Universität St.Gallen

Temporal stability and psychological foundations of cooperation preferences

Stefan Volk, Christian Thöni, Winfried Ruigrok

January 2011 Discussion Paper no. 2011-01

School of Economics and Political Science, Department of Economics

University of St. Gallen

Editor:	Martina Flockerzi University of St. Gallen School of Economics and Political Science Department of Economics Varnbüelstrasse 19 CH-9000 St. Gallen Phone +41 71 224 23 25 Fax +41 71 224 31 35 Email seps@unisg.ch
Publisher:	School of Economics and Political Science Department of Economics University of St. Gallen Varnbüelstrasse 19 CH-9000 St. Gallen Phone +41 71 224 23 25
Electronic Publication:	Fax +41 71 224 31 35 http://www.seps.unisg.ch

Temporal stability and psychological foundations of cooperation preferences

Stefan Volk, Christian Thöni, Winfried Ruigrok

Author's address:

Stefan Volk
University of Tübingen
Faculty of Business Administration and Economics
Melanchthonstrasse 30
D-72074 Tübingen, Germany
Email stefan.volk@uni-tuebingen.de

Christian Thöni
University of St. Gallen
School of Economics and Political Science
Varnbüelstrasse 14
CH-9000 St.Gallen, Switzerland
Email christian.thoeni@unisg.ch

Winfried Ruigrok
University of St. Gallen
School of Management
Dufourstrasse 40a
CH-9000 St.Gallen, Switzerland
Email winfried.ruigrok@unisg.ch

Abstract

A core element of economic theory is the assumption of stable preferences. We test this assumption in public goods games by repeatedly eliciting cooperation preferences in a fixed subject pool over a period of five months. We find that cooperation preferences are very stable at the aggregate level, but less so at the individual level. Nevertheless, individual preferences are sufficiently stable to predict future behavior fairly accurately. Our results also provide evidence on the psychological foundations of cooperation preferences. The personality dimension 'Agreeableness' is closely related to both the type and the stability of cooperation preferences.

Keywords

Social preferences, preference stability, conditional cooperation, free riding, personality, Big-Five.

JEL Classification

C91, C72, H41

1. Introduction

In his influential essay on “The Methodology of Positive Economics” (1953, p. 4), Milton Friedman explains that the task of economic theory “is to provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances.” In brief, economic theory is about making predictions, which are enormously facilitated if peoples' tastes are stable across time and circumstances. The assumption of stable preferences belongs accordingly to the core elements of economic theory, or, as Gary Becker (1976, p. 5) puts it: “The combined assumptions of maximizing behavior, market equilibrium, and stable preferences, used relentlessly and unflinchingly, form the heart of the economic approach as I see it.”

On the same page Becker also argues that “preferences are assumed not to change substantially over time”. He explains that the assumption of preference stability both across different circumstances and over time “provides a stable foundation for generating predictions about responses to various changes, and prevents the analyst from succumbing to the temptation of simply postulating the required shift in preferences to “explain” all apparent contradictions to his predictions.” Whether the assumption of preference stability is reasonable for real world preferences is ultimately an empirical question. However, empirical evidence on preference stability is surprisingly scarce.

In this paper we present experimental evidence on the stability of social preferences, more precisely on the stability of preferences for cooperation in a social dilemma. We measure cooperation preferences by Fischbacher et al.’s (2001, henceforth FGF) variant of the four player one-shot public goods game and repeat the measurement with the identical subject pool after two and a half and five months. The FGF method has the advantage of measuring cooperativeness independent of subjects’ beliefs, which are presumably more volatile than preferences.[1] Furthermore, we also measure subjects’ personality using the Big-Five model, which is the standard trait framework for research in personality (e.g., Goldberg, 1993). Our experimental setup enables us to investigate (i) aggregate-level and individual-level stability of cooperation

[1] There is of course no way to measure preferences directly. In a strict sense it is thus impossible to test preference stability. We also cannot identify the proximate mechanism which causes subjects to contribute, be it interdependent preferences or reciprocity (see Cox, 2004; Falk and Fischbacher, 2006; Fehr and Schmidt, 1999; Dufwenberg and Kirchsteiger, 2004). However, we argue that the FGF design is a suitable mechanism to elicit revealed preferences for cooperation. If people have stable social preferences, they should behave identical at all three points in time, irrespective of what they think others will do. See also Thöni, Tyran, and Wengström (2009) on the distinction between beliefs and preferences in cooperation.

preferences over time, and (ii) the relationship between stable personality traits and cooperation preferences.

Our paper contributes to a nascent body of literature on the stability of social preferences. However, the existing literature has largely focused on the distribution of heterogeneous preference types across different experiments at the aggregate level.[2] The few studies that have examined individual-level stability can be divided into two categories.

The first category includes studies investigating the consistency and stability of social preferences across different games or variants of the same game in a given experimental session. Andreoni and Miller (2002) and Fisman et al. (2007) apply the axioms of revealed preferences to decisions observed in modified dictator games. They find that subjects' choices can be rationalized by a well-behaved utility function. Fischbacher and Gächter (2010) elicit individual cooperation preferences in a strategy method experiment and observe contributions in ten consecutive one-shot games in the direct response method. They find that data on individual cooperation preferences allow the prediction of the development of contributions over time very accurately.[3] Blanco et al. (in press) investigate the stability of preferences across different games by observing the same sample of subjects in a number of different one-shot games. They find that about one third of their subjects exhibit stable preferences in the sense that they consistently follow a plausible behavioral norm.[4]

The second category of research includes a very small number of studies examining longitudinal stability of social preferences over time. Muller et al. (2008) elicit subjects' cooperation preferences repeatedly in a series of five consecutively played two-stage public goods experiments within one session. They find that 37 percent of their 60 subjects remain in the same preference category for all five measurements. Studies allowing for more time between the measurements are particularly scarce.[5] Brosig et al. (2007) implement sequential prisoner's dilemma games and repeat the initial experiment with the same subject pool two times with one

---

[2] For experiments using the FGF design see Kocher et al. (2008), Herrmann and Thöni (2009), Thöni et al. (2009). Homo-/heterogeneity between subject pools in repeated public goods games are studied in Gächter et al. (2010).

[3] Burlando and Guala (2005) and Gächter and Thöni (2005) investigate preference stability in a setting where they regroup subjects according to their cooperation preferences. They use initial public goods games to identify different types and form groups of alike subjects and observe contributions in repeated public goods games.

[4] See also de Oliveira et al. (2009). A related strand of literature investigates whether different elicitation methods (strategy method vs. direct response method) lead to identical inference about subjects' preferences, see Brandts and Charness (2009) for an overview or Fischbacher and Gächter (2009) for the FGF design.

[5] Horowitz (1992), Andersen et al. (2005) and Zeisberger et al. (in press) investigate the stability of risk preferences over time. Meier and Sprenger (2010) present data on the stability of time preferences.

month in between each repetition. They find that 43 percent of the subjects choose the same response in all three waves, and all of these subjects act consistently selfish. On average they observe a decay of cooperative behavior over time.

In contrast to Brosig et al. (2007), we observe highly stable aggregate results. The distribution of cooperation preferences is basically unchanged across time. At the individual level we observe a more diverse picture. If we classify our subjects into three categories (Conditional Cooperators, Free Riders and Others) we find that about two thirds of the subjects remain in the same category for two consecutive measurements. Half of the subjects remain in the same category for all three measurements. Furthermore, we find a relationship between preference stability and dispositional characteristics. In particular, the personality dimension Agreeableness, which is related to one's tendency to be cooperative rather than competitive (Costa and McCrae, 1992), appears to be part of the psychological foundations of cooperation preference heterogeneity and stability. We find that Free Riders and Conditional Cooperators are characterized by different levels of Agreeableness and that higher levels of Agreeableness are associated with more stable preferences for conditional cooperation, while lower levels are associated with more stable preferences for free riding. The latter findings contribute to an emerging research stream that seeks to identify "personality and attitude variables that allow the classification of subjects into different "types" whose decision behavior in social dilemmas may be described by alternative models" (Rapoport and Suleiman, 1993, p. 193).[6] The remainder of the paper is organized as follows. In Section 2 we describe our experimental design and our questionnaire, in Section 3 we present the results and Section 4 concludes.

