The P-experiment is incentive compatible because a random draw selects three group members for whom the unconditional contribution is payoff relevant and one group member for whom the conditional contribution, given the average unconditional contributions of the other three members, is payoff relevant. The payoffs are equal to (1). See FGF or Fischbacher and Gächter (2009) for further details.

[6] The likelihood in period 1 that a player would meet another player once again during the remaining nine periods was 72 percent. The likelihood that the *same* group of four players would meet was 2.58 percent. Since the experiment was conducted anonymously, however, subjects were unable to recognize whether they were matched with a particular player in the past.

[7] Participants had a financial incentive for correct beliefs, but it was small, to avoid hedging. If their estimation was exactly right, subjects received three experimental money units ($\approx$ \$0.8) in addition to their other experimental earnings. They received two additional money units if their estimation deviated by only one point from the other group members' actual average contribution, one money unit if they deviated by two points, and no additional money if their estimation was off the actual contribution by more than three points.

the other subjects. On average, participants earned 35 Swiss francs (CHF) (roughly $30, including a show-up fee of 10 CHF).[8] Each session lasted roughly 90 minutes.

II. Results

We organize the discussion of our results as follows. In Section A, we document the decline of cooperation. In Section B, we present the extent of heterogeneity in people's cooperation preferences and actual contribution patterns. In the remaining sections, we analyze behavior in the C-experiment. We show how subjects form their beliefs (Section C) and how their contribution decisions are related to the elicited preferences in the P-experiment (Section D). We conclude in Section E with a simulation study in which we assess how the belief process and subjects' preferences affect the decline of cooperation.

A. *The Decline of Cooperation*

Figure 1 sets the stage for our analysis, which aims to explain the decline of cooperation. The figure shows the temporal patterns of cooperation and beliefs for each of our six sessions separately.

The figure conveys four unambiguous messages. First, contributions and beliefs decline in six out of six sessions. Second, behavior in the six sessions is very similar. Third, contributions are lower than beliefs in almost all instances. Finally, mean period contributions and beliefs are significantly positively correlated in all six sessions (Spearman rank correlation tests, $p < 0.007$). This finding is consistent with previous observations in public goods games where beliefs were elicited (e.g., Weimann 1994; Croson 2007; Neugebauer et al., forthcoming).

B. *Heterogeneous Preferences and Contributions*

Recall that we have a complete contribution schedule from each participant that indicates how much he or she is prepared to contribute as a function of others' contributions. A simple way of characterizing heterogeneity is to look at the slope (of a linear regression) of the schedule and the mean contribution in the schedule. For instance, a free rider's schedule consists of zero contributions for all contribution levels of other group members. Therefore, his slope and mean contribution are zero. An unconditional cooperator, who contributes 20 tokens for all others' contribution levels has a mean contribution of 20 and a slope of 0. A perfect conditional cooperator, who contributes exactly the amount others contribute, has a slope of one and a mean contribution of ten tokens. Figure 2A depicts the results separately for the C-P and P-C experiments. The *x*-axis shows the slope of the schedules and the *y*-axis the average contribution in the schedule. The dots in Figure 2A correspond to individual observations, and the size of a dot to the number of observations it represents.

Figure 2A shows two things. First, there is a large degree of heterogeneity.[9] Free riders (located at 0–0) and perfect conditional cooperators (at 1–10) are relatively the largest group of subjects. We also find a few participants who contribute an unconditional positive amount (along the *y*-axis, at $x = 0$). Many participants have a positive mean contribution and a positive slope; a

[8] During the experiment subjects earned their payoffs in "points" (according to (1) and the earnings from correct belief estimates). At the end of the experiment, we exchanged the accumulated sum of points at an exchange rate of one point = CHF 0.35 for the points earned in the P-experiment and at a rate of 1 point = CHF 0.07 for the points earned in the C-experiment.

[9] This evidence is consistent with other studies using different methods. See footnote 2.

