ESSAYS ON EXPERIMENTAL AND BEHAVIORAL ECONOMICS

A Dissertation
Presented to
The Academic Faculty

By

Jeongbin Kim
B.A., Yonsei University, 2008
M.A., Seoul National University, 2011

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in the Department of Economics at Brown University

Providence, Rhode Island
May 2017

This dissertation by Jeongbin Kim is accepted in its present form

by the Department of Economics as satisfying the

dissertation requirements for the degree of Doctor of Philosophy.

Date ______________ ____

Pedro Dal Bó, Advisor

Recommended to the Graduate Council

Date ______________ ____

Louis Putterman, Reader

Date ______________ ____

Geoffroy de Clippel, Reader

Approved by the Graduate Council

Date ______________ ____

Andrew G. Campbell, Dean of the Graduate School

VITA

Jeongbin Kim was born in Seoul, South Korea in 1983. He graduated with a Bachelor of Arts in Economics from Yonsei University in 2008. His Master of Arts in Economics was granted from Seoul National University in 2011. After receiving his Ph.D. from Brown in 2017, Jeongbin will join California Institute of Technology as a postdoctoral scholar.

To my family, Jieun and Yu-Joon.

ACKNOWLEDGEMENTS

I am grateful to my advisor, Pedro Dal Bó, for his patient guidance, continuing support, and invaluable scholar's attitude that he has shown to me. All the discussions that we have had over the last six years have shaped and guided my journey to economics. Without him, my doctorate would not have come to completion. I'm very proud of the fact that I'm his first student who is doing infinitely repeated game experiments. I also thank Louis Putterman for his unwavering support and encouragement. His critiques and suggestions for improving my papers made all aspects of my potential as an experimental economist immeasurably better. I especially thank Geoffroy de Clippel for inspiring my job market paper, which is the first chapter of this dissertation.

I am incredibly grateful to my wife, Jieun, and my son, Yu-Joon. Without their sacrifice, love, and support, I would likely not have ended up in graduate school. By observing Jieun's patience, I was able to recognize that time preferences would be really essential in repeated games. I'm looking forward to seeing how patient Yu-Joon will be. I would also like to thank my family (especially my parents in Gimcheon and Seoul, and my brother) for their love and pray. I thank all of you who have prayed for me through my time at Brown.

I give glory to God for everything.

Preface

In this dissertation, I investigate the determinants of cooperation in social dilemma experiments. To study the effects of time preferences and information about others' pro-social behavior on cooperation, I use laboratory and online experiments.

In the first chapter, I explore the effects of time preferences on cooperation in infinitely repeated prisoner's dilemma game experiments. I implement a novel experimental design in which subjects play repeated games in an experimental session in the laboratory, but stage game payoffs are paid over a long period of time. I exogenously vary subjects' discount factors by changing the timing of stage game payoffs (weekly or monthly) and I vary present bias by introducing a delay for the first stage game payoffs. I find that the rate of cooperation is higher when subjects are paid every week, implying that higher discount factors promote higher cooperation. I also find that the rate of cooperation is higher when there is a delay for the first stage game payoffs, suggesting that present bias reduces cooperation.

In the second chapter, I study the relationship between individuals' time preferences and cooperation in infinitely repeated prisoner's dilemma experiments. In Amazon's Mechanical Turk, I implement a novel experimental design in which subjects play one repeated game over weeks - one stage game each week. Results are that, first, consistent with a model of quasi-hyperbolic discounting and its application to repeated games, the degree of present bias is negatively correlated with cooperation. Surprisingly, there is no relationship between discount factors and cooperation. Second, subjects with time consistent preferences are less likely to deviate from their plan of action. Third, subjects with time varying preferences are more likely to break cooperative relationships. Finally, the degree of present bias and the discount factor measured at the beginning of the experiment can predict attrition and the length of participation of subjects in later weeks.

In the third chapter, I employ a two-phase experimental design to study one mecha-

nism of how systematically manipulated beliefs can promote or deter cooperation. I use the decision in the first phase trust game to create five environments that differ in the information subjects have about the trust/trustworthiness of their group members, where they are asked to play a voluntary contribution mechanism (VCM) in the second phase. By exploring correlations between behavior and beliefs in VCM, I conclude that high trusting environments also imply high trustworthiness, both of which positively adjust people's beliefs on the cooperativeness of others, which in turn results in higher contribution. In particular, trustworthiness is predictive of conditional cooperation in VCM, and conditional cooperators raise their contributions when believing the corresponding environment is trustworthy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THE EFFECTS OF TIME PREFERENCES ON COOPERATION: EXPERIMENTAL EVIDENCE FROM INFINITELY REPEATED GAMES

1.1 Introduction

Many situations in the real world involve strategic interactions among people over time. For instance, a team consisting of multiple employees is usually the working unit for projects in firms, and it is often the case that a partnership among two or more parties leads a business. If these interactions within a team or a partnership happen repeatedly and periodically over time, they can be well-approximated by the environment which has been studied in infinitely repeated games.

Contributions in the theory of infinitely repeated games have shown that time preferences (discount factor and present bias) are essential in determining cooperative behavior. For instance, if discount factors are high enough, it is possible to credibly punish opportunistic behavior and reward cooperation (see Fudenberg and Maskin, 1986 and Abreu et al., 1990). Present bias has negative effects on cooperation in repeated prisoner's dilemma games (see Chade et al., 2008). In the literature of infinitely repeated games experiments, many papers try to test for theoretical predictions by changing the continuation probability which corresponds to the discount factor in theory.[1] For example, Dal Bó (2005) shows that the continuation probability matters for cooperation in the infinitely repeated prisoner's dilemma games. However, the causal effects of time preferences of humans on cooperative behavior in repeated games remain to be studied.

A couple of studies try to answer the question of whether individuals' time preferences

[1]Infinitely repeated games are induced in the laboratory by using the probability of continuation which determines a length of a repeated game (see Roth and Murnighan, 1978). After each stage game, there is at least one more stage game with a fixed probability, and this fixed probability is known to all subjects. See Dal Bó and Fréchette (forthcoming) for a survey of this literature.

are correlated with cooperation in repeated games.[2] Davis et al. (2016) measure and relate subjects' discount factors to their behavior in repeated games and find no evidence of robust correlations. This result is not surprising because, as pointed out in Kim (2016), the conventional way of inducing repeated games in the laboratory should make time preferences irrelevant for cooperation. The reason is that, as payoffs for all rounds are usually paid at the end of the experiment, there is no time horizon over which payoffs from stage games need to be discounted.

Taking this into account, Kim (2016) implements a novel experimental design which incorporates a time horizon into a repeated game. In Amazon's Mechanical Turk, subjects play one repeated prisoner's dilemma game over weeks - one stage game each week. He finds that the degree of present bias is negatively correlated with cooperation, but discount factors have no correlation with cooperation.[3] Although this may be the most intuitive design for implementing a repeated game over time, some limitations may exist. First, as shown in Dal Bó and Fréchette (2011), subjects' experience with repeated games plays an important role, so learning may matter for subjects to find their optimal behavior consistent with their time preferences. Second, a possibility of attrition that a counterpart may not return to the experiment might influence behavior. Third, a correlation between time preferences and cooperation does not imply a causal effect of time preferences on cooperation as subjects with different time preferences may also differ in other dimensions.