2. Methods

2.1 Experimental design and procedures

The data reported in this paper come from an experimental part and a questionnaire part. In the experimental part we used the design introduced by FGF to measure cooperation preferences in public goods games. In order to analyze whether cooperation preferences are stable over time, we

[6] The relation between personality measures and other-regarding, prosocial behavior has been studied in public goods games (Kurzban and Houser, 2001; Perugini et al., 2010), dictator games (Ben-Ner et al., 2004a,b), ultimatum games (Brandstätter and Königstein, 2001) and a series of distribution games (Bartling et al., 2009). Ashton et al. (1998) and Hirsh and Peterson (2009) study the correlation between personality measures and nonincenticed measures for cooperation. Boone et al. (1999; 2002) correlate personality measure like 'Locus of control' with cooperation in prisoners' dilemma games.

replicated the protocol used by FGF and repeated our initial study (Wave 1) in a random matching mode two and a half (Wave 2) and five months (Wave 3) after the first study by inviting the same participants back to the laboratory. All sessions used an identical protocol in which subjects were randomly assigned to groups of four members. Group compositions were unknown to the subjects and not revealed after the experiments. The basic decision situation was a standard linear public goods game. Each subject received an endowment of 20 tokens and chose a contribution $c_i \in \{0,1,...,20\}$ to a linear public good with a marginal per capita return of 0.4. The payoff function was given by

$$\pi_i = 20 - c_i + 0.4\sum_{j=1}^{4} c_j ,$$

where the public good is equal to the sum of the contributions of all four group members.

In the experiment subjects had to make two types of contribution decisions, an 'unconditional' and a 'conditional' contribution to the public good. The unconditional contribution was a single decision about how many of the 20 tokens to invest into the public good. For the conditional contribution, subjects had to fill in a table showing the 21 possible average contribution levels of the other three group members (rounded to integers). They were asked to state for each of the 21 possibilities their corresponding contribution. After all subjects had made both types of decisions a random mechanism determined which of the two decisions became outcome relevant. The random mechanism (throw of a die) selected in each group one subject that contributed according to his or her 'contribution table', while the other three group members contributed according to their unconditional contribution.

We conducted our experiments in the computer lab of a European university, using the software z-Tree (Fischbacher, 2007). In order to maximize comparability with the original experiment by FGF we replicated the same protocol and used the same parameters. We observed the decisions of 72 subjects at times 1 and 2. Four of these did not return for the third study leaving us with 68 subjects with complete information at all three points in time. The participants were randomly seated at separated computer terminals, received written instructions and had to answer a number of control questions. At the end of the experiments the participants were informed about their final payoff and paid privately. On average, subjects earned about US$ 13 in each of the three experiments.[7]

2.2 Questionnaire

In the experiment of Wave 1 we assessed individual differences in personality traits by a measure of the Big-Five Model of personality. Personality traits have an estimated annual stability coefficient of 0.98, indicating high stability over lifetime (Conley, 1985). Research has accordingly shown that subjects completing a personality questionnaire more than once will tend to obtain highly similar scores (e.g., McCrae and Costa, 1990). In order to reduce the potential for transient measurement errors (e.g., Chmielewski and Watson, 2009) we therefore measured personality characteristics in Wave 1 and did not repeat the measurement in the subsequent waves.

The Big-Five model is the standard trait framework for research in personality (Goldberg, 1993; John and Srivastava, 1999) and has enjoyed increasing popularity across a wide variety of disciplines including economics.[8] The model specifies that five overarching dimensions account for the biggest part of between-subject variation in stable personality traits. These dimensions are Extraversion (sociable, active, energetic), Agreeableness (cooperative, considerate, trusting), Conscientiousness (dependable, organized, persistent), Emotional Stability (calm, secure, unemotional) and Openness to Experience (imaginative, intellectual, artistically sensitive). While we measured all five dimensions, one of them, Agreeableness, is particularly relevant for our study of cooperation preferences. Agreeableness is centrally related to one's tendencies to strive for cooperation rather than competition (Costa and McCrae, 1992). Moreover, individuals high on Agreeableness are generally more inclined to forgo self-interests in favor of collective interests than individuals low on Agreeableness (e.g., Buss, 1991; Graziano and Eisenberg, 1997; Koole et al., 2001). We therefore expected Agreeableness to be positively correlated with subjects' cooperativeness in our experiments.

[7] Prior to the experiment reported here subjects played a one-shot public goods game and a one-shot public goods game with punishment. Subjects did, however, not receive any information about other subjects' decisions nor about their payoff. At the end of the experiments they were informed only of their aggregate payoff in all three experiments.

[8] See, e.g., Ameriks et al. (2007); Bartling et al. (2009), Lo et al. (2005).

We elicited personality traits by a Ten-Item Personality Inventory (TIPI), which includes a subscale for each of the five personality dimensions (Gosling et al., 2003). Each subscale is composed of two items and each item contains a pair of two trait descriptors. For example, the first item is: "*I see myself as* extraverted, enthusiastic." Subjects have to rate on a 7-point scale ranging from 1:'disagree strongly' up to 7:'agree strongly' the extent to which the pair of traits applies to them. The ten items are 'extraverted, enthusiastic' and 'reserved, quiet' (reverse-scored) for the Extraversion subscale; 'sympathetic, warm' and 'critical, quarrelsome' (reverse-scored) for the Agreeableness subscale; 'dependable, self-disciplined' and 'disorganized, careless' (reverse-scored) for the Conscientiousness subscale; 'calm, emotionally stable' and 'anxious, easily upset' (reverse-scored) for the Emotional Stability subscale; and 'open to new experiences, complex' and 'conventional, uncreative' (reverse-scored) for the Openness to Experience subscale.

3. Results

We organize our results as follows: We start by conducting an instrument check of our experiment and look at aggregate stability and individual stability. We then conduct an instrument check of our personality measure before we explore the impact of the Big-Five personality dimensions on preference heterogeneity and preference stability. Finally, we investigate the predictive power of preference stability. Specifically, we use the data from the first and second wave to predict individual behavior in the third wave and evaluate the prediction with the true data.

3.1 Instrument check public goods game

Since we use the FGF design, we start by checking whether we replicate previous findings from the literature in our Wave 1 experiment. Following the classification of preference types proposed by FGF, we divide our subjects into four categories. All subjects with a contribution schedule that has a significant positive slope ($p<0.01$, Spearman rank correlation) or shows a monotonically increasing pattern are classified as 'Conditional Cooperators'. 'Free Riders' contribute nothing in any case. Subjects who increase their contribution in the contribution schedule up to some maximum and decrease it thereafter are called 'Hump- shaped'. The remaining subjects fall into the category 'Others'. We (FGF) observe 58 (50) percent Conditional

Cooperators, 21 (30) percent Free Riders, 3 (14) percent Hump-shaped and 14 (6) percent Others. We run a χ^2 test on the joined data to check whether our distribution of types systematically differs from the one observed by FGF, and indeed the differences are borderline significant with $\chi^2(3) = 6.44$; $p = 0.092$. A closer look at the percentage numbers reveals that the greatest difference between our data and the FGF data is the smaller number of Hump-shaped patterns we observe. This is not completely unexpected, as other replication studies of the FGF design (e.g., Herrmann and Thöni, 2009; Kocher et al., 2008; Thöni et al., 2009) also find a lower percentage of Hump-shaped patterns. Due to the infrequent occurrence in our data we include the Hump-shaped patterns into the category Others. If we compare our data to FGF using this reduced classification we observe no significant differences ($\chi^2(3) = 0.771$; $p = 0.680$). We conclude that, with regard to the most interesting categories, our Wave 1 experiment produced very similar results in comparison to the previous literature.

3.2 Aggregate stability

Before we classify our data from the three waves into categories we look at the overall conditional cooperation scheme. Figure 1 shows the average conditional cooperation scheme for the three waves. The horizontal axis depicts the average contribution of the other three subjects, on which a subject can condition the own contribution. The vertical axis shows the average of all conditional contribution decisions in a wave.