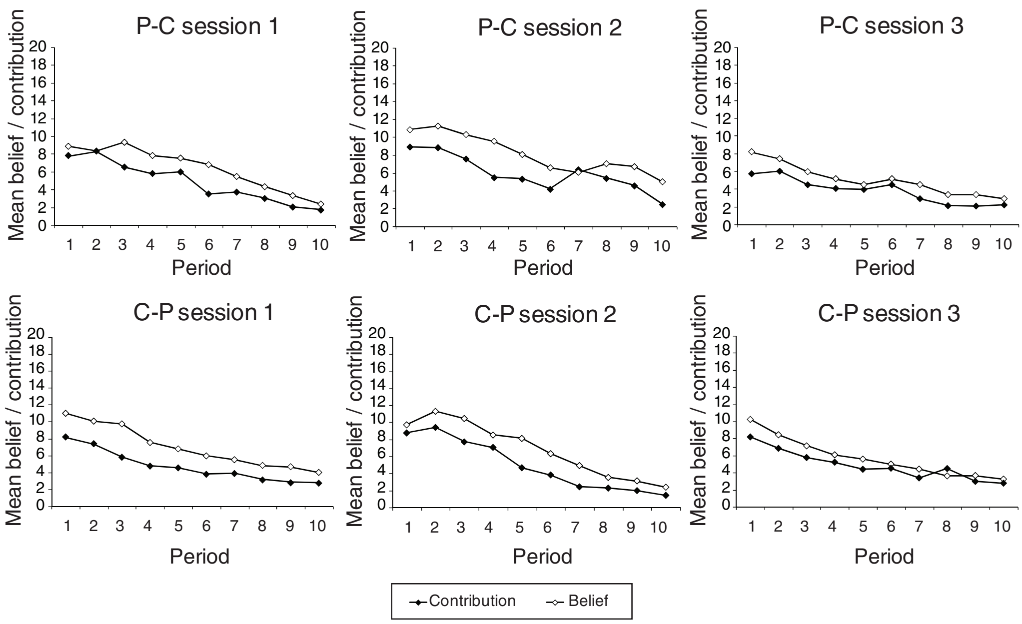

Figure 1. Mean Beliefs and Contributions over Time

few participants have a negatively sloped schedule (that is, they contribute more the less others contribute). Second, the distribution between the C-P and the P-C sessions is very similar. Mann-Whitney tests do not allow the rejection of the null hypotheses that both means and slopes are equally distributed between the treatments ($p > 0.87$).[10] Thus, elicitation of preferences before subjects actually experienced contributions to the public good (in the P-C sessions) or after (in the C-P sessions) did not affect the elicited preferences. This is an important finding for our interpretation that the P-experiment elicits cooperation preferences. It shows that participants in the C-P sessions who have experienced actual contribution behavior do not express different cooperation preferences than do participants in the P-C sessions who are inexperienced in actual game playing when they express their preferences.[11]

Figure 2B shows a scatter plot of individual slopes of linear regressions (estimated without intercepts) of contributions on beliefs on the x-axis, and average contributions in the C-experiment on the y-axis. The construction of Figure 2B is similar to Figure 2A. As in Figure 2A, we distinguish between the C-P and P-C sessions. We find no sequence effect, neither with respect to average contributions nor with respect to slopes (Mann-Whitney tests, $p > 0.21$).[12]

[10] In Figure 2A we looked at the slope and mean contribution of a subject's schedule. However, qualitatively, we get very similar results if we look at Spearman rank order correlation coefficients, linear correlation coefficients, and slopes and intercepts of linear regressions. In all cases p-values of Mann-Whitney tests that compare the C-P and the P-C experiments yield $p > 0.275$. The absolute number of [0–0]-combinations in the P-experiments was as follows: 18 in the C-P sessions and 14 in the P-C sessions. Six (seven) people are located at [1–10] in the C-P (P-C) sessions.

[11] The elicited contribution schedules in our study are also not significantly different from FGF (χ^2-test, $p = 0.729$). See an earlier version (Fischbacher and Gächter 2009) for further details.

[12] In the C-experiments depicted in Figure 2B, nine people are located at [0–0] in each of the C-P and the P-C sessions.

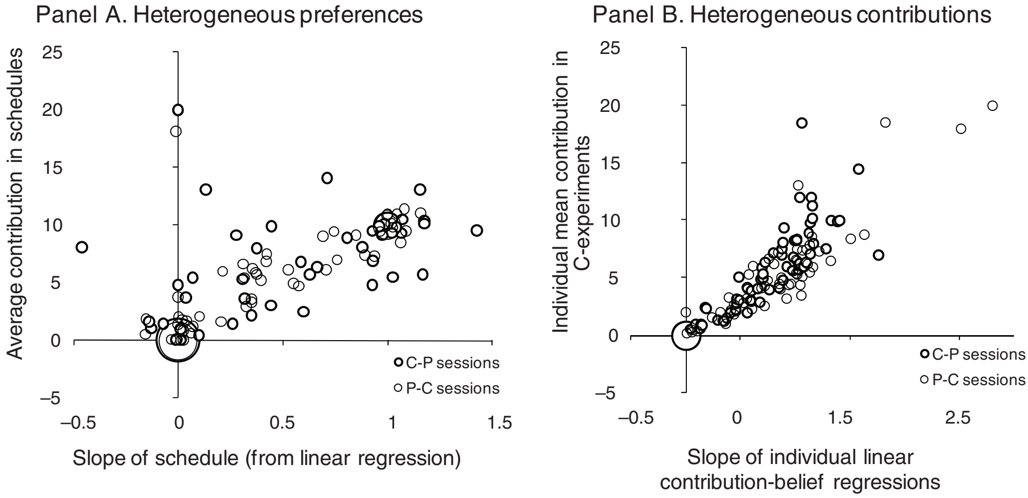

Figure 2. Heterogeneous Contribution Preferences (*panel A*) and Actual Contributions as a Function of Beliefs (*panel B*)

Figure 2B reveals considerable heterogeneity in contribution behavior. Individual average contributions (depicted on the *y*-axis) vary between 0 and 20 tokens, although most participants contribute fewer than ten tokens on average. Fourteen percent of all participants contribute exactly zero in all ten periods. We find that the individual estimated slopes of the schedules from the P-experiment (Figure 2A) and the slopes of individual linear contribution-belief regressions in the C-experiments (Figure 2B) are highly significantly positively correlated (Spearman's $\rho = 0.39, p = 0.0000$). Average cooperation levels in the P-experiment and in the C-experiment are highly correlated as well (Spearman's $\rho = 0.40$, $p = 0.0000$). We interpret this as a first piece of evidence that expressed cooperation preferences and actual cooperation behavior are correlated at the individual level.

Before we investigate the link between beliefs, preferences, and contributions, we look at how people form beliefs in the C-experiment. Understanding belief formation is important because previous evidence, and our own, suggests that beliefs have an influence on contributions (see Section IID). It is therefore possible that belief formation contributes to the decay of cooperation if people reduce their beliefs per se, that is, independently of contributions. In Section E we will address this possibility and a competing hypothesis suggested by our findings on contribution preferences (Figure 2A): contributions decline because people are imperfect conditional cooperators.

C. *The Formation of Beliefs*

With the help of three econometric models, we investigate the question of how people form their beliefs about their group members' contribution in a given period. The estimation method is OLS with robust standard errors clustered on sessions as the independent units of observation.[13]

[13] We estimated all models with random and fixed effects specifications, as well as with Tobit, with very similar results. For instance, the correlation coefficient of predicted values of the Tobit estimation and the OLS is 0.9995. Since the estimation results are very similar, we report the OLS results only for ease of interpretation.