To deal with all of these problems, this paper introduces a novel approach that is easy to implement in the laboratory and allows us to exogenously change the subjects' time preferences. In this experiment, subjects play 20 repeated games and one repeated game is randomly selected for payment. For the selected repeated game, subjects receive

[2] Several papers investigate correlations between other personal characteristics and cooperation in repeated games. See Sabater-Grande and Georgantzis (2002), Proto et al. (2014), and Davis et al. (2016) for risk aversion. See Proto et al. (2014) for intelligence. For social preferences, see Dreber et al. (2014).

[3] The experimental design in Kim (2016) allows for examining the effects of dynamically changing preferences on behavior. For instance, he finds that time consistent subjects are less likely to deviate from their plan of action, and subjects with time varying preferences are more likely to break cooperative relationships.

stage game payoffs over time. In a Weekly treatment, subjects receive payoffs for the first round on the same day that the session was conducted. Then, they receive payoffs for the following stage games once a week. In a Monthly treatment, subjects receive stage payoffs once a month.

Changing the timing of stage game payoffs exogenously varies subjects' discount factors with which they evaluate payoffs. Subjects in the Weekly treatment should have higher discounting factors than subjects in the Monthly treatment. This enables us to examine the causal effects of discount factors on cooperation.

In addition, we have a Delay treatment in which there is a month of a front end delay for the first round payoffs. The introduction of a front end delay should eliminate present bias as all payoffs are in the future. Comparing behavior in these two treatment with monthly payments enables us to examine the causal effects of present bias on cooperation.

We confirm that the average discount factor over a week is higher than that over a month by eliciting subjects' discount factors over a week and a month. Given this, we find strong evidence of the effects of time preferences on cooperation, which is consistent with the general results in theory. First, the rate of cooperation is higher in the Weekly treatment than in the Monthly treatment. This implies that the more patient the subjects are, the higher the rate of cooperation is. Second, subjects cooperate significantly more in the Delay treatment than subjects in the Monthly treatment. This confirms the negative effect of present bias on cooperation. In both comparisons, as subjects gain experience, the differences of cooperation rates across the two treatments get more evident, supporting the argument that learning matters for cooperation in repeated games.

1.2 Experimental Design

The experiments consist of two phases: a time preferences elicitation phase and an infinitely repeated games phase. At the beginning of the experiment, subjects receive the instructions for phase 1, indicating that phase 2 will follow. At this point, however, they are not informed about what they will be asked to do in phase 2. The information about payoff realizations for phase 1 is presented at the end of the experiment to avoid the possibility that a realized outcome in phase 1 may affect subjects' behavior in phase 2.

To control for different transaction costs between immediate and future payments, all payments from the two phases are paid through VENMO, which is an online website and a mobile application for money transfers.[4] VENMO is widely used among Brown undergraduates. The fact that payments from this experiment are paid through VENMO is clearly announced in the invitation email for recruiting subjects.

1.2.1 Time Preferences Elicitation

In phase 1 we elicit subjects' time preferences.[5] We use the Random Binary Choice (RBC) mechanism which is procedurally identical to the Becker-DeGroot-Marschak (BDM) mechanism.[6] Subjects are asked to make decisions for 8 blocks. In each block, there are 2,000 questions (or rows) in each of which subjects are asked to choose between Option A (sooner payment) and Option B (later payment). For example, in one block, the amount for the sooner payment is fixed as \$8.00 and the amount for the later payment increases from \$0.01 to \$20.00 with \$0.01 increments between questions.

Instead of having to answer all the questions, subjects are asked to state the value at which they switch from the option with a fixed amount to the option with varying values. At the value stated and above, all the options with varying amounts are chosen. The 8 blocks differ in timings of sooner and later payments and in whether subjects are asked to switch from Option A to Option B, or vice versa. In particular, subjects are asked to decide between: (1) payment today and payment in 1 week, (2) payment today and

[4] To pay subjects through VENMO we need to have subject' name, email address, or phone number. We ask subjects about these in the questionnaire at the end of the experiment.

[5] A growing body of literature has studied how we can measure individuals' time preferences and whether measured time preferences are correlated with behavior over time. See Frederick et al. (2002) for a critical review of early attempts to measure time preferences, and Urminsky and Zauberman (2016) for a recent survey.

[6] See Azrieli et al. (2016) for the incentive compatibility of the RBC mechanism. Truth telling is a dominant strategy for this mechanism.

payment in 1 month, (3) payment in 1 month and payment in 1 month and 1 week, and (4) payment in 1 month and payment in 2 months. For the 4 blocks, the sooner option pays \$8.00, and the later option pays \$0.01 for the question 1 and the amount increases with the increment of \$0.01, reaching \$20.00 at the question 2,000. For other 4 blocks, the later option pays \$20.00 and the amount for the sooner option varies between \$0.01 and \$20.00.[7] To make subjects' decisions incentive compatible, only one question in one of the blocks will be randomly selected at the end of the experiment. Depending on the realized outcome, subjects will be paid through VENMO on the designated date.

1.2.2 Infinitely Repeated Prisoner's Dilemma

In phase 2, we induce infinitely repeated games by following the conventional experimental protocol in the literature.[8] A repeated game (or a match) consists of a sequence of stage games (or rounds), and the length of a match is determined by a random termination rule. After each round, there is a 75% probability that the match will continue for at least another round, with a 25% probability that the match will be terminated after that round. The stage game is a prisoner's dilemma and Table 2.1 presents the stage game payoffs for all treatments. Subjects play 20 matches and are randomly matched with another subject after the end of each match. One of the matches is randomly selected for payment at the end of the experiment.

Table 1.1: Stage game payoffs

	The other's choice	
Your choice	1	2
1	\$4.00, \$4.00	\$1.00, \$5.00
2	\$5.00, \$1.00	\$2.00, \$2.00

A novelty of the experimental design is that subjects receive their payoffs for each round over time. In particular, there are two frequencies of interactions over which payoffs for

[7] The instructions and the screen shots are available in the Appendix.

[8] See Dal Bó and Fréchette (2016) for a survey of the literature.

each round will be paid: weekly and monthly. In the "Weekly" treatment, subjects receive payoffs for the first round on the same day that the session is conducted. Then, they receive payoffs for the second round in one week, payoffs for the third round in two weeks, and so on. In the "Monthly" treatment, subjects receive payoffs for the first round on the same day that a session is conducted and receive payoffs for the following rounds once a month. In each session, only one of two frequencies of interactions is implemented to pay subjects, and payoffs and the frequency of interactions are commonly known to subjects.