The curves for the three waves have a similar positive slope, indicating that on average subjects are Conditional Cooperators. There seems to be a slight positive trend, i.e. the level of the average conditional contribution is higher in later waves. We therefore test for systematic differences between the three times of measurement using the average conditional cooperation of a subject in a wave as observation. Our results show no significant differences between the three waves.[9]

[9] Because we focus on average effects, we do not make use of the within subject comparisons available in our data at this point and use Wilcoxon rank sum tests. The corresponding p-values are $p = .349$ ($p = .506$) for the comparison between Wave 1 and Wave 2 (Wave 2 and Wave 3).

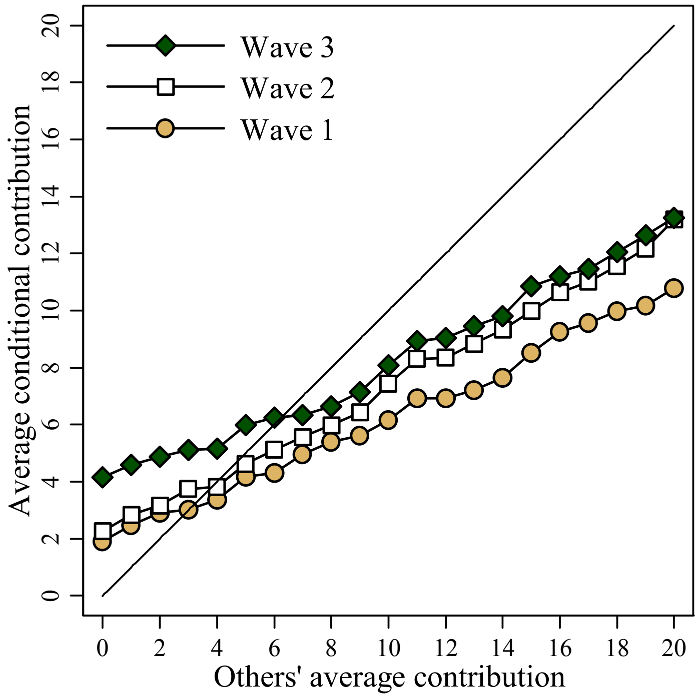


Figure 1: Average conditional cooperation for the three waves.

Overall, conditional contributions are thus relatively similar over the course of five months in the same subject pool. Let us now check whether the same holds true for the distribution of types in the population. The three bars in Figure 2 show the frequency of Conditional Cooperators, Free Riders and Others in the three waves of our experiment (ignore for the moment the dizzying lines between the bars). The distribution of types is apparently very similar across the three times of measurement. Free Riders account for between 19 percent (Wave 3) and 25 percent (Wave 1) of the population, while Conditional Cooperators account for between 56 percent (Wave 3) and 60 percent (Wave 2). The distribution of preference types is not significantly different across the three times of measurement ($\chi^2(4) = 1.994$; $p = 0.737$). This is in contrast to the results of Brosig et al. (2007) who report a decay of cooperation in the later measurements and speculate that their subjects learned the free rider strategy over time. While the learning hypothesis can neither explain our observations nor the observations of Muller et al. (2008), our data point to a different explanation for preference stability/instability at the individual level. As we will report in the next sections, our findings suggest that there is heterogeneity of individual behavior in terms of both preference types and preference stability (i.e., some individuals seem to have stable preferences, while others do not), which can be linked to individual differences in dispositional personality traits.

To conclude, our observations show high aggregate stability over time. Observed changes in the distribution of types as well as in average conditional contribution schemes are both small and insignificant. In all three waves we observe shares of Conditional Cooperators (56 – 60 percent) and Free Riders (19 – 25 percent), which are comparable to the shares reported in a number of previous studies. These studies have found that between 50 and 60 percent of people can be classified as cooperators, while between 20 and 30 percent are Free Riders (e.g., Fischbacher and Gächter, 2010; Kurzban and Houser, 2005). In a next step, we investigate whether this aggregate stability is caused by stability of the contribution patterns at the individual level.

3.3 Individual stability

In order to test for individual stability, we first check whether subjects classified into one of the three classes in the first (second) wave remain in that class in the second (third) wave. With 72 subjects in Waves 1 and 2 and 68 subjects in Wave 3, there are 140 opportunities for individual preferences to change type between two times of measurement. In 93 cases (66.4 percent) subjects do not change their type between two points in time. Preference stability is somewhat stronger between Wave 2 and Wave 3 (69.1 percent versus 63.9 percent between Wave 1 and 2).[10] Given that we have three measures for each subject we can also check whether a subject belongs to the same type in all three waves. Exactly half of the 68 subjects we observe three times are classified as the same type in all three waves. 35.3 percent display stable Conditional Cooperation preferences, 10.3 percent stable Free Rider preferences and 4.4 percent are consistently classified as Others. Consequently, underneath a high degree of aggregate stability there are quite some changes going on at the individual level. On the other hand, we can clearly refute the hypothesis that types are random. To do so we simulate the stability of types under the assumption that each player randomly picks a type in Wave 2 and Wave 3 with a probability that equals the observed relative frequency of the three types in our data. In the simulation we observe

[10] This result is remarkably similar to the findings of Muller et al. (2008), who examine 60 subjects playing a comparable game in 5 repetitions within the same experimental session. In their design there are 240 opportunities for individual preferences to change and in 66.7 percent of cases subjects remain in the same category in the next game.

that in 20.6 percent of the cases subjects are of the same type in two adjacent waves (100 runs, standard deviation 2.8 percentage points). The hypothesis of random types can be rejected at any reasonable level of significance.

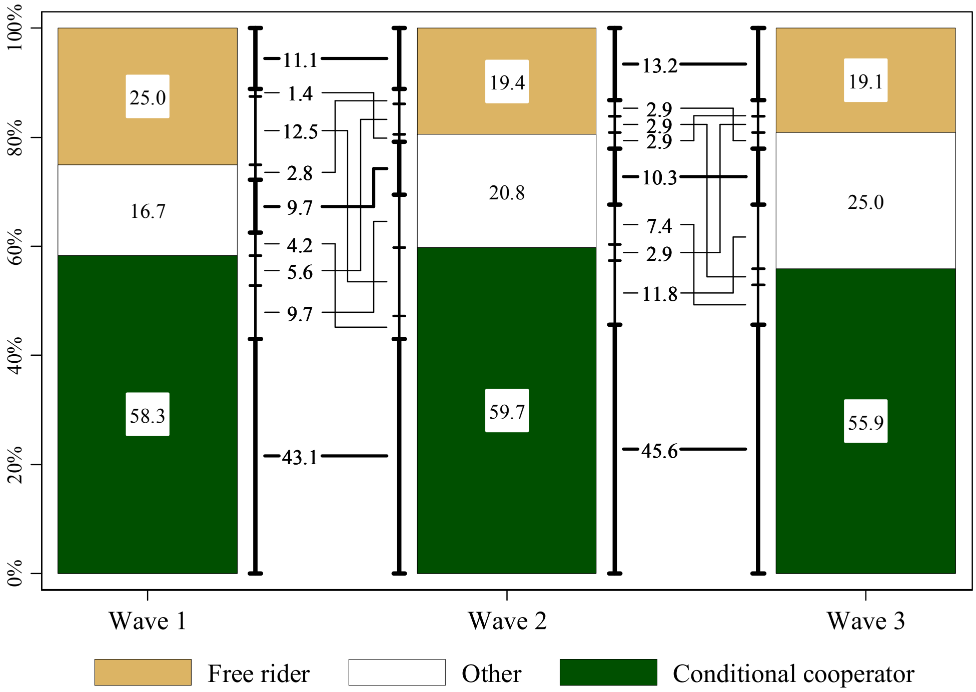


Figure 2: Distribution of types in the three waves and transitions between the times of measurement. All numbers denote percentages of the whole population in a given wave.

What are the patterns of the type changes? The three bars in Figure 2 show the fractions of Conditional Cooperators, Free Riders and Others at the three points in time. The numbers between the bars show the fraction of subjects that remains in a class or changes to another class in the subsequent wave. All numbers indicate the percentage of subjects relative to the whole population. For example, 58.3 percent of the subjects are classified as conditionally cooperative in Wave 1. About three fourths of these (and 43.1 percent of all subjects) remain Conditional Cooperators in Wave 2, while 5.6 percent of all subjects change from Conditional Cooperator to Free Rider from Wave 1 to Wave 2 and 9.7 percent of all subjects move from Conditional Cooperator to the category Others. Bold lines correspond to observations that remain in a category.