Table 1—Forming Beliefs

	Dependent variable: Belief about other group members' contribution		
Model	(1)	(2)	(3)
Period	−0.761***	−0.079	
	(0.090)	(0.042)	
Others' contributions ($t-1$)		0.394***	0.415***
		(0.023)	(0.020)
Belief ($t-1$)		0.549***	0.569***
		(0.037)	(0.036)
Constant	10.711***	0.835*	0.118
	(0.864)	(0.398)	(0.148)
Observations	1,260	1,260	1,260
R^2	0.26	0.64	0.64

Notes: OLS regressions with data from period 2 to 10. Robust standard errors (clustered on sessions) in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Model 1 in Table 1, which includes only "period," simply confirms the impression from Figure 1 that beliefs decline significantly over time. However, this model cannot explain *why* there is a downward trend. Models 2 and 3 present our models of belief formation. We argue that people form their beliefs in period t on the basis of their beliefs in period $t-1$ and the observation of others' contributions in period $t-1$. To see this, take periods 1 and 2. In period 1 a subject can rely only on his or her intuitive ("home-grown") beliefs about others' contributions. In period 2, he or she also makes an observation about others' actual contribution in period 1. A subject may therefore update his or her period 2 belief on the basis of his or her period 1 belief and the observed period 1 contributions by others. A similar logic might hold in all remaining periods.

Model 2 presents the estimates of this belief formation model. We find that both "belief ($t-1$)" and "others' contribution ($t-1$)" are highly significantly positive; "period" is insignificant, an order of magnitude smaller than in model 1, and not significantly different from zero. We also estimated model 2 separately for periods 1 to 5 and periods 6 to 10. The estimated coefficients are very similar in both halves of the experiment (Chow-test, $p > 0.1$).[14] In other words, the way people form beliefs does not change over time.

Model 3 is the same as model 2 except that we drop the insignificant variable "period." The sum of coefficients of "belief ($t-1$)" and "others' contribution ($t-1$)" is insignificantly different from 1 ($F(1,5) = 0.41, p = 0.549$).[15] We will use this model in our simulations below.

[14] We also applied an Arellano-Bond linear, dynamic panel-data estimation method (Manuel Arellano and Stephen R. Bond 1991). However, there is still significant second order correlation ($p < 0.05$, Arellano-Bond test) in the residuals implying that its estimates are inconsistent (Arellano and Bond 1991, 281–82). Moreover, in simulations similar to those we discuss in Section IIE, it turned out that the Arellano-Bond estimates cannot explain the data patterns at all, whereas model 3 can. As a further robustness check of model 2, we included up to three lags for the variable "others' contributions." Only the first lag is significant; the higher lags are very small and insignificant.

[15] The period coefficient in model 2 is insignificantly different from zero but highly significantly different from the period coefficient in model 1 (seemingly unrelated regressions, $p < 0.001$). For reasons of comparability with model 2, we estimated model 1 for periods 2–10 only. The period coefficient in model 1 for all periods equals −0.753.

Table 2—Explaining Contributions

	Dependent variable: Contribution					
Model	(1)	(2)	(3)	(4a)	(4b)	(4c)
Periods used	1–10	1–10	1–10	1–10	1–5	6–10
Subjects excluded[a]	No	No	No	Yes	Yes	Yes
Period	−0.639 (0.071)***	−0.060 (0.056)				
Predicted contribution		0.242 (0.069)**	0.242 (0.069)**	0.443 (0.073)***	0.385 (0.074)***	0.614 (0.082)***
Belief		0.644 (0.071)***	0.666 (0.059)***	0.545 (0.065)***	0.582 (0.065)***	0.376 (0.116)**
Constant	8.343 (0.545)***	0.005 (0.569)	−0.473 (0.244)	−0.318 (0.312)	−0.204 (0.541)	−0.116 (0.378)
Observations	1,400	1,400	1,400	1,260	630	630
R^2	0.10	0.34	0.34	0.38	0.33	0.33

Note: Robust standard errors in parentheses.

[a] Models 4a to 4c exclude (confused) subjects who, on the basis of the P-experiment, could not be classified according to the FGF classification as either a "free rider," "conditional cooperator," or a "triangle contributor."

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Given these results, we can summarize the belief formation process as follows: a subject's belief in a given period is a weighted average of what he or she *believed* about others in the previous period and his or her *observation* of the others' contributions in the previous period. We will use this result when we investigate in Section IIE the role of belief formation for explaining the dynamics of voluntary cooperation.

D. *Explaining Contributions*

In this section we investigate determinants of people's contributions econometrically. We have three explanatory variables—"period," "predicted contribution," and "belief." We estimated three basic models, which we document in Table 2. The estimation method is OLS with robust standard errors clustered on sessions as the independent units of observation.[16]

Model 1, which includes only "period," confirms the impression from Figure 1 that contributions decline significantly over time. Model 2 adds the variables "predicted contribution" and "belief." We find that both variables matter significantly. In other words, because "belief" is significant, there is conditional cooperation on top of contribution preferences in the C-experiments. However, because "period" is an order of magnitude smaller than in model 1 and not significantly different from zero, the decline of cooperation must be due to "predicted contribution" and "belief." Model 3 is the same as model 2, but drops the insignificant variable "period."[17]

[16] As with belief formation, we estimated all models with random and fixed effects specifications, as well as with Tobit. Since the estimation results are very similar, we report only the OLS results.