In addition, we have one more treatment with monthly payments, in which there is a month of a delay for payoffs for the first round.[9] In the "Delay" treatment, payoffs for the first round are paid to subjects in 1 month from the date on which a session is conducted, payoffs for the second round in 2 months, and so on. Taken together, we have 3 treatments in total, each of which has 4 sessions. To control for the possibility that different sequences of rounds realized for matches may have different effects on the evolution of cooperation, we randomly create 4 different sets of sequences of rounds for 20 matches and apply each set of sequences to one session in each treatment. Therefore, the nth session in each treatment share the same sequences of rounds for 20 matches. Table 1.2 summarizes the information of all treatments. The instructions for all of the treatments are available in the Appendix.

Table 1.2: Treatments information

Treatment	Frequency	Delay	Prob. of continuation	# of sessions
Weekly	1 week	No	0.75	4
Monthly	1 month	No	0.75	4
Delay	1 month	1 month	0.75	4

[9] A delay for the first round payoffs refers to the front end delay. Introducing a front end delay enables decisions for the first round to happen in the future, and hence, to be free from influences of present bias.

1.3 Research Questions

1.3.1 Frequency of Interactions, Patience, and Cooperation

The main focus of this experiment is whether implementing two different frequencies of interactions can allow us to have subjects with different round to round discount factors. In other words, we expect that on average, subjects in the Weekly treatment will have higher discount factors than subjects in the Monthly treatment. Eliciting subjects' time preferences over weekly and monthly time horizons in phase 1 enable us to investigate whether the discount factors for weekly and monthly frequencies of interactions are different. Therefore, the following question needs to be answered first.

Question 1. *Are the discount factors for the weekly frequency of interactions higher than the discount factors for the monthly frequency of interactions?*

Previous works in the theory of infinitely repeated games have shown that discount factors play an important role in determining whether cooperation can be supported as an equilibrium outcome. Fudenberg and Maskin (1986) show that individually rational payoffs can be supported in a subgame perfect equilibrium if the discount factor, δ, is sufficiently high. Abreu et al. (1990) show that the set of subgame perfect equilibrium payoffs expands in δ.[10] In this experiment, weekly and monthly frequencies of interactions will make it possible that round payoffs can be discounted relatively less in the Weekly treatment and more in the Monthly treatment.[11] Taken together, this leads us to the following question.

[10]Under the assumptions of common knowledge about payoffs in Table 2.1 and a common discount factor, the threshold of a discount factor, δ, over which cooperation can be supported as an equilibrium outcome is 0.33. Accounting for the continuation probability of 0.75, the threshold of an actual discounting factor would be $0.44 (= 0.33/0.75)$.

[11]Note that subjects' discount factors in this experiment might be different. Lehrer and Pauzner (1999) show that the set of feasible payoffs can be larger than the convex hull of the stage game payoffs as players can be better off by trading payoffs over time. Moreover, subjects may know the discount factors of other subjects. However, the characterization of the set of equilibrium payoffs under incomplete information of discount factors needs to be studied.

Question 2. *Do subjects cooperate more in the Weekly treatment than in the Monthly treatment?*

Since there exists ample experimental evidence that a substantial proportion of people have present bias, it is important to examine the effect of present bias on cooperation. Chade et al. (2008) study infinitely repeated games under quasi-hyperbolic discounting. In general, they show that the set of equilibrium payoffs does not increase in δ and β. However, they prove that for a class of games including the prisoner's dilemma in which the minimax point of the stage game coincides with a Nash equilibrium, with δ (β) fixed, the set of equilibrium outcomes expands as β (δ) increases.[12] This may imply that the existence of present biased subjects could have have a detrimental effect on cooperation.[13] Comparing cooperation in the Monthly treatment with cooperation in the Delay treatment will shed light on the following question.

Question 3. *Do subjects cooperate more in the Delay treatment than in the Monthly treatment?*

1.3.2 Who is going to cooperate?

Elicited subjects' time preferences in phase 1 can be used to check whether there is correlations between time preferences and cooperation. It may be natural to expect that more patient and less present biased subjects are more likely to cooperate. However, a multiplicity of equilibria may make it hard to draw clear predictions. In this experiment, the threshold over which cooperation can be supported as an equilibrium outcome is 0.44, so most subjects may have weekly and monthly discount factors higher than this threshold. Given this, cooperation and defection can then be supported as equilibrium

[12] Chade et al. (2008) assume $\beta \leq 1$. Whether the same results can hold when there is no restriction on β needs to be studied.

[13] Kim (2016) show that subjects with present bias are less likely to cooperate than subjects without present bias. The possibility that other subjects may be present biased can also affect subjects' beliefs about others' cooperativeness.

outcomes, and whether subjects' time preferences can guide them to coordinate on some specific equilibrium outcomes is a challenging problem. Therefore, we will try to answer the following question.

Question 4. *Are subjects' β and δ correlated with cooperation?*

1.4 Results

We conducted 12 sessions (4 sessions for each treatment) between September and October 2016. A total of 226 Brown University students participated in the experiments, with an average of 18.83 subjects per session, a maximum of 24 and a minimum of 14. The average earning for subjects is $33.03, with a maximum $57 and a minimum $14. The average number of rounds per repeated game was 4.2, with a maximum of 17 and a minimum of 1. We used z-Tree (Fischbacher, 2007) to program our experiment.[14]

1.4.1 The distribution of Discount factors and Present bias

The parameters of time preferences of interest are β and δ.[15] What we can infer from subjects' decisions for each block may be their intertemporal rate of substitution between two different time points. Estimating precise point estimates of β and δ based on structural assumptions and information about subjects' utility function and liquidity constraints is not a primary purpose of this paper.[16] Instead, we will put more rigorous efforts toward assessing whether subjects have different β and δ depending on the frequency of interactions.

[14]For inducing repeated games, we used a modified version of the z-Tree code which was originally developed by Sevgi Yuksel and Emanuel Vespa. The original z-Tree code was used in Yuksel and Fréchette (2016) and Vespa (2015).

[15]In a model of quasi-hyperbolic discounting, a person with $\beta - \delta$ preferences evaluates a stream of payoffs with the sequence of quasi-hyperbolic discount, 1, $\beta\delta$, $\beta\delta^2$,.... If β equals 1, this model is equivalent to the standard model of exponential discounting. If $\beta < (>) 1$, it can capture the notion of present (future) bias. See Laibson (1997).

[16]See Andersen et al. (2008) and Andreoni and Sprenger (2012) for recent developments in measuring time preferences. See also Dean and Sautmann (2014) for the discussion about accounting for financial shocks on liquidity constraints.