Overall, Figure 2 shows that the type of the Conditional Cooperator is clearly the most stable category, with three fourth of the subjects remaining in that class between two waves.

Surprisingly, the category of Free Riders seems to be rather unstable, especially between Wave 1 and Wave 2. Among the 25 percent Free Riders in Wave 1, more than half change their type in Wave 2, most of those who change join the conditionally cooperative group. The fraction of Free Riders which stay in that category increases to almost 70 percent when we compare Wave 2 and Wave 3. However, the group of conditionally cooperative subjects is still more stable with about 77 percent of the subjects remaining in the same class between Wave 2 and Wave 3.

In conclusion, our results indicate that the very high aggregate stability of the distribution of types is not accompanied by an equally strong stability of types at the individual level. On the other hand, the pattern we observe reflects certainly some non-trivial degree of individual stability. Overall, this finding corresponds with results of other recent studies. Blanco et al. (in press) and Brosig et al. (2007), for example, examine the consistency of individual and aggregate behavior with Fehr and Schmidt's (1999) model of inequality aversion. Both studies find that the model has predictive power at the aggregate level but less so at the individual level. In a next step, we will use our questionnaire data to explore the impact of the Big-Five personality dimensions on preference heterogeneity and preference stability.

3.4 Instrument check personality measure

We first conduct an instrument check to assess whether our personality measure generates results similar to those found in other studies using the same instrument. Table 1 reports the means (ø) and standard deviations (sd) for the five TIPI subscales for the present study and three previous studies employing the TIPI (Donnellan et al., 2006; Ehrhart et al., 2009; Gosling et al., 2003). As Table 1 shows, the personality scores obtained in our study are very similar to those reported in the other three studies, indicating that the TIPI items performed well in our setting.

	Present study (*n*=72)		Gosling et al., 2003 (*n*=1799)		Donnellan et al., 2006 (*n*=329)		Ehrhart et al., 2009 (*n*=902)	
TIPI subscales	ø	sd	ø	sd	ø	sd	ø	sd
Extraversion	4.8	1.7	4.4	1.5	5	1.5	4.8	1.5
Agreeableness	5.3	1.2	5.2	1.1	5.4	1.1	4.9	1.1
Conscientiousness	5.6	1.2	5.4	1.3	5.6	1.2	5.9	1.1
Emotional Stability	4.8	1.4	4.8	1.4	3.2	1.5	4.9	1.4
Openness to Experience	5.9	1	5.4	1.1	5.5	1.1	5.4	1.1

Table 1: Means (ø) and standard deviations (sd) for the five TIPI subscales for the present study and three previous studies employing the TIPI.

3.5 Personality and cooperation preferences

In this section, we investigate whether the Big-Five personality dimensions are related to cooperation preferences. We use a multinomial logit model with the Big-Five personality dimensions as explanatory variables to explain the observed cooperation preference type in Wave 1. Table 2 shows the results of these estimates. The omitted case is the Conditional Cooperator type. Table 2 shows that two Big-Five dimensions are significantly related to the observed type in Wave 1, Agreeableness and Openness to Experience.

Subjects scoring high on Agreeableness are significantly less likely to be classified as Free Riders compared to Conditional Cooperators. The same is true for the category Others. Furthermore, scoring high on Openness to Experience is associated with a higher probability of being classified as a Free Rider compared to being classified as a Conditional Cooperator.

	Multinomial Logit	
	Free Rider	Other
Extraversion	-0.363	0.569
	(0.363)	(0.429)
Agreeableness	-0.898**	-0.924**
	(0.390)	(0.420)
Conscientiousness	-0.020	0.384
	(0.329)	(0.415)
Emotional Stability	0.076	-0.241
	(0.342)	(0.405)
Openness to Experience	0.782**	-0.235
	(0.395)	(0.360)
Constant	-1.060***	-1.468***
	(0.340)	(0.402)
Model chi-square	20.1	
Prob > chi2	0.028	
Pseudo R-squared	0.146	
N	72	

Table 2: Multinomial Logit estimate for type of conditional cooperation scheme in Wave 1. Omitted case is Conditional Cooperator. Independent variables are the five personality dimensions, demeaned and standardized. Standard errors in parentheses; ** denotes significance at 5 percent.

What is the size of the personality effect? We use the multinomial logit model to predict probabilities of the three types dependent on the Agreeableness score. Panel A in Figure 3 presents the results. The horizontal axis shows the range of Agreeableness scores we observe in our experimental data. To facilitate the interpretation we de-mean and standardize the values of Agreeableness. The size of the estimated effect is quite remarkable. A subject scoring 2.7 standard deviations lower than the average on Agreeableness has an estimated probability of about 13 percent to be classified as Conditional Cooperator. For the average subject the probability is about 63 percent and increases up to 86 percent for a subject scoring 1.4 standard deviations higher than the average on Agreeableness. Conversely, the estimated probability of being classified as Free Rider drops from 51 percent to a mere 8 percent if we compare the least to the most agreeable subject in our sample.

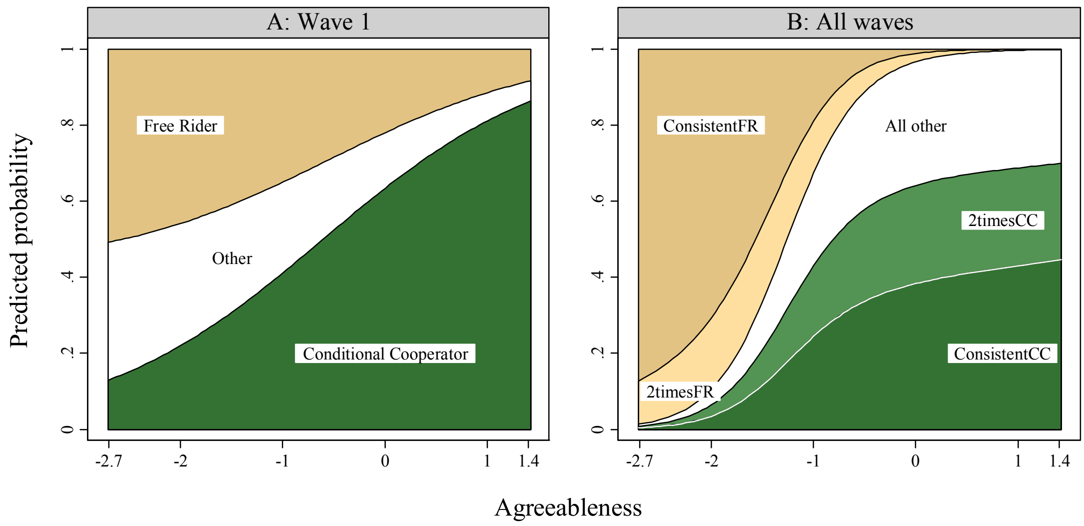


Figure 3: Panel A: Estimated probability of being classified as either 'Conditional Cooperator', 'Free Rider', or 'Other' dependent on the Big-Five dimension Agreeableness (demeaned and normalized) in Wave 1. Panel B: Estimated probability of being classified into one of five classes during all three waves of the experiment dependent on Agreeableness. 'Consistent FR (CC)' are subjects classified as Free Rider (Conditional Cooperator) in all three waves; 2timesFR (CC) are subjects classified as Free Rider (Conditional Cooperator) in two out of the three waves; 'All other' denotes all remaining patterns.

What about time effects? We can use our personality measures elicited in the experiment of Wave 1 to predict the type in Wave 2, or Wave 3. It turns out that the influence of Agreeableness on the types Free Rider and Conditional Cooperator is strong and significant in all three waves. The positive connection between Agreeableness and the type Other observed in Table 2 is, however, not present in the subsequent waves. Likewise, the positive influence of

Openness to Experience on the category Free Rider is insignificant in Wave 2 and only marginally significant in Wave 3.