[17] One might worry about multicollinearity in models 2 and 3 because beliefs enter the calculation of "predicted contributions." Although "belief" and "predicted contribution" are correlated (Spearman $\rho = 0.395$), the variance inflation factor as an indicator of multicollinearity is 1.22, which is below the critical value of ten (Damodar N. Gujarati 2003).

The observations from models 2 and 3 raise two related questions: (i) Why is the coefficient on "predicted contribution" surprisingly low (it should be one if the elicited preferences would predict perfectly)? (ii) Why do people condition their contribution decision based not only on their preferences according to their predicted contribution but also on their beliefs? Regression models 4a to 4c shed light on the first question. These models are the same as model 3 but exclude "confused" subjects (10 percent) who, according to the classification proposed by FGF, could not be classified as free riders, conditional cooperators, or triangle contributors. Model 4 shows that the coefficient on "predicted contributions" increases substantially (and significantly according to a seemingly unrelated regression, $p = 0.012$) once the confused subjects are excluded (the coefficient is still less than one, though, and beliefs still matter in addition to preferences). Regression models 4b and 4c reveal that there is also a time trend in the relative importance of "predicted contributions" and "beliefs." "Beliefs" seem to be more important in the first half of the experiment than in the second half, where the coefficient on "predicted contributions" is substantially and significantly higher than in the first half ($p < 0.001$), and vice versa for the coefficient on "beliefs" ($p = 0.039$).[18]

Why do beliefs matter in addition to contribution preferences? First, note that in models 4a to 4c, the constant is not significantly different from zero, and the sum of the coefficients for "predicted contribution" and "belief" add up to a number not significantly different from one (e.g., in Regression model 4a the sum equals 0.998; $F(1,5) = 0.08, p = 0.7863$). Hence, according to these regressions, subjects contribute a weighted average of "predicted contribution" and "belief." A contribution that matches the belief is, by definition, perfectly conditionally cooperative. Thus, subjects behave according to a contribution pattern that is in between their elicited contribution schedule and perfect conditional cooperation. Since most people's elicited contribution preferences lie below the diagonal, that is, below the schedule of perfect conditional cooperators, this intermediate contribution pattern lies above the predicted cooperation. This means that subjects are more conditionally cooperative in the C-experiment than predicted from their decisions in the P-experiment. Models 4b and 4c show that this is the case particularly in earlier periods; in later periods "predicted contribution" becomes more important than "beliefs." One potential explanation for why beliefs matter in addition to "predicted contribution" is subjects' willingness to invest in cooperation in order to induce high beliefs and contributions in the population, even in our random matching design (For some evidence in this regard, see Anabela Botelho et al., 2009.)

E. *Why Do Contributions Decline?*

Because both beliefs and contribution preferences matter for determining contributions, the question arises as to how they each contribute to explaining the decline of cooperation. In this section we use simulation methods to understand why contributions decline. We distinguish among three fundamental possible causes: what individual preferences are, how beliefs change over the course of the game, and heterogeneity (in preferences or beliefs). We use simulation methods because they allow us to use counterfactual assumptions, which are helpful in disentangling the role of preferences and beliefs.

The simulations are based on a two-stage process. In the first stage, the simulated players form a belief about the other players' contributions. Then, players decide on a contribution, which they (partially) base on their beliefs. Before we address our two main questions, we make the following basic observations on the conditions under which such a two-stage process can explain the decline of cooperation. Contributions will not decline if contributions or beliefs are independent of experience and time, that is, if people are *unconditionally cooperative* or if they

[18] However, the coefficients of "predicted contributions" and "beliefs" do not significantly differ from each other in the two regressions: ($F(1, 5) = 2.42, p = 0.181$ and $F(1, 5) = 1.60, p = 0.262$)

have *unconditional beliefs*. Thus, to explain the decline of cooperation, both beliefs and contributions need to be *conditional* on the behavior of others. Of course, conditionality is only a necessary condition for the decline of cooperation. Suppose, for instance, that belief updating is naïve (that is, beliefs are equal to the average contribution in the previous period) and contributions match beliefs. In this case, cooperation will be stable. Thus, cooperation will decline only if either beliefs are lower than naïve beliefs or if people are imperfect conditional cooperators. Our simulations will shed light on the relative importance of these two possibilities.

To be able to disentangle, in our simulations, the roles of beliefs and contributions for the decay of cooperation, we make (counterfactual) assumptions about cooperation behavior and belief formation. With regard to cooperation behavior, we will assume that the simulated players (i) contribute according to preferences we have observed in the P-experiment, or (ii) they (counterfactually) are all perfect conditional cooperators. The second dimension concerns belief updating. Here we assume that players either (i) update according to the weighted-average model outlined above (model 3 of Table 1), or (ii) (counterfactually) form their beliefs naïvely. Thus, we have 2×2 combinations of assumptions about cooperation behavior and belief formation. Starting from the benchmark of perfect conditional cooperation and naïve beliefs, we can hold one dimension constant and change the other to see whether belief formation or contributions are responsible for the decay of cooperation.

The simulations use the exact matching structure that was in place in each period of a given session.[19] As starting values, we use the actual contributions and beliefs in period 1. The details of our 2×2-methodology are as follows:

(i) In our benchmark model, the pCC_N-*model*, we assume that all players are perfect conditional cooperators, that is, players match their beliefs exactly: $Contribution(t) = Belief(t)$. Beliefs are formed naïvely (denoted by subindex N), that is, $Belief(t) = Others' \ Contribution(t-1)$. Under these assumptions contributions are obviously stable at the initial level of contributions.

(ii) The pCC_A-*model* keeps the assumption of perfect conditional cooperation but assumes that beliefs are formed according to the actual beliefs estimated in model 3 in Table 1 (denoted by subindex A). In this model, contributions will drop only if beliefs per se become inherently pessimistic.[20] Thus, this model reveals the extent to which the belief formation process can be responsible for the decay of cooperation.