We simply define and estimate δ from decisions that are made on blocks which do not include the option that would pay subjects today. We can then estimate β from the decisions in other blocks.[17] The 8 blocks of decisions allow us to have 2 measures of each weekly β and δ, and monthly β and δ. Then, we take the average of these parameters and have β_w, β_m, δ_w and δ_m.[18]

Table 1.3: Description of β and δ

Parameters	Mean	S.D.	Wilcoxon Signed Rank	Spearman's ρ	Present (future) bias (%)	# of Obs.
β_w	0.977	0.207	-0.431	0.236***	33.50 (23.30)	206
β_m	1.007	0.198			30.10 (24.76)	206
δ_w	0.902	0.139	5.933***	0.771***		206
δ_m	0.873	0.153				206

Notes:
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 1.3 presents the description of these 4 parameters. First, while β_w is slightly less than 1, β_m is slightly above 1. However, these differences are not significant assessed by the Wilcoxon Signed Rank test, implying that subjects do not have a tendency that β_m is higher than β_w.[19] These two measures are significantly correlated, but the coefficient is not high. For the discount factors, we have clearer patterns. The average discount factor over a week (δ_w) is significantly higher than the average discount factor over over a month (δ_m). The significance level is quite robust that if we use the paired t-test, the result does not change. Also, δ_w and δ_m and highly and significantly correlated.[20] Moreover, we find

[17]For example, we can compute a weekly discounting factor from the decisions between sooner options that will pay in 1 month and later options that will pay in 1 month and 1 week. After computing δ, the decisions that are made between sooner options that will pay today and later options that will pay in 1 week enables us to compute $\beta \times \delta$. For simplicity, we assume that subjects have linear utilities and their decisions for each block are narrow bracketed in the sense that their intertemporal decisions are not affected by the conditions outside the laboratory. Allowing for the degree of risk aversion that is usually observed in the literature does not affect the results of this paper.

[18]In calculating these parameters, we exclude some subjects who do not seem to understand the tasks. For instance, some people report that their δ is 800 or 0. The exclusion of such subjects results in 206 subjects (91.2%).

[19]We also use the paired t-test to assess whether the mean of β_w and β_m are significantly different and find that the differences are only marginally significant (p-value = 0.0875).

[20]Note that the monthly discount factor inferred from δ_w is lower than δ_m, i.e., $\delta_w^4 < \delta_m$. This is

that a substantial portion of subjects have present or future bias. This may be due to the differences in methods eliciting these parameters. For instance, using the multiple price list (MPL), Kim (2016) finds that 16.4% and 12.5% of subjects have present and future bias, respectively, consistent with the results from previous papers using the MPL.[21]

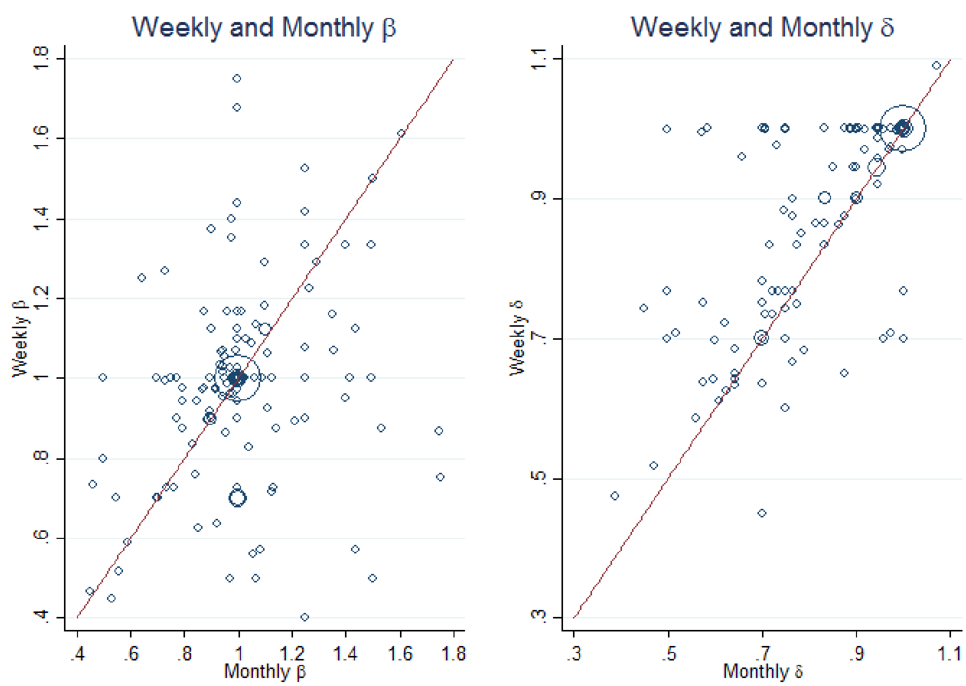


Figure 1.1: Distribution of β and δ

Figure 3.1 represents the distribution of β_w, β_m, δ_w, and δ_m. As shown in left panel, it seems that there is no clear relationship between β_w and β_m. 37.86% subjects have the equal β_w and β_m. 31.07% of subjects have β_w (β_m) higher than β_m (β_w). In contrast, the right panel shows that most subjects have δ_w higher than δ_m. Only 8.25% of subjects have δ_w lower than δ_m, confirming that subejcts discount more over longer time horizons.

consistent with ample experimental evidence that the implicit discount factor over longer time horizons is higher than the implicit discount factor over shorter time horizons. See Frederick et al. (2002).

[21] Freeman et al. (2016) compare the BDM with the MPL and find that discount factors inferred from the BDM are higher than those inferred from the MPL. As they introduce a day of the frond end delay for the earlier options, they do not investigate whether these two different mechanisms end up with different proportions of subjects with present and future bias.

1.4.2 Do subjects cooperate more when the the frequency of interactions is higher?

To answer this question we compare subjects' behavior in the Weekly treatment with that in the Monthly treatment. Figure 1.2 shows the evolution of the rate of cooperation over matches in each treatment. Note that our primary focus is on looking at the first rounds cooperation since different matches may end up with a different number of rounds, and the rate of cooperation may depend on the number of rounds.