How are personality measures associated with preference stability over time? Given that we have three types of cooperation preferences and three times of measurement, there are ten possible combinations of types. A subject who is classified as Conditional Cooperator in all three waves is categorized as *ConsistentCC*. Likewise, we observe three times Free Riders (*ConsistentFR*). Subjects who are classified as either Conditional Cooperator or Free Rider in two out of the three waves are categorized as *2timesCC* and *2timesFR* respectively. All remaining combinations are in the same group (*All other*).[11] We run a multinomial logit model for these five categories. Like before, only the personality dimension Agreeableness shows significant results. Panel B of Figure 3 shows the predicted probabilities dependent on Agreeableness. Clearly, the probability of *ConsistentFR* is now very strongly connected to the Agreeableness measure. It ranges from 87 percent for the lowest observed Agreeableness to virtually zero for an Agreeableness score that is a standard deviation above the mean. The category *2timesFR* shows a similar pattern. On the other end of the scale we observe that the probability of *ConsistentCC* rises from virtually zero to 43 percent for the highest Agreeableness score observed in our sample. The category *2timesCC* displays a similar pattern. Finally, unlike in Panel A, the remaining observations are more likely to be associated with higher scores of Agreeableness.

Taken together, our results suggest that the personality dimension Agreeableness can contribute to our understanding of the psychological foundations of cooperation preference heterogeneity and stability. Concerning preference heterogeneity, we find that Free Riders and Conditional Cooperators are characterized by different levels of Agreeableness. Regarding preference stability, we find that higher (lower) levels of Agreeableness are associated with more stable preferences for conditional cooperation (free riding).

3.6 Predicting types

In the beginning of the paper we have argued that economic theory is about making predictions and that making predictions is enormously facilitated if peoples' tastes are stable. So far we have shown that cooperation preferences are stable at the aggregate level. If we are interested in

[11] We do not introduce a separate category for subjects classified as Other in all three waves because (i) only three subjects do so, and (ii) it is unclear how we should interpret a pattern that is consistently inconsistent.

aggregate outcomes, a prediction that simply extrapolates the observed share would be fairly accurate. At the individual level we find that about one third of the subjects changes preference type between two waves. What does this mean for the predictive power of potential models predicting individual types? In the following, we use the data from Wave 1 and Wave 2 to predict a subject's type in Wave 3. We then evaluate the predictions with the data from Wave 3. We calculate the accuracy for different methods of predicting individual types in Wave 3, making increasingly use of the information we have about our subjects from Waves 1 and 2. The accuracy of a model is simply the number of subjects for whom we correctly predict the type in relation to all subjects.

A first benchmark is a prediction that completely ignores preference stability and does not take information from Waves 1 and 2 into account. Given that there are three types of cooperation preferences we would make the uneducated guess of predicting each of the three types with equal probability. In this case we would predict the correct type in 33.3 percent of the cases (in expectation).[12]

If we only consider information about the aggregate outcome in Waves 1 and 2 to predict individual behavior in Wave 3, our best estimate for a subject's type would be the type that occurs most frequently, which is the Conditional Cooperator. In this case we would be correct for all Conditional Cooperators and wrong for all other types. Thus our prediction would be correct in 55.9 percent of the cases, which is the fraction of Conditional Cooperators observed in Wave 3.

In a next step we make use of individual information from Waves 1 and 2. We run a multinomial logit model using the data from Wave 2 with two dummies for a subject's type in Wave 1 as explanatory variables ($CC_{t\text{-}1}$ and $FR_{t\text{-}1}$). This model allows predicting probabilities for the types in Wave 3, dependent on their type in Wave 2. We then assign to each individual the type with the highest predicted probability and list our observations in a so-called confusion matrix. Model 1 in Table 3 shows the results. The model predicts the conditional cooperative type for 54 subjects and the type Other for the remaining 14 subjects observed in Wave 3. None of the subjects is predicted to fall into the category Free Riders. The reason for this is that there is a substantial fraction of the Free Riders in Wave 1 that move to another category in Wave 2 (see Figure 2). The nine numbers in the center of the table for Model 1 demonstrate the accuracy of

[12] If p_i is the predicted probability of type i ($i = 1,\ldots,n$) and q_i is the true probability, then the prediction is correct with probability $\sum p_i q_i$. If all p_i are identical then the sum reduces to $1/n$. Thus for three types the hit rate is one third, independent of the true probabilities.

the prediction. Among the 54 subjects who are predicted to be Conditional Cooperators, 33 turn out to be of that type in Wave 3, 11 are Free Riders and 10 are Others. Bold numbers on the main diagonal are the observations for which the prediction is correct. The predictive success of the model is simply the number of correct predictions in relation to all observations in Wave 3 (68). The rightmost column shows the accuracy of Model 1, which is in this case 58.8 percent.

In Model 2 we use the Big-Five personality dimensions to estimate a multinomial logit model with a data from Wave 2. This model predicts only the types Conditional Cooperator and Free Rider. This reflects the results discussed in the previous section that personality measures are informative for the distinction between Free Riders and Conditional Cooperators, but not when it comes to the category Others. Model 2 does not take into account the transition probabilities. The personality measures allow us to predict the correct type in 63.2 percent of the cases in Wave 3.[13]

	Predicted type		Observed type Cond. Coop. 38	 Free Rider 13	 Other 17	Accuracy
Model 1	Cond. Coop.	54	**33**	11	10	
Covariates: $CC_{t\text{-}1}$, $FR_{t\text{-}1}$	Free Rider	0	0	**0**	0	58.8%
	Other	14	5	2	**7**	
Model 2	Cond. Coop.	62	**38**	8	16	
Covariates: Big-Five	Free Rider	6	0	**5**	1	63.2%
	Other	0	0	0	**0**	
Model 3	Cond. Coop.	45	**32**	4	9	
Covariates: $CC_{t\text{-}1}$, $FR_{t\text{-}1}$, Big-Five	Free Rider	9	1	**7**	1	67.6%
	Other	14	5	2	**7**	
Model 4	Cond. Coop.	46	**33**	3	10	
Covariates: $CC_{t\text{-}1}$, $FR_{t\text{-}1}$,	Free Rider	8	0	**8**	0	70.6%
Agreeableness, interaction	Other	14	5	2	**7**	

Table 3: Confusion matrices for four models to predict types in Wave 3 based on observations in Wave 1 and Wave 2.

In Model 3 we combine the personality measures with the lagged type of a subject. This model predicts the occurrence of all three types and achieves an accuracy of 67.6 percent. Finally, in Model 4 we introduce interaction terms between the personality measures and the lagged type. This allows accounting for the fact that personality and the probability of changing

[13] We use the experimental data from Wave 2 in order to estimate with the same dependent variable as in the three other models. Alternatively, we could predict the type using the data from Wave 1 or both waves. A model using only Wave 1 data predicts the occurrence of all three types but results in an identical hit rate. If we use the joint data from Waves 1 and 2 the hit rate reaches 66.2 percent.

type between two points in time are likely to be related. To keep the number of explanatory variables low we only interact the most important personality dimension (Agreeableness) with the lagged type. Model 4 improves the prediction relative to Model 3 and achieves an accuracy of 70.6 percent.

Accuracy might not be the only performance measure we are interested in. An alternative approach is to ask which model performs best in detecting a type (precision). If we look at the type Conditional Cooperator, clearly Model 2 scores best on precision. All 38 Conditional Cooperators observed in Wave 3 are predicted to be of that type. For the type Free Rider Model 4 performs best, identifying 8 out of the 13 we observe in Wave 3. In all models precision in predicting the type Other is relatively low. We conclude that, despite the fact that a substantial fraction of the subjects change their type between waves, information about subjects' past behavior in combination with personality measures is highly informative when we aim at predicting individual types.

4. Conclusions

Economic theory almost always implicitly assumes that preferences are stable. Yet, empirical evidence on the stability of preferences is very limited. Furthermore, economists routinely classify preferences into different types without establishing the underpinnings of preference heterogeneity. The goal of this study was to investigate the stability of cooperation preferences over time and to explore the psychological foundations of preference heterogeneity and preference stability. The main contribution of our study is threefold.