(iii) The aCC_A-*model* keeps the actual beliefs assumption but replaces perfect conditional cooperation by actual conditional cooperation as elicited in the P-experiment (denoted aCC). The weights on contribution preferences and beliefs correspond to the estimated parameters of model 3 in Table 2. This simulation model shows the combined predicted effects of actual belief updating and actual contribution preferences for the decline of cooperation.

(iv) Finally, in the aCC_N-*model*, the simulated players determine their contributions according to the actual conditional contribution schedules, but beliefs are formed naïvely. By assuming naïve beliefs, this model reveals the extent to which the cooperation preferences themselves contribute to the decline of cooperation.

[19] That is, in simulation models 3 and 4 described below, we replace each human participant by his or her contribution schedule and observe how contributions evolve given our assumptions about belief updating.

[20] "Virtual learning" (Roberto Weber 2003), that is, learning with no feedback by just thinking about the problem for several periods, is a possible reason for this "pessimism."

We address the role of heterogeneity for the decline of cooperation with two counterfactual models in which we remove heterogeneity from the contribution process, thus, ending up with 3×2 combinations of assumptions. By comparing these models with the actual contributions, we can assess the role of heterogeneity. The models differ only in their assumptions of the belief formation process. One model assumes naïve beliefs and is thus comparable to the aCC_N-*model*; the other model uses the estimated belief formation and is thus comparable to aCC_A:

(v) In the iCC_N-*model* players are assumed to be identical conditional cooperators (*iCC*). As a schedule, we use an *average* linear one: *Contribution* $= \alpha + k \times$ *Contribution of others*. The estimates from the data of our P-experiment return $\alpha = 0.956$ and $k = 0.425$. Therefore, in this model *Contribution* $= 0.956 + 0.425 \times$ *Contribution of others*. The iCC_N-*model* assumes that belief formation is naïve. Thus, in comparison to the *aCC-models*, the iCC_N-*model* informs us about the role of preference heterogeneity of players under naïve belief formation.

(vi) Finally, the iCC_A-*model* is the same as the iCC_N-*model* but assumes actual belief formation.

Figure 3 depicts the simulation results. We compare the simulation results to the actual average contributions ("data"). Panel A illustrates the implications of our simulation models. For this purpose we use the average over all six sessions. Panel B depicts the predictive success of the aCC_A-*model* at the session level. We illustrate the results of the aCC_A-*model* because it is the only simulation model that does not use counterfactual assumptions.

Since the average initial contribution is eight tokens, the pCC_N-*benchmark* implies that contributions are stable at the initial level. The pCC_A-*model* predicts that contributions will decline to the extent beliefs decline. The simulation results return stable contributions. To understand this finding, recall that according to our model of actual belief formation (model 3 of Table 1) $Belief(t) = 0.415 \times Others'\ contribution(t-1) + 0.569 \times Belief(t-1) + 0.118$. By assumption of perfect conditional cooperation $Contribution(t-1) = Belief(t-1)$ for all players and therefore $Belief(t) = (0.415 + 0.569) \times Belief(t-1)$. The sum of the two coefficients (0.415 and 0.569) equals 0.984, which is insignificantly different from 1 (F-test, $p = 0.549$). This observation implies that, under perfect conditional cooperation, beliefs remain constant. Hence, we conclude that the belief formation process per se does not contribute to the decline of cooperation.[21] Put differently, beliefs decline because contributions decline, and not because people become inherently more pessimistic over time, irrespective of contribution behavior.

The aCC_A-*model* and the aCC_N-*models* replace the assumption of perfect conditional cooperation with people's actual contribution preferences as elicited in the P-experiment. Both simulation models predict a decline of cooperation. Since beliefs per se are not responsible for the unraveling of cooperation, we conclude that the decline is triggered by people's contribution preferences. The aCC_N-*model* predicts a much faster decline than we actually observe in the data. By contrast, the aCC_A-*model* tracks the actual data quite well. To evaluate this model statistically, we regress the actual contributions on the predicted contributions (using OLS). We find that the model coefficient (robust s.e.) equals 1.002 (0.030), which is not significantly different from one. The model constant (robust s.e.) equals –0.036 (0.188), which is not significantly different from zero. Thus, on average, the aCC_A-*model*

[21] If we disregard the constant and take the coefficient of 0.984 literally, the belief formation process per se can account for a decline of cooperation of at most 14 percent (i.e., $1 - (1 - 0.984)^9 = 0.135$) over the nine remaining periods after period 1.