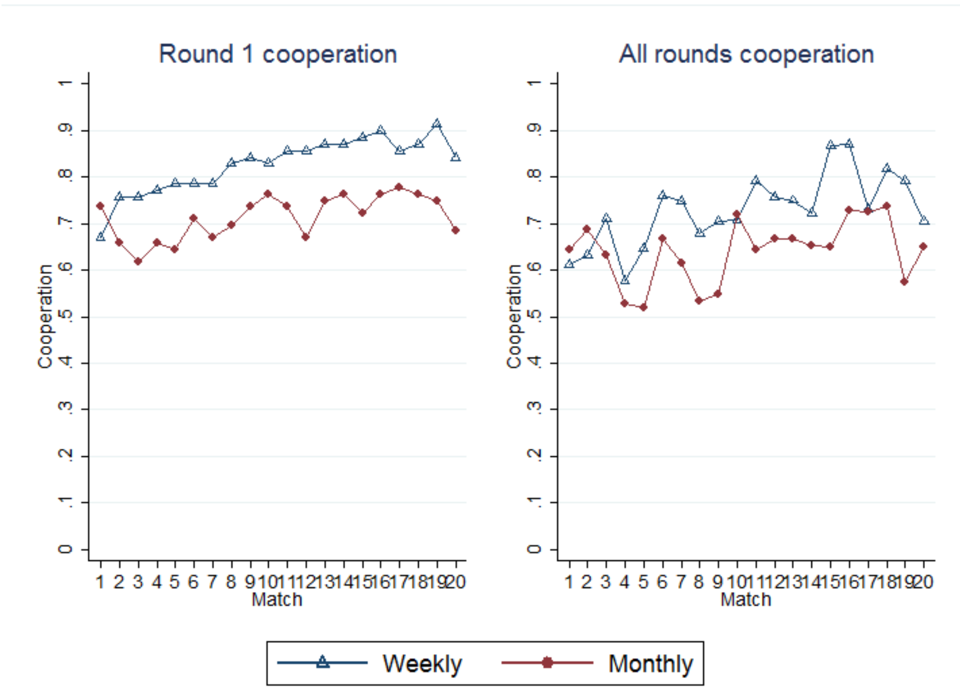


Figure 1.2: Average cooperation of the Weekly and Monthly treatments

The left panel of Figure 1.2 represents the rate of round 1 cooperation in each match. It seems that the effects of time horizons on cooperation tend to be clear as subject gain experience. In the first match, the rate of cooperation in the Monthly treatment is higher than that in the Weekly treatment. However, at the onset of the second match, the rate of cooperation in the Weekly treatment becomes higher than that in the Monthly treatment, and the difference of cooperation between the two treatments seem to get more persistent in the later matches. The right panel of Figure 1.2 represents the rate of cooperation for all rounds, and the similar pattern appears in a less clear manner.

Table 1.4: Percentage of Cooperation in Weekly and Monthly Treatments

First match					
First round			All rounds		
Weekly		Monthly	Weekly		Monthly
0.67	<	0.74	0.61	<	0.64

Last match					
First round			All rounds		
Weekly		Monthly	Weekly		Monthly
0.84	>***	0.68	0.71	>	0.63

All matches					
First round			All rounds		
Weekly		Monthly	Weekly		Monthly
0.83	>***	0.71	0.72	>	0.63

Notes:
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 1.4 supports this pattern by statistically assessing the differences between the two treatments.[22] There is no significant difference in round 1 cooperation in the first match. However, this changes as subjects gain experience: round 1 cooperation in the match is significantly higher in the Weekly treatment than in the Monthly treatment. Moreover, round 1 cooperation of all matches in the Weekly treatment is also significantly higher than that in the Monthly treatment.[23] Resonating with the right panel of Figure 1.2, looking at all rounds cooperation does not result in significant differences.

This result suggests that as predicted by t, the rate of cooperation is higher in the treatment in which the time horizon for payment is shorter, i.e., the frequency of interactions is higher. To reveal the differences in cooperative behavior across the treatments,

[22]Unless specified, statistical significance throughout the paper is assessed by probit regressions with a binary variable indicating one of the two relevant categories. Standard errors are clustered at the level of the session.

[23]Statistical differences do not hinge on the selection of specific matches. For instance, for the first and the last five matches, the rate of cooperation in the Weekly treatment is higher than the in the Monthly treatment with the p-values of 0.167 and 0.013, respectively.

however, subjects may need to gain experience for learning how to play repeated games.

1.4.3 Do subjects cooperate more when there is a delay for the first round payoff?

Next, we examine the effect of a front end delay on cooperation by comparing behavior in the Delay and the Monthly treatments. Figure 1.3 shows the rate of cooperation over matches in both treatments. The left panel of Figure 1.3 considers round 1 cooperation, and in the first match, round 1 cooperation is slightly higher in the Delay treatment in the Monthly treatment. As subjects get more experienced, they show clearly different cooperative behavior. After the decrease of cooperation for the first 4 matches, it is clear that the difference of cooperation between the two treatments becomes unambiguous as the trend of cooperation in the Delay treatment is increasing in general while the rate of cooperation in the Monthly treatment fluctuates around 70 percent. The right panel of Figure 1.3 shows the rate of cooperation for all rounds with which the general pattern is similar to the comparison of round 1 cooperation between those treatment.

The statistical significance of these differences is assessed in Table 1.5. While the difference of round 1 cooperation in the first match is not significant, round 1 cooperation in the last match is significantly higher in the Delay treatment than in the Monthly treatment. Considering round 1 cooperation in all matches also reveals significantly higher cooperation in the Delay treatment than in the Monthly treatment. Taken together, we can conclude that as subjects gain experience, they cooperate more when there is a front end delay for the round 1 payoff.

1.4.4 Do different strategies prevail across treatments?

To delve into differences in cooperative behavior across treatments to a greater extent, it is interesting to see whether subjects in different treatments use different strategies in the later matches when the subjects' behavior is more likely to be stabilized. We esti-

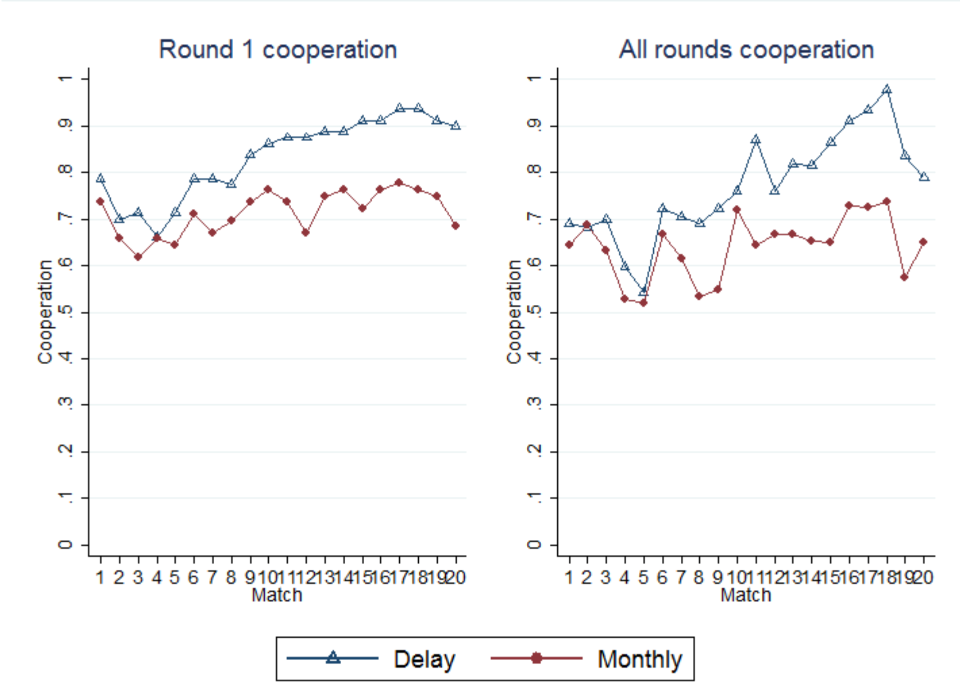


Figure 1.3: Average cooperation of the Delay and Monthly treatments

mate strategies used in the last ten matches following the Strategy Frequency Estimation Method (SFEM) in Dal Bó and Fréchette (2011).[24] In this estimation we assume 6 strategies: Always Defect (AD), Always Cooperate (AC), Grim (G), Tit for Tat(TFT), Win Stay Lose Shift(WSLS), and a trigger strategy with two periods of punishment (T2).[25]

Table 1.6 presents the estimates of the frequency of each strategy.[26] There are some interesting patterns over treatments. First, the proportion of AD is the highest in the

[24]The SFEM estimates a mixture model in which the frequency of each strategy from a pre-specified set of strategies is measured, assuming that each subjects uses the same (mixed) strategies in every repeated game with a possibility of mistakes. See Dal Bó and Fréchette (2011) for the estimation procedure in detail.