First, we report high aggregate stability of cooperation preferences over time. This finding complements Blanco et al. (in press) who report high aggregate stability of social preferences across different games. Taken together these results suggest that theories of social preferences can "provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances", as stipulated by Friedman (1953, p. 4). Our finding of high aggregate stability is also interesting, as it suggests that the heterogeneity of other-regarding preferences reported in previous studies is a temporally stable phenomenon. In line with earlier research, we find that the most important preference types are Conditional Cooperators and Free Riders, which has important theory and policy consequences (see Gächter, 2007 for a comprehensive discussion).

Second, we provide evidence on the stability of individual preferences over time. Individual-level preference stability is an important prerequisite of signaling games or reputation mechanisms. Learning another players' type through signaling makes only sense if there is such a thing as a stable type (see e.g. the models by Ellingsen and Johannesson, 2008, or Levine, 1998). We find that individual-level stability of cooperation preferences over time is lower than aggregate stability. Still, individual stability is high enough so that knowledge about an individual's past behavior offers substantial informational content when it comes to predicting future behavior. Our results also highlight another interesting aspect. While previous work has focused on one dimension of preference heterogeneity, i.e., heterogeneity in terms of preference types, our findings suggest there is a second dimension researchers should pay attention to, i.e., heterogeneity in terms of preference stability. Some individuals seem to have stable preferences, while others do not. However, given the small number of studies on the within-subject stability of cooperation preferences, more research is certainly needed for a proper understanding of this phenomenon.

Third, we provide evidence on the psychological foundations of cooperation preferences. Our results suggest that Free Riders and Conditional Cooperators identified in our experiments are indeed subjects with different personalities, as measured by the Big-Five dimension 'Agreeableness'. The relation between personality and cooperation preferences is particularly strong for consistent Conditional Cooperators and consistent Free Riders. Information about subjects' personality traits improve the accuracy of our predictions for Wave 3 types beyond what was possible relying only on behavioral data.

In conclusion, our results indicate that there is heterogeneity of individual behavior in terms of both preference types and preference stability, which can be linked to individual differences in personality. Applying an individual differences approach to investigate the importance of dispositional factors may accordingly prove useful for our understanding of the heterogeneity and stability of cooperation preferences. We believe that future theoretical and empirical research on social preferences would benefit substantially from incorporating personality and attitude variables. We therefore call for more interdisciplinary research that spans the behavioral economics and personality psychology fields to cross-fertilize insights from both disciplines.

REFERENCES

Ameriks, J., Caplin, A., Leahy, J., Tyler, T., 2007. Measuring self-control problems. American Economic Review 97, 966-972.

Andersen, S., Harrison, G. W., Lau, M. I., Rutström, E. 2008. Lost in state space: Are preferences stable? International Economic Review 49, 1091-1112.

Andreoni, J., Miller, J., 2002. Giving according to Garp: An experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.

Ashton, M.C., Paunonen, S.V., Helmes, E., Jackson, D.N., 1998. Kin altruism, reciprocal altruism, and the Big Five personality factors. Evolution and Human Behavior 19, 243-255.

Bartling, B., Fehr, E., Marechal, M.A., Schunk, D., 2009. Egalitarianism and competitiveness. American Economic Review 99, 93–98.

Becker, G. S., 1976. The economic approach to human behavior. University of Chicago Press, Chicago.

Ben-Ner, A., Kong, F., Putterman, L., 2004a. Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. Journal of Economic Psychology 25, 581-589.

Ben-Ner, A., Putterman, L., Kong, F., Magan, D., 2004b. Reciprocity in a two-part dictator game. Journal of Economic Behavior & Organization 53, 333-352.

Blanco, M., Engelmann, D., Normann, H.-T., in press. A within-subject analysis of other-regarding preferences. Games and Economic Behavior.

Boone, C., De Brabander, B., Carree, M., de Jong, G., van Olffen, W., van Witteloostuijn, A., 2002. Locus of control and learning to cooperate in a Prisoner's Dilemma game. Personality and Individual Differences 32, 929–946.

Boone, C., De Brabander, B., van Witteloostuijn, A., 1999. The impact of personality on behavior in five prisoner's dilemma games. Journal of Economic Psychology 20, 343-377.

Brandstätter, H., Königstein, M., 2001. Personality influences on ultimatum bargaining decisions. European Journal of Personality 15(S1), S53-S70.

Brandts, J., Charness, G., 2009. The strategy method: A survey of experimental evidence. Mimeo, Department of Business Economics U. Autònoma de Barcelona and Department of Economics, UC Santa Barbara.

Brosig, J., Riechmann, T., Weimann, J., 2007. Selfish in the end? An investigation of consistency and stability of individual behavior. MPRA Paper 2035, University Library of Munich.

Burlando, R. M., Guala, F., 2005. Heterogeneous agents in public goods experiments. Experimental Economics 8, 35-54.

Buss, D. M., 1991. Evolutionary personality psychology. Annual Review of Psychology 42, 459-491.

Chmielewski, M., Watson, D., 2009. What is being assessed and why it matters: The impact of transient error on trait research. Journal of Personality and Social Psychology 97, 186-202.

Conley, J. J., 1985. Longitudinal stability of personality traits: A multitrait–multimethod–multioccasion analysis. Journal of Personality and Social Psychology 49, 1266–1282.

Costa, P. T., McCrae, R. R., 1992. Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO FFI) professional manual. Psychological Assessment Resources, Odessa, FL.

Cox, J. C., 2004. How to identify trust and reciprocity. Games and Economic Behavior 46, 260–281.

De Oliveira, A. C. M, Croson R. T. A., Eckel, C., 2009. Are Preferences Stable Across Domains? An Experimental Investigation of Social Preferences in the Field. CBEES Working Paper #2008-3.

Donnellan, M. B., Oswald, F. L., Baird, B. M., Lucas, R. E., 2006. The mini IPIP scales: Tiny yet effective measures of the Big Five factors of personality. Psychological Assessment 18, 192–203.

Dufwenberg, M., Kirchsteiger, G. 2004. A theory of sequential reciprocity. Games & Economic Behavior 47, 268-298.

Ehrhart, M. G., Ehrhart, K. H., Roesch, S. C., Chung-Herrera, B. G., Nadler, K., Bradshaw, K., 2009. Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. Personality and Individual Differences 47, 900-905.

Ellingsen, T., Johannesson, M., 2008. Pride and prejudice: The human side of incentive theory. American Economic Review 98, 990-1008.

Falk, A., Fischbacher, U., 2006. A Theory of Reciprocity. Games and Economic Behavior 54, 293-316.

Fehr, E., Schmidt, K. M., 1999. A theory of fairness, competition and cooperation. Quarterly Journal of Economics 114, 817- 868.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10, 171-178.

Fischbacher, U., Gächter, S., 2009. The behavioral validity of the strategy method in public good experiments. CeDEx Discussion Paper No. 2009-25, University of Nottingham.

Fischbacher, U., Gächter, S., 2010. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. American Economic Review 100, 541-556.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Economic Letters 71, 397–404.

Fisman, R., Kariv, S., Markovits, D., 2007. Individual preferences for giving. American Economic Review 97, 1858-1876.

Friedman, M., 1953. The methodology of positive economics, in: Friedman, M. (Ed.), Essays in positive economics. University of Chicago Press, Chicago, pp. 3-43.

Gächter, S., 2007. Conditional cooperation. Behavioral regularities from the lab and the field and their policy implications, in: Frey, B.S., Stutzer, A. (Eds.), Economics and psychology. A promising new cross-disciplinary field. MIT Press, Cambridge.

Gächter, S., Herrmann, B., Thöni, C., 2010. Culture and cooperation. Philosophical Transactions of the Royal Society B 365, 2651-2661.

Gächter, S., Thöni, C., 2005. Social learning and voluntary cooperation among like-minded people. Journal of the European Economic Association 3, 303-314.

Goldberg, L. R., 1993. The structure of phenotypic personality traits. American Psychologist 48, 26–34.

Gosling, S. D., Rentfrow, P. J., Swann, W. B., 2003. A very brief measure of the Big-Five personality domains. Journal of Research in Personality 37, 504-528.

Graziano, W. G., Eisenberg, N. H., 1997. Agreeableness: A dimension of personality, in: Hogan, R., Johnson, J., Briggs, S. (Eds.), Handbook of personality psychology. Academic Press, London, pp. 795-824.

Herrmann, B., Thöni, C., 2009. Measuring conditional cooperation: A replication study in Russia. Experimental Economics 12, 87-92.