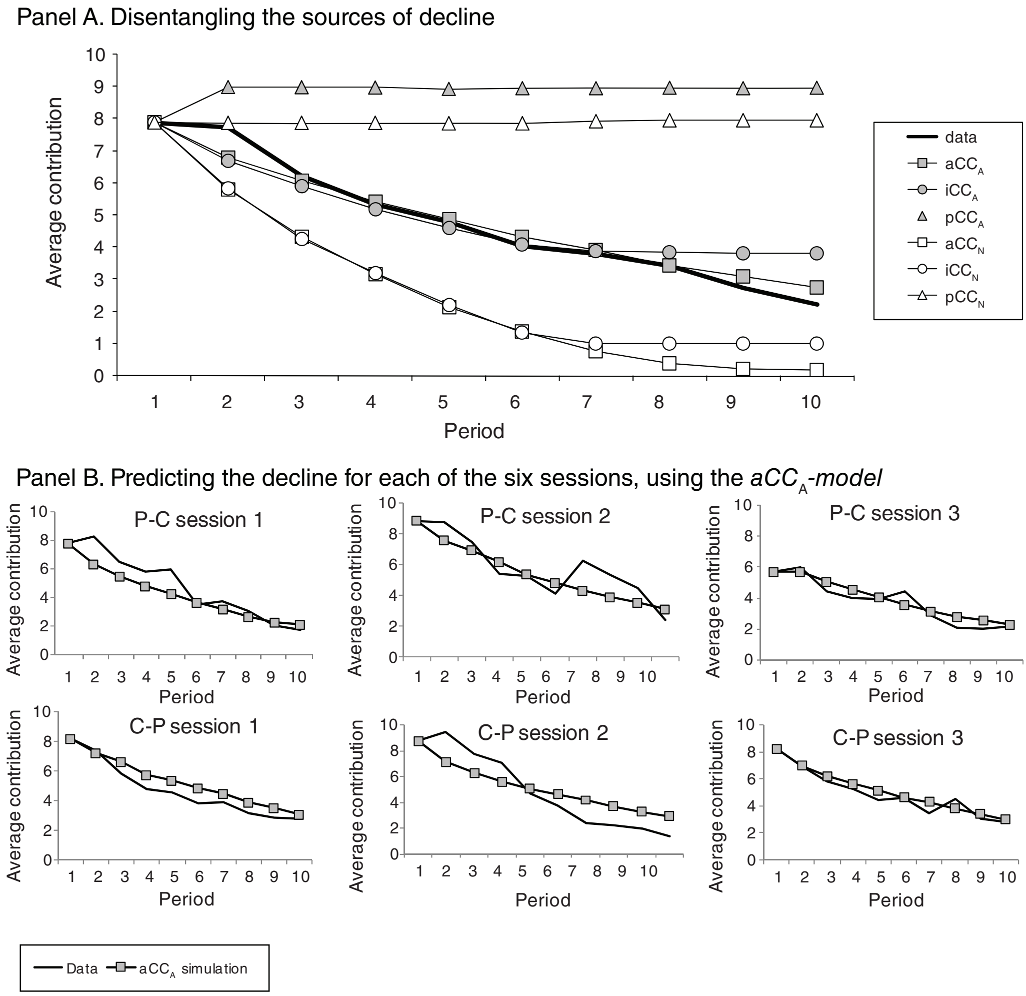

FIGURE 3. EXPLAINING THE DECLINE OF COOPERATION – SIMULATION RESULTS

predicts the data very well. This is also apparent from Figure 3B, which compares the session averages with the predictions from the *aCC_A-model* applied to the respective session.[22]

Finally, we can assess the importance of preference heterogeneity by comparing the *aCC-models* with the *iCC-models* (under both naïve and actual belief updating). By construction, the *iCC-models* eliminate preference heterogeneity by replacing the individual preference schedules by the average preference schedule, whereas the *aCC-models* use the individual preference schedules. The comparison shows that preference heterogeneity is surprisingly unimportant in explaining the decay of cooperation because the *aCC-models* and the *iCC-models* match each

[22] All other models perform worse. In the *iCC_A-model* as the second best model, the regression of actual contributions on the predicted contributions has a slope of 1.06, which is insignificantly different from one, and the constant equals -0.432, which is weakly significantly different from 0 ($p = 0.061$). In all other models, the model constant is highly significantly different from zero or the slope is highly significantly different from 1 ($p < 0.01$). Furthermore, the correlation between actual and predicted contribution is highest for the *aCC_A-model* (Spearman correlation $= 0.46$). It is lower in all other models (all Spearman correlations $p < 0.40$).

other closely under both naïve and actual belief updating. Heterogeneity matters only toward the end of the experiment. In the *iCC-models,* contributions stop declining toward the end while the models with heterogeneous preferences also correctly predict the decline in the last periods. Due to more realistic belief updating, the iCC_A*-model* matches the data better than the iCC_N*-model*.

VII. Summary and Concluding Remarks

Our goal in this study was to investigate the role of social preferences and beliefs about others' contributions for the dynamics of free riding in public goods experiments. We achieved this by eliciting preferences in one specially designed game (the P-experiment) and observing contributions and beliefs in ten one-shot standard public goods games with random matching (the C-experiment).

Our findings show that contributions decline (in randomly composed groups) because, on average, people are imperfect conditional cooperators. After some time, all types *behave* like income-maximizing free riders, even though only a minority is *motivated* by pure income-maximization alone.

Scholars across the social sciences (Paul Samuelson 1954; Mancur Olson 1965; Garrett Hardin 1968) have long noted the inherent fragility of voluntary cooperation due to free-rider incentives in the voluntary provision of public goods, collective actions, and common pool resource management. Our study shows that voluntary cooperation is inherently fragile, even if most people are not free riders but conditional cooperators. Because most of them are only imperfectly conditionally cooperative, however, cooperation is bound nevertheless to unravel; the presence of free riders just speeds up the decline. Other mechanisms, like punishment or rewards (Elinor Ostrom, Walker, and Roy Gardner 1992; Fehr and Gächter 2000; Sefton, Shupp, and Walker 2007), communication (Isaac and Walker 1988; Ostrom, Walker, and Gardner 1992; Jeannette Brosig, Ockenfels, and Weimann 2003), tax-subsidy mechanisms (Josef Falkinger et al. 2000; Jürgen Bracht, Charles Figuieres, and Marisa Ratto 2008), or, in general, good institutional designs (Ostrom 1990) are necessary to sustain cooperation.

REFERENCES

Ambrus, Attila, and Parag A. Pathak. 2009. "Cooperation over Finite Horizons: A Theory and Experiments." Unpublished.

Andreoni, James. 1988. "Why Free Ride? Strategies and Learning in Public Goods Experiments." Journal of Public Economics, 37(3): 291–304.