[25]WSLS is a strategy that begins with cooperation and then depends on the combination of behavior chosen in the previous round. If both cooperate or defect, then cooperation will be selected. Otherwise, WSLS will defect. T2 begins with cooperation, and if the other defects, then T2 triggers two rounds of defection. After the punishment phase, T2 goes back to cooperation. AD, AC, G, and TFT are the strategies that are most frequently identified in Dal Bó and Fréchette (2011). Dal Bó and Fréchette (2015) show that these strategies are robustly identified if T2 is replaced with another strategy such as STFT which is equivalent to TFT except that it defects in round 1.

[26]The frequency of T2 is computed by the fact that the frequencies for all strategies sum to one, and the coefficient of gamma captures the degree of noise with infinite gamma implying that behavior would be purely random. This result does not hinge on the selection of specific matches. We have very similar results if we estimate the strategies using the five last matches.

Table 1.5: Percentage of Cooperation in Delay and Monthly Treatments

First match					
First round			All rounds		
Delay		Monthly	Delay		Monthly
0.79	>	0.74	0.69	>	0.64
Last match					
First round			All rounds		
Delay		Monthly	Delay		Monthly
0.90	>***	0.68	0.79	>	0.63
All matches					
First round			All rounds		
Delay		Monthly	Delay		Monthly
0.83	>***	0.71	0.77	>**	0.63

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Monthly treatment and only less than 3 percent of the data is identified as AD in the Delay treatment. Second, a substantial proportion of subjects use cooperative strategies. In particular, on average, more than 75 percent of the data in all treatments can be identified either G or TFT. While G is more frequently identified in the treatment that pays subjects monthly, TFT is the highest in the Weekly treatment. That is, subjects in the Weekly treatment are more likely to use a strategy that punishes a defector for limited length. Subjects rarely use WSLS, and T2 seems to be used only in the Delay treatment. This result may imply that although the rate of cooperation is similar across the Weekly and the Delay treatments, the mechanisms to support cooperation in each treatment may depend on the frequency of interactions and the existence of a front end delay for the round 1 payoff. We leave more thorough investigations of this issue for future research.

Table 1.6: Estimation of Strategies Used by Treatment (Last 10 matches)

	Weekly treatment	Monthly treatment	Delay treatment
AD	0.087*	0.131*	0.027
	(0.050)	(0.074)	(0.026)
AC	0.097	0.124	0.107
	(0.065)	(0.098)	(0.092)
G	0.244***	0.268**	0.287
	(0.149)	(0.103)	(0.181)
TFT	0.572**	0.477***	0.515***
	(0.113)	(0.145)	(0.187)
WSLS	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)
T2	0.000	0.000	0.064
Gamma	0.368***	0.467***	0.329***
	(0.048)	(0.049)	(0.052)
Obs.	3,150	3,406	3,544

Notes: Bootstrapped standard errors in parentheses.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

1.4.5 Are individual time preferences related to cooperation?

In this section, we turn our attention to the individual level analysis of whether measured time preferences of subjects are related to their behavior in repeated games. For the regression below, we take the average of β_w and β_m for β and δ_w and δ_m for δ as we find similar average values for these variables.[27]

Table 1.7 presents the marginal effects from probit regressions in which the correlation of time preferences with round 1 cooperation are examined in each treatment. Surprisingly, only little evidence of such correlations is found. In the Weekly treatment, there is no correlation between time preferences and cooperation in the first match. However, as subjects gain experience, there is significant and positive correlation between δ and cooperation, implying that more patient subjects are more likely to cooperate in the last match. We find no significant correlation between δ and cooperation in the treatments

[27]In calculating δ, there are other two alternatives. One is to have monthly δ taking the average of δ_w^4 and δ_m, and the other is to take the average of δ_w and $\delta_w^{\frac{1}{4}}$ for weekly δ. As shown in the Appendix, using any value of δ does not change the results presented in Table 1.7.

Table 1.7: β, δ, and Round 1 Cooperation (Probit - Marginal effects)

	Weekly			Monthly			Delay		
	(1) Match 1	(2) Match 20	(3) All matches	(4) Match 1	(5) Match 20	(6) All matches	(7) Match 1	(8) Match 20	(9) All matches
β	-0.019	-0.055	-0.154	0.637	0.249	0.287	0.249	0.245	0.312**
	(0.418)	(0.641)	(0.243)	(0.648)	(0.542)	(0.364)	(0.265)	(0.199)	(0.145)
δ	-0.059	0.641***	0.233	-0.730	-0.291	0.200	0.139	0.174	0.157
	(0.270)	(0.193)	(0.188)	(0.643)	(0.419)	(0.410)	(0.143)	(0.310)	(0.288)
Obs.	63	63	1,260	67	67	1,340	76	76	1,520

Notes: Dependent variable: cooperation=1, defection=0. Clustered standard errors in parentheses. Marginal effects are taken at the mean.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

with monthly payment. β is only significantly related to cooperation in the Delay treatment only if we consider round 1 cooperation in all matches. This result may imply the difficulty of inferring cooperativeness from personal characteristics.[28]

One reason which makes investigating the relationship between δ and cooperation difficult may be the fact that in this experiment, cooperation and defection can then be supported as equilibrium outcomes for most subjects.[29] This may imply that what we could test in this paper is whether δ can predict who is going to coordinate on the cooperative outcome among others, rather than whether δ can draw the boundaries between unconditional defectors and potential cooperators. As shown in Dal Bó and Fréchette (2011), finding a sufficient criterion for coordination on the cooperative outcome is a

[28] There are several articles which also try to capture the connection between personal characteristics and cooperation in laboratory infinitely repeated game experiments. Sabater-Grande and Georgantzis (2002), Proto et al. (2014), and Davis et al. (2016) examine whether risk aversion affects cooperation. While Proto et al. (2014) and Davis et al. (2016) do not find a significant relationship between risk aversion and cooperation, risk aversion is negatively correlated with cooperation in Sabater-Grande and Georgantzis (2002) in which subjects are assigned into groups based on their risk aversion. Proto et al. (2014) study the relationship between intelligence and cooperation and find that, only for a high continuation probability, a group of subjects with higher IQ test scores cooperate more than a group of subjects with lower IQ test scores as they gain experience. Dreber et al. (2014) show that there is no relationship between giving behavior in a dictator game and cooperation when cooperation can be supported as an equilibrium outcome.