Hirsh, J., Peterson, J. B., 2009. Extraversion, neuroticism, and the prisoner's dilemma. Personality and Individual Differences 46, 254-256.

Horowitz, J. K., 1992. A test of intertemporal consistency. Journal of Economic Behavior & Organization 17, 171–82.

John, O. P., Srivastava, S., 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives, in: Pervin, L.A., John, O.P. (Eds.), Handbook of personality: Theory and research. Guilford Press, New York, pp. 102–138.

Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., Sutter, M., 2008. Conditional cooperation on three continents. Economics Letters 101,175-178.

Koole, S. L., Jager, W., van den Berg, A. E., Vlek, C. A. J., Hofstee, W. K. B., 2001. On the social nature of personality: Effects of extraversion, agreeableness, and feedback about collective resource use on cooperation in a resource dilemma. Personality and Social Psychology Bulletin 27, 289–301.

Kurzban, R., Houser, D., 2001. Individual differences in cooperation in a circular public goods game. European Journal of Personality 15, S37-S52.

Kurzban, R., Houser, D., 2005. Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. PNAS 102, 1803-1807.

Levine, D. K., 1998. Modeling altruism and spitefulness in experiments. Review of Economic Dynamics 1, 593-622.

Lo, A., Repin, D.V., Steenbarger, B.N., 2005. Fear and greed in financial markets: A clinical study of day traders. American Economic Review 95, 352–359.

McCrae, R. R., Costa, P. T., 1990. Personality in adulthood. Guilford Press, New York.

Meier, S., Sprenger, C., 2010. Stability of time preferences. IZA Discussion paper No. 4756.

Muller, L., Sefton, M., Steinberg, R., Vesterlund, L., 2008. Strategic behavior and learning in repeated voluntary contribution experiments. Journal of Economic Behavior and Organization 67, 782-793.

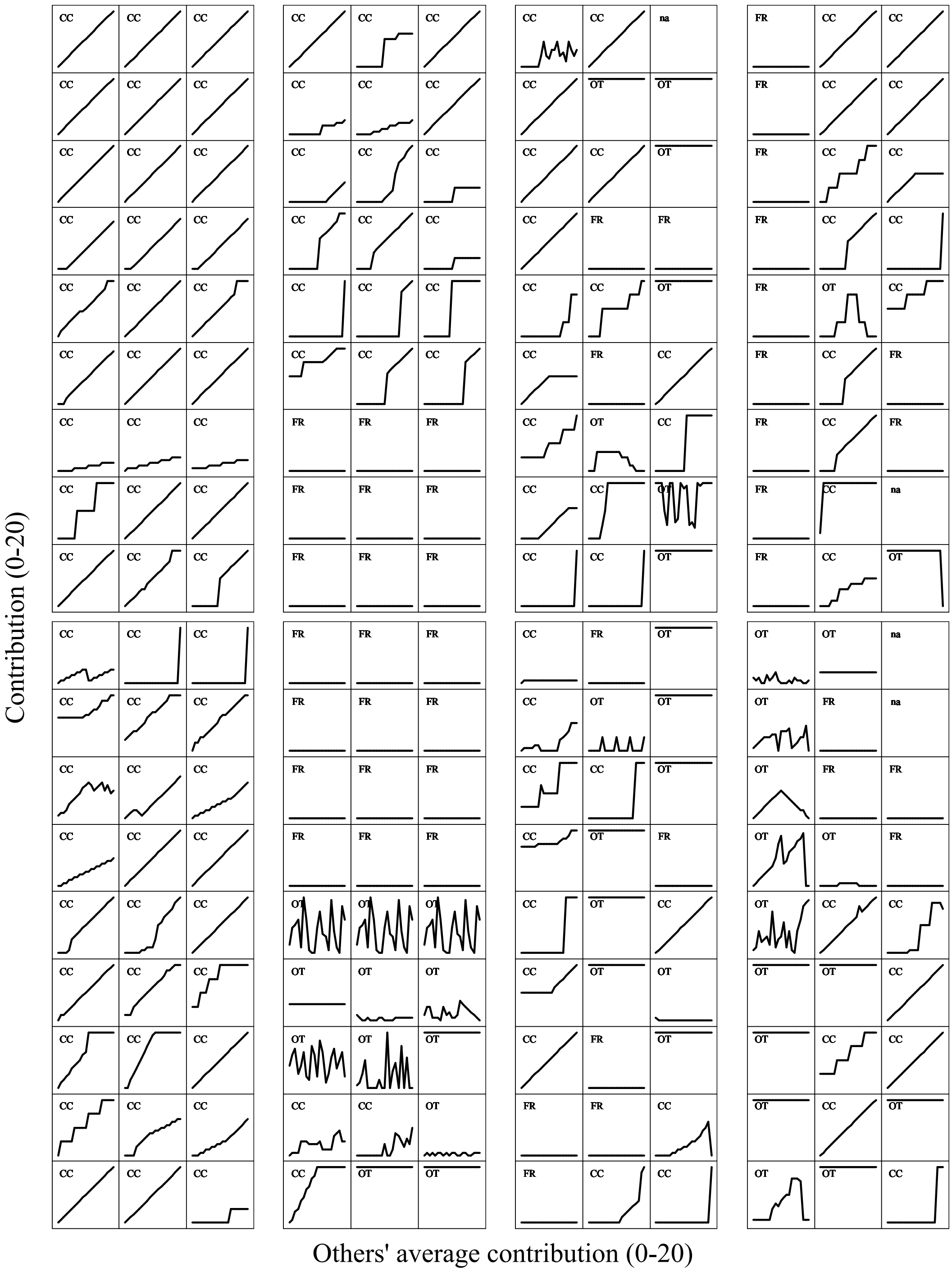
Perugini, M., Tan, J. H., Zizzo, D. J., 2010. Which is the more predictable gender? Public good contribution and personality. Economic Issues 15, 83-111.

Rapoport, A., Suleiman, R., 1993. Incremental contribution in step-level public goods games with asymmetric players. Organizational Behavior and Human Decision Processes 55, 171–194.

Thöni, C., Tyran, J., Wengström, E., 2009. Microfoundations of social capital. Discussion paper no. 09-24, University of Copenhagen.

Zeisberger, S., Vrecko, D., Langer T, in press. Measuring the time stability of Prospect Theory preferences. Theory and Decision.

Appendix A:
Conditional contribution schedules per subject



Appendix A: Individual data. Three horizontally adjacent panels show the conditional cooperation scheme for a subject in Wave 1, Wave 2 and Wave 3 respectively. Letters denote the classification: CC for Conditional Cooperator, FR for Free Rider and OT for other types.

Appendix B:
Instructions (not intended for publication)

If you read the following instructions carefully, you can, depending on your decisions, earn a considerable amount of money. It is therefore very important that you read these instructions with care. The instructions which we have distributed to you, are solely for your private information. It is prohibited to communicate with the other participants during the experiment. Should you have any questions please ask us. If you violate this rule, we shall have to exclude you from the experiment and from all payments.

You will participate in this experiment three times over the course of the next months. The decisions you make in one experiment will have no influence on the outcomes or payoffs of the following experiments. Your decisions in the first experiment will therefore not influence outcomes or payoffs of the second or third experiment and vice versa. In each of the three independent experiments, the participants will be randomly divided into groups of four members. You will therefore be in a group with 3 other participants. The people sitting right next to you are most likely not in your group. The composition of the groups will change by random after each experiment. In each experiment your group will therefore consist of different participants. Nobody will know who is in which group. Neither before, nor after the experiment, will you learn which people are/were in your group. The decisions you made in the first experiment will be unknown to your group mates in the second experiment. The same holds for the third experiment.

This experiment will be conducted under full anonymity. During and after the experiment your identity in the experiment will remain undisclosed. Except us, the experimenters, nobody can associate certain decisions with certain people. During the experiment, your entire earnings will be calculated in Points. At the end of the experiment the total amount of Points you have earned will be paid to you in cash.

The decision situation

You will later learn how the actual experiment will be conducted. We first introduce you to the basic decision situation. At the end of the description of the decision you will find control questions that help you to gain an understanding of the decision situation. You will be a member of a group of 4 people. Each group member has to decide on the division of 20 Points. You can either put these 20 Points into your private account or you can invest them fully or partially into a project. Each Point you do not invest into the project will automatically remain in your private account.