Andreoni, James. 1995. "Cooperation in Public-Goods Experiments: Kindness or Confusion?" *American Economic Review*, 85(4): 891–904.

Andreoni, James A., and John H. Miller. 1993. "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence." *Economic Journal*, 103(418): 570–85.

Arellano, Manuel, and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58(2): 277–97.

Ashley, Richard, Sheryl Ball, and Catherine Eckel. Forthcoming. "Motives for Giving: A Reanalysis of Two Classic Public Goods Experiments." *Southern Economic Journal*.

Bardsley, Nicholas, and Peter G. Moffatt. 2007. "The Experimetrics of Public Goods: Inferring Motivations from Contributions." *Theory and Decision*, 62(2): 161–93.

Bayer, Ralph-C, Elke Renner, and Rupert Sausgruber. 2009. "Confusion and Reinforcement Learning Experimental in the Public Goods Games." University of Nottingham Centre for Decision Research and Experimental Economics Discussion Paper 2009-18.

Botelho, Anabela, Glenn W. Harrison, Lígia M. Costa Pinto, and Elisabet E. Rutström. 2009. "Testing Static Game Theory with Dynamic Experiments: A Case Study of Public Goods." *Games and Economic Behavior*, 67():253–65.

Bracht, Jürgen, Charles Figuieres, and Marisa Ratto. 2008. "Relative Performance of Two Simple Incentive Mechanisms in a Public Goods Experiment." *Journal of Public Economics*, 92(1–2): 54–90.

Brandts, Jordi, and Arthur Schram. 2001. "Cooperation and Noise in Public Goods Experiments: Applying the Contribution Function Approach." *Journal of Public Economics*, 79(2): 399–427.
Brosig, Jeannette, Joachim Weimann, and Axel Ockenfels. 2003. "The Effect of Communication Media on Cooperation." *German Economic Review*, 4(2): 217–41.
Burlando, Roberto, and Francesco Guala. 2005. "Heterogeneous Agents in Public Goods Experiments." *Experimental Economics*, 8(1): 35–54.
Burlando, Roberto, and John D. Hey. 1997. "Do Anglo-Saxons Free-Ride More?" *Journal of Public Economics*, 64(1): 41–60.
Carpenter, Jeffrey P. 2007. "Punishing Free-Riders: How Group Size Affects Mutual Monitoring and the Provision of Public Goods." *Games and Economic Behavior*, 60(1): 31–51.
Cookson, Richard. 2000. "Framing Effects in Public Goods Experiments." *Experimental Economics*, 3(1): 55–79.
Cooper, Russell, Doug DeJong, Robert Forsythe, and Thomas Ross. 1996. "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games." *Games and Economic Behavior*, 12(2): 187–218.
Croson, Rachel.1996. "Partners and Strangers Revisited." *Economics Letters*, 53(1): 25–32.
Croson, Rachel. 2007. "Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Goods Games." *Economic Inquiry*, 45(2): 199–216.
Croson, Rachel, Enrique Fatas, and Tibor Neugebauer. 2005. "Reciprocity, Matching and Conditional Cooperation in Two Public Goods Games." *Economics Letters*, 87(1): 95–101.
Croson, Rachel, and Jen Shang. 2008. "The Impact of Downward Social Information on Contribution Decisions." *Experimental Economics*, 11(3): 221–33.
Duffy, John, and Jack Ochs. 2009. "Cooperative Behavior and the Frequency of Social Interaction." *Games and Economic Behavior*, 66(2): 785–812.
Egas, Martijn, and Riedl, Arno. 2008. "The Economics of Altruistic Punishment and the Maintenance of Cooperation." *Proceedings of the Royal Society B – Biological Sciences*, 275(1637): 871–78.
Falkinger, Josef, Ernst Fehr, Simon Gächter, and Rudolf Winter-Ebmer. 2000. "A Simple Mechanism for the Efficient Provision of Public Goods: Experimental Evidence." *American Economic Review*, 90(1): 247–64.
Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review*, 90(4): 980–94.
Ferraro, Paul J., and Christian A. Vossler. 2005. "The Dynamics of Other-Regarding Behavior and Confusion: What's Really Going on in Voluntary Contributions Mechanism Experiments?" Georgia State University Andrew Young School of Policy Studies Working Paper 2005–001.
Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics*, 10(2): 171–78.
Fischbacher, Urs, and Simon Gächter. 2009. "On the Behavioral Validity of the Strategy Method in Public Goods Experiments." University of Nottingham Centre for Decision Research and Experimental Economics Discussion Paper 2009–25.
Fischbacher, Urs, Simon Gächter, and Ernst Fehr. 2001. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters*, 71(3): 397–404.
Frey, Bruno S., and Stephan Meier. 2004. "Social Comparisons and Pro-Social Behavior: Testing "Conditional Cooperation" In a Field Experiment." *American Economic Review*, 94(5): 1717–22.
Gächter, Simon. 2007. "Conditional Cooperation: Behavioral Regularities from the Lab and the Field and Their Policy Implications." In *Psychology and Economics: A Promising New Cross-Disciplinary Field*, ed. Bruno S. Frey and Alois Stutzer, 19–50. Cambridge, MA: MIT Press.
Gächter, Simon, and Ernst Fehr. 1999. "Collective Action as a Social Exchange." *Journal of Economic Behavior and Organization*, 39(4): 341–69.
Gächter, Simon, Elke Renner, and Martin Sefton. 2008. "The Long-Run Benefits of Punishment." *Science*, 322(5907): 1510.
Goeree, Jacob K., Charles A. Holt, and Susan K. Laury. 2002. "Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior." *Journal of Public Economics*, 83(2): 255–76.
Gujarati, Damodar N. 2003. *Basic Econometrics, 4th Edition*. New York: MacGrawHill.
Guttman, Joel M. 1986. "Matching Behavior and Collective Action: Some Experimental Evidence." *Journal of Economic Behavior and Organization*, 7(2): 171–98.
Hardin, Garrett. 1968. "The Tragedy of the Commons." *Science*, 162(3859): 1243–1148.
Herrmann, Benedikt, and Christian Thöni. 2009. "Measuring Conditional Cooperation: A Replication Study in Russia." *Experimental Economics*, 12(1): 87–92.
Herrmann, Benedikt, Christian Thöni, and Simon Gächter. 2008. "Antisocial Punishment across Societies." *Science*, 319(5868): 1362–67.