[29] Note that the threshold over which cooperation can be supported as an equilibrium outcome is 0.44 and most subjects have discount factors higher than this threshold.

challenging problem. Making δ a dichotomous criterion for cooperation by raising the threshold may shed light on this issue.

Another possible reason may be strategic uncertainty. If a subjects has very pessimistic beliefs about other's cooperativeness, defection could be an optimal behavior regardless of the level of δ. However, little is know about the extent to which strategic uncertainty is related to individuals' time preferences. Therefore, controlling for the effects of strategic uncertainty no cooperation may make it possible to examine the relationship between time preferences and cooperation. We leave these possibilities for future research.

1.5 Conclusions

In this paper we implement a novel experimental design for repeated games in the laboratory. Subjects make all decisions for several repeated games in the same experimental session, but stage game payoffs are paid to them over several weeks or months. Varying the frequency of interactions (weekly or monthly) allows us to investigate the effects of discounting on cooperation. Given the observation that the average elicited discount factor over a week is higher than that over a month, subjects cooperate more when the frequency of interactions is higher. This confirms that higher discount factors can lead to higher cooperation rates. We also introduce a month of a front end delay to study the effects of present bias on cooperation. We find that in the treatments with monthly payments, the rate of cooperation is higher when there is the front end delay. As subjects gain experience, the differences in behavior across the treatments become more persistent. We also relate subjects' time preferences to their cooperative behavior in repeated games. Surprisingly, we find no robust relationship between time preferences and cooperation.

Taken together, these results shed light on how we can better understand cooperative behavior in strategic interactions over time. The facts that varying the frequency of interactions affect cooperation, but there is no robust correlation between individuals' patience and cooperation may imply that we need to take into account different channels through which time preferences affect cooperation. For instance, it may be the case that changing the frequency of interactions not only affects patience, but also influences subjects' beliefs about the average level of others' patience. By the same token, introducing a delay for the first round payoff not only eliminates the effects of present bias on cooperation, but also helps subjects have beliefs about their counterparts that they are less likely present biased. Therefore, disentangling the extent to which preferences and beliefs promote cooperation in repeated games could be an interesting avenue for future research.

Another interesting question could be whether people's cooperative behavior in repeated games over time can explain their working attitude toward a team in the real work place. As many firms and other organizations have teams as the working unit of projects, understanding team members' cooperative behavior over time is an important question. As shown in this paper, measuring and relating individuals' time preferences to their behavior is not enough to infer meaningful understanding since strategic interactions involve more complicated aspects than individual decision making. Therefore, relating measured cooperativeness in the repeated game experiments in this paper to workers' administrative records may make it possible to predict who is going to shirk or exert significant efforts on team projects. We also leave this possibility for future research.

1.6 Appendix

1.6.1 Phase 1 instructions

Instructions (phase 1)

Welcome

You are about to participate in a session on decision-making, and you will be paid for your participation. What you earn depends partly on your decisions and partly on chance. The payment you earn will be paid to you through VENMO.

The entire session will take place through computer terminals. Please do not talk or try to communicate in any way with other participants during the session.

The entire session consists of two phases. The instructions for phase 1 are given below. After phase 1 ends, you will be given the instructions for phase 2.

We will start with a brief instruction period for phase 1. During this instruction period, you will be given a description of the main features of phase 1. If you have any questions during this period, raise your hand. Your question will then be answered publicly so everyone can hear.

General Instructions

1. In phase 1, you will be asked to make decisions for 8 blocks of questions. In each block, there are 2,000 questions. For each question, you can choose either: Option A, which pays you sooner, or Option B, which pays you later.

2. After you answer all questions, I will randomly pick one question and pay you the option you chose on that question. Each question is equally likely to be chosen for payment. Obviously, you have no incentive to lie on any question, because if that question gets chosen for payment, then you would end up with the option you like less.

3. For example, the questions in one block are as follows (note that each row corresponds to a question, and so you will have to choose an option in each row):

Questions	Payment Option A (Pays the Amount Below Today)	Payment Option B (Pays the Amount Below in 1 month)
1	\$8.00	\$0.01
2	\$8.00	\$0.02
3	\$8.00	\$0.03
⋮	⋮	⋮
1,999	\$8.00	\$19.99
2,000	\$8.00	\$20.00

I assume you will choose Option A for at least the first few questions, but at some point switch to choosing Option B. In order to save time, you can answer at which dollar value you'd switch. I can then 'fill out' your answers to all 2,000 questions based on your switch point (choosing Option A for all questions before your switch point, and Option B for all questions at or after your switch point). I will still draw one question randomly for payment. Again, if you lie about your preferred switch point, you might end up with an option that you like less.

4. The 8 blocks will differ in two ways: (1) the timings of sooner and later payments:

- Between payment **today** and payment in **1 week**.
- Between payment **today** and payment in **1 month**.
- Between payment in **1 month** and payment in **1 month and 1 week**.
- Between payment in **1 month** and payment in **2 months**.

and (2) whether you are asked to switch from A to B or B to A.

Payment

1. At the end of the experiment, one question in one of the blocks will be randomly selected for payment and will be displayed on your screen. Depending on your decision for that question, you will be paid on the designated date through VENMO. If in the question that is randomly selected, your decision was to receive a payment today, then you will be paid through VENMO within a few hours of the end of the experiment. If, on the other hand, your decision was to receive a payment in the future, you will be paid on the designated date through VENMO.

2. In addition, you will receive a $5 show-up fee through VENMO after the experiment.

- Are there any questions?

Before we start, let me remind you that:

- There are 8 blocks of questions in each of which you will be asked to state your switch point.
- Only one question in one of the blocks will be randomly selected for payment.
- Depending on your decision, you will be paid on the designated date through VENMO.
- A $5 show-up fee will be paid to you through VENMO after the experiment.

1.6.2 Phase 2 instructions

Instructions (phase 2) - Weekly treatment

We will start with a brief instruction period for phase 2. During this instruction period you will be given a description of the main features of phase 2. If you have any questions during this period, raise your hand. Your question will then be answered publicly so that everyone can hear.

General Instructions

1. In phase 2 you will be asked to make decisions in several rounds. Each sequence of rounds is referred to as a match. You will be randomly paired with another person for a match.

2. The length of a match is determined randomly. After each round, there is a 75% probability that the match will continue for at least another round. This is as if we were to randomly choose an integer between 1 and 100 and continue if the number chosen is less than or equal to 75 and end if the number chosen is larger than 75. So, for instance, if you are in round 2, the probability that there will be a third round is 75%, and if you are in round 9, the probability that there will be another round is also 75%.

3. Once a match ends, you will be randomly paired with another person for a new match. You will have 20 matches in phase 2.

4. In each round, you will be asked to choose between action 1 and 2. The payoffs are determined by your action and the action chosen by the person paired with you. The payoffs are described in the table below:

	The other's choice	
Your choice	1	2
1	\$4.00, \$4.00	\$1.00, \$5.00
2	\$5.00, \$1.00	\$2.00, \$2.00

- The first entry in each cell represents your payoff, while the second entry represents the payoff of the person you are paired with. That is, if:

You select 1 and the other selects 1, you each make \$4.00.

You select 1 and the other selects 2, you make \$1.00 while the other makes \$5.00.

You select 2 and the other selects 1, you make \$5.00 while the other makes \$1.00.

You select 2 and the other selects 2, you each make \$2.00.

- Once you and the person you are paired with have made your choices, those choices will be highlighted and your payoff for the round will appear.

Payment

1. At the end of the experiment, one of the matches will be randomly selected for payment.

2. For the selected match, you will receive payment for the first round today. After that, you will receive payment for the following rounds once a week. That is, you will receive payment for the second round in 1 week, payment for the third round in 2 weeks, and so on. The schedule of payment is summarized in the table below.

Payoffs (round)	Payment schedule (from today)
1st round payoff	Today
2nd round payoff	in 1 week
3rd round payoff	in 2 weeks
⋮	⋮

3. In the same way that payments are made for phase 1, you will be paid on the designated dates through VENMO.

- Are there any questions?

Before we start, let me remind you that:

- The length of a match is randomly determined. After each round, there is a 75% probability that the match will continue for at least another round. You will play with the same person for the entire match.

- After a match is finished, you will be randomly paired with another person for a new match. You will have 20 such matches.

- One match will be randomly selected for payment.
You will receive your payment for the first round today. After that, you will receive payment for the following rounds once a week. That is, you will receive payment for the second round in 1 week, payment for the third round in 2 weeks, and so on.

Instructions (phase 2) - Monthly treatment

We will start with a brief instruction period for phase 2. During this instruction period you will be given a description of the main features of phase 2. If you have any questions during this period, raise your hand. Your question will then be answered publicly so that everyone can hear.

General Instructions

1. In phase 2 you will be asked to make decisions in several rounds. Each sequence of rounds is referred to as a match. You will be randomly paired with another person for a match.

2. The length of a match is determined randomly. After each round, there is a 75% probability that the match will continue for at least another round. This is as if we were to randomly choose an integer between 1 and 100 and continue if the number chosen is less than or equal to 75 and end if the number chosen is larger than 75. So, for instance, if you are in round 2, the probability that there will be a third round is 75%, and if you are in round 9, the probability that there will be another round is also 75%.

3. Once a match ends, you will be randomly paired with another person for a new match. You will have 20 matches in phase 2.

4. In each round, you will be asked to choose between action 1 and 2. The payoffs are determined by your action and the action chosen by the person paired with you. The payoffs are described in the table below:

	The other's choice	
Your choice	1	2
1	\$4.00, \$4.00	\$1.00, \$5.00
2	\$5.00, \$1.00	\$2.00, \$2.00

- The first entry in each cell represents your payoff, while the second entry represents the payoff of the person you are paired with. That is, if:

You select 1 and the other selects 1, you each make \$4.00.

You select 1 and the other selects 2, you make \$1.00 while the other makes \$5.00.

You select 2 and the other selects 1, you make \$5.00 while the other makes \$1.00.

You select 2 and the other selects 2, you each make \$2.00.

- Once you and the person you are paired with have made your choices, those choices will be highlighted and your payoff for the round will appear.

Payment

1. At the end of the experiment, one of the matches will be randomly selected for payment.

2. For the selected match, you will receive payment for the first round today. After that, you will receive payment for the following rounds once a month. That is, you will receive payment for the second round in 1 month, payment for the third round in 2 months, and so on. The schedule of payment is summarized in the table below.

Payoffs (round)	Payment schedule (from today)
1st round payoff	Today
2nd round payoff	in 1 month
3rd round payoff	in 2 months
⋮	⋮

3. In the same way that payments are made for phase 1, you will be paid on the designated dates through VENMO.

- Are there any questions?

Before we start, let me remind you that:

- The length of a match is randomly determined. After each round, there is a 75% probability that the match will continue for at least another round. You will play with the same person for the entire match.

- After a match is finished, you will be randomly paired with another person for a new match. You will have 20 such matches.

- One match will be randomly selected for payment.
You will receive your payment for the first round today. After that, you will receive payment for the following rounds once a month. That is, you will receive payment for the second round in 1 month, payment for the third round in 2 months, and so on.

Instructions (phase 2) - Delay treatment

We will start with a brief instruction period for phase 2. During this instruction period you will be given a description of the main features of phase 2. If you have any questions during this period, raise your hand. Your question will then be answered publicly so that everyone can hear.

General Instructions

1. In phase 2 you will be asked to make decisions in several rounds. Each sequence of rounds is referred to as a match. You will be randomly paired with another person for a match.

2. The length of a match is determined randomly. After each round, there is a 75% probability that the match will continue for at least another round. This is as if we were to randomly choose an integer between 1 and 100 and continue if the number chosen is less than or equal to 75 and end if the number chosen is larger than 75. So, for instance, if you are in round 2, the probability that there will be a third round is 75%, and if you are in round 9, the probability that there will be another round is also 75%.

3. Once a match ends, you will be randomly paired with another person for a new match. You will have 20 matches in phase 2.

4. In each round, you will be asked to choose between action 1 and 2. The payoffs are determined by your action and the action chosen by the person paired with you. The payoffs are described in the table below:

	The other's choice	
Your choice	1	2
1	\$4.00, \$4.00	\$1.00, \$5.00
2	\$5.00, \$1.00	\$2.00, \$2.00

- The first entry in each cell represents your payoff, while the second entry represents the payoff of the person you are paired with. That is, if:

You select 1 and the other selects 1, you each make \$4.00.

You select 1 and the other selects 2, you make \$1.00 while the other makes \$5.00.

You select 2 and the other selects 1, you make \$5.00 while the other makes \$1.00.

You select 2 and the other selects 2, you each make \$2.00.

- Once you and the person you are paired with have made your choices, those choices will be highlighted and your payoff for the round will appear.

Payment

1. At the end of the experiment, one of the matches will be randomly selected for payment.

2. For the selected match, you will receive payment for the first round in 1 month from today. After that, you will receive payment for the following rounds once a month. That is, you will receive payment for the second round in 2 months, payment for the third round in 3 months, and so on. The schedule of payment is summarized in the table below.

Payoffs (round)	Payment schedule (from today)
1st round payoff	in 1 month
2nd round payoff	in 2 months
3rd round payoff	in 3 months
⋮	⋮

3. In the same way that payments are made for phase 1, you will be paid on the designated dates through VENMO.

- Are there any questions?

Before we start, let me remind you that:

- The length of a match is randomly determined. After each round, there is a 75% probability that the match will continue for at least another round. You will play with the same person for the entire match.

- After a match is finished, you will be randomly paired with another person for a new match. You will have 20 such matches.

- One match will be randomly selected for payment.
You will receive your payment for the first round in 1 month from today. After that, you will receive payment for the following rounds once a month. That is, you will receive payment for the second round in 2 months, payment for the third round in 3 months, and so on.

1.6.3 Phase 1 screen shots