Your income from the private account

For each Point you put into your private account, you will earn exactly one Point. For example, if you put 20 Points into your private account (which implies that you do not invest anything into the project) you will earn exactly 20 Points from the private account. If you put 6 Points into your private account, you will receive an income of 6 Points from the private account. Nobody except you earns something from your private account.

Your income from the project

From the amount you invest into the project each group member will get the same payoff. Of course, you will also get a payoff from the Points the other group members invest into the project. The income for each group member will be determined as follows:

Income from the project = sum of contributions to the project × 0.4

For example, if the sum of all contributions to the project is 60 Points, then you and all other members of your group will get a payoff of 60 × 0.4 = 24 Points from the project. If the four group members together contribute 10 Points to the project, you and all other members of your group will get a payoff of 10 × 0.4 = 4 Points from the project.

Your total income

Your total income in Points is the sum of your income from the private account and your income from the project:

Income from the private account (= 20 – contribution to the project) + Income from the project (= 0.4 × sum of contributions to the project) = Total income

Control questions

Please answer the following control questions. They will help you to gain a better understanding of the calculation of your income that varies with your decision about how you distribute your 20 Points.

1. Each group member has 20 Points. Assume that none of the four group members (including you) contributes anything to the project. What will your total income be? What will the total income of each of the other group members be?

2. Each group member has 20 Points. You invest 20 Points into the project. From the other three members of the group each one contributes 20 Points to the project. What will your total income be? What will the total income of each of the other group members be?

3. Each group member has 20 Points. The other 3 members contribute in total 30 Points to the project.
What will your total income be, if you have – in addition to the 30 Points – invested 0 Points into the project?
What will your total income be, if you have – in addition to the 30 Points – invested 8 Points into the project?
What will your total income be, if you have – in addition to the 30 Points – invested 15 Points into the project?

4. Each group member has 20 Points. Assume that you invest 8 Points into the project.
What will your total income be, if the other group members –in addition to your 8 Points– together contribute 7 Points to the project?
What will your total income be, if the other group members –in addition to your 8 Points– together contribute 12 Points to the project?
What will your total income be, if the other group members –in addition to your 8 Points– together contribute 22 Points to the project?

The experiment

You will now learn how the actual experiment will be conducted. The experiment involves the decision situation we just explained to you. Each participant has to make two types of decisions. In the following we will call them "unconditional contribution" and "contribution table". With the unconditional contribution to the project you have to decide how many of the 20 Points you want to invest into the project. After you have determined your unconditional contribution your second task is to fill in a contribution table. In the contribution table you have to indicate for each possible average contribution of the other three group members (rounded to the next integer) how many Points you want to contribute to the project. You can condition your contribution on the contribution of the other group members. This will be immediately clear to you if you take a look at the following table. In the experiment, this table will be presented to you:

Contribution Table:

0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

The numbers are the possible (rounded) average contributions of the other three group members to the project. You simply have to insert into each entry field how many Points you want to contribute to the project – conditional on the indicated average contribution of the others. You have to make an entry into each entry field. For example, you will have to indicate how much you contribute to the project if the other three contribute on average 0 Points to the project, how much you contribute if the other three contribute on average 1, 2, or 3 Points etc. In each entry field you can insert all integer numbers from 0 to 20.

After all participants of the experiment have made an unconditional contribution and have filled in their contribution table, in each group a random mechanism will select a group member. For the randomly determined subject only the contribution table will be the payoff-relevant decision. For the other three group members that are not selected by the random mechanism, only the unconditional contribution will be the payoff-relevant decision. When you make your unconditional contribution and when you fill in the contribution table you of course do not know whether you will be selected by the random mechanism. You will therefore have to think carefully about both types of decisions because both can become relevant for you. Two examples should make this clear.

EXAMPLE 1: Assume that you have been selected by the random mechanism. This implies that your relevant decision will be your contribution table. For the other three group members the unconditional contribution is the relevant decision. Assume they have made unconditional contributions of 0, 2, and 4 Points. The average contribution of these three group members, therefore, is 2 Points. If you have indicated in your contribution table that you will contribute 1 Point if the others contribute 2 Points on average, then the total contribution to the project is given by 0+2+4+1=7 Points. All group members, therefore, earn 0.4×7=2.8 Points from the project plus their respective income from the private account. If you have instead indicated in your contribution table that you will contribute 19 Points if the others contribute two Points on average, then the total contribution of the group to the project is given by 0+2+4+19=25. All group members therefore earn 0.4×25=10 Points from the project plus their respective income from the private account.

EXAMPLE 2: Assume that you have not been selected by the random mechanism which implies that for you and two other group members the unconditional contribution is taken as the payoff-relevant decision. Assume your unconditional contribution is 16 Points and those of the other two group members are 18 and 20 Points. The average unconditional contribution of you and the two other group members, therefore, is 18 Points. If the group member who has been selected by the random mechanism indicates in her contribution table that she will contribute 1 Point if the other three group members contribute on average 18 Points, then the total contribution of the group to the project is given by 16+18+20+1=55 Points. All group members will therefore earn 0.4×55=22 Points from the project plus their respective income from the private account. If instead the randomly selected group member indicates in her contribution table that she contributes 19 if the others contribute on average 18 Points, then the total contribution of that group to the project is 16+18+20+19=73 Points. All group members will therefore earn 0.4×73=29.2 Points from the project plus their respective income from the private account.

The random selection of the participants will be implemented as follows. Each group member is assigned a group-member number between 1 and 4. The computer will randomly select one participant. This participant will, after all participants have made their unconditional contribution and have filled in their contribution table, throw a 4-sided die. The number that shows up will be entered into the computer. If the selected participant throws the membership number that has been assigned to you, then for you your contribution table will be relevant and for the other group members the unconditional contribution will be the payoff-relevant decision. Otherwise, your unconditional contribution is the relevant decision.

After you made the unconditional contribution and filled in the contribution table we ask you to make some estimates about your peer group members. The other members of your group of four also made a decision about their unconditional contribution. Your task is to estimate the average unconditional contribution to the project (rounded to an integer) of the other members of your group. You will be paid for the accuracy of your estimates. If your estimate is exactly right (that is, if your estimate is exactly the same as the actual average unconditional contribution of the other members of your group), you will get 3 Points in addition to your other income from the experiment. If your estimate deviates by one Point from the correct result, you will get 2 additional Points, a deviation by 2 Points still earns you 1 additional Point. If your estimate deviates by 3 or more Points from the correct result, you will not get any additional points.

Your income in the experiment

Your total income in Points in the experiment is the sum of your income from your private account and the income from the project plus your income from right estimates:

Income from your private account (= 20 – contribution to the project) + Income from the project (= 0.4 × sum of all contributions to the project) + Income from your right estimates = Total income

Control questions

Please answer the following control questions. The following contribution table is filled with hypothetical values:

0	1
1	12
2	8
3	8
4	5
5	9
6	1
7	20
8	6
9	8
10	0
11	19
12	14
13	20
14	11
15	12
16	8
17	1
18	15
19	0
20	17

Assume that you have been selected by the random mechanism. This implies that your relevant decision will be your contribution table. For the other three group members the unconditional contribution is the relevant decision.

1. Assume you filled in the table above and the other three group members made unconditional contributions of 10, 12, and 14 Points. What will be your contribution according to the table above? What will be the total contribution to the project of your group accordingly?

2. Assume you filled in the table above and the other three group members made unconditional contributions of 4, 8, and 18 Points. What will be your contribution according to the table above? What will be the total contribution to the project of your group accordingly?

Assume that you have not been selected by the random mechanism which implies that for you and two other group members the unconditional contribution is taken as the payoff-relevant decision.

3. Assume your unconditional contribution is 6 Points and those of the other two group members are 8 and 10 Points. What will be the contribution of the fourth group member according to the table above? What will be the total contribution of your group to the project accordingly?

4. Assume your unconditional contribution is 0 Points and those of the other two group members are 4 and 5 Points. What will be the contribution of the fourth group member according to the table above? What will be the total contribution of your group to the project accordingly?