Houser, Daniel, and Robert Kurzban. 2002. "Revisiting Kindness and Confusion in Public Goods Experiments." *American Economic Review*, 92(4): 1062–69.

Isaac, R. Mark, and James M. Walker. 1988. "Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism." *Economic Inquiry*, 26(4): 585–608.

Isaac, R. Mark, James M. Walker, and Susan H. Thomas. 1984. "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations." *Public Choice*, 43(2): 113–49.

Kelley, Harold, and Anthony Stahelski. 1970. "Social Interaction Basis of Cooperators' and Competitors' Beliefs About Others." *Journal of Personality and Social Psychology*, 16(1): 190–219.

Keser, Claudia, and Frans van Winden. 2000. "Conditional Cooperation and Voluntary Contributions to Public Goods." *Scandinavian Journal of Economics*, 102(1): 23–39.

Kocher, Martin G., Todd Cherry, Stephan Kroll, Robert J. Netzer, and Matthias Sutter. 2008. "Conditional Cooperation on Three Continents." *Economics Letters*, 101(3): 175–78.

Kurzban, Robert, and Daniel Houser. 2005. "Experiments Investigating Cooperative Types in Humans: A Complement to Evolutionary Theory and Simulations." *Proceedings of the National Academy of Sciences of the United States of America*, 102(5): 1803–07.

Laury, Susan K., James M. Walker, and Arlington W. Williams. 1995. "Anonymity and the Voluntary Provision of Public Goods." *Journal of Economic Behavior and Organization*, 27(3): 365–80.

Masclet, David, Charles Noussair, Steven Tucker, and Marie-ClaireVilleval. 2003. "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism." *American Economic Review*, 93(1): 366–80.

Muller, Laurent, Martin Sefton, Richard Steinberg, and Lise Vesterlund. 2008. "Strategic Behavior and Learning in Repeated Voluntary Contribution Experiments." *Journal of Economic Behavior and Organization*, 67(3–4): 782–93.

Neugebauer, Tibor, Javier Perote, Ulrich Schmidt, and Malte Loos. 2009. "Selfish-Biased Conditional Cooperation: On the Decline of Contributions in Repeated Public Goods Experiments." *Journal of Economic Psychology*, 30(1): 52–60.

Nikiforakis, Nikos, and Hans-Theo Normann. 2008. "A Comparative Statics Analysis of Punishment in Public-Good Experiments." *Experimental Economics*, 11(4): 358–69.

Ockenfels, Axel. 1999. *Fairness, Reziprozität und Eigennutz – Ökonomische Theorie und Experimentelle Evidenz*. Tübingen: Mohr Siebeck.

Ockenfels, Axel, and Joachim Weimann. 1999. "Types and Patterns: An Experimental East-West-German Comparison of Cooperation and Solidarity." *Journal of Public Economics*, 71(2): 275–87.

Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action, the Political Economy of Institutions and Decisions*. Cambridge: Cambridge University Press.

Ostrom, Elinor, James M. Walker, and Roy Gardner. 1992. "Covenants with and without a Sword: Self-Governance Is Possible." *American Political Science Review*, 86(2): 404–17.

Palfrey, Thomas R., and Jeffrey E. Prisbrey. 1997. "Anomalous Behavior in Public Goods Experiments: How Much and Why?" *American Economic Review*, 87(5): 829–46.

Park, Eun-Soo. 2000. "Warm-Glow Versus Cold-Prickle: A Further Experimental Study of Framing Effects on Free-Riding." *Journal of Economic Behavior and Organization*, 43(4): 405–21.

Samuelson, Paul A. 1954. "The Pure Theory of Public Expenditure." *The Review of Economics and Statistics*, 36(4): 387–89.

Sefton, Martin, Robert Shupp, and James M. Walker. 2007. "The Effect of Rewards and Sanctions in Provision of Public Goods." *Economic Inquiry*, 45(4): 671–90.

Selten, Reinhard. 1967. "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperimentes." In *Beiträge Zur Experimentellen Wirtschaftsforschung,* ed. Heinz Sauermann, 136–68. Tübingen: J.C.B. Mohr (Paul Siebeck).

Selten, Reinhard, and Rolf Stoecker. 1986. "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning-Theory Approach." *Journal of Economic Behavior and Organization*, 7(1): 47-70.

Sonnemans, Joep, Arthur Schram, and Theo Offerman. 1999. "Strategic Behavior in Public Good Games: When Partners Drift Apart." *Economics Letters*, 62(1): 35–41.

Sugden, Robert. 1984. "Reciprocity: The Supply of Public Goods through Voluntary Contributions." *Economic Journal*, 94(376): 772–87.

Weber, Roberto A. 2003. "'Learning' with No Feedback in a Competitive Guessing Game." *Games and Economic Behavior*, 44(1): 134–44.

Weimann, Joachim. 1994. "Individual Behavior in a Free Riding Experiment." *Journal of Public Economics*, 54(2): 185–200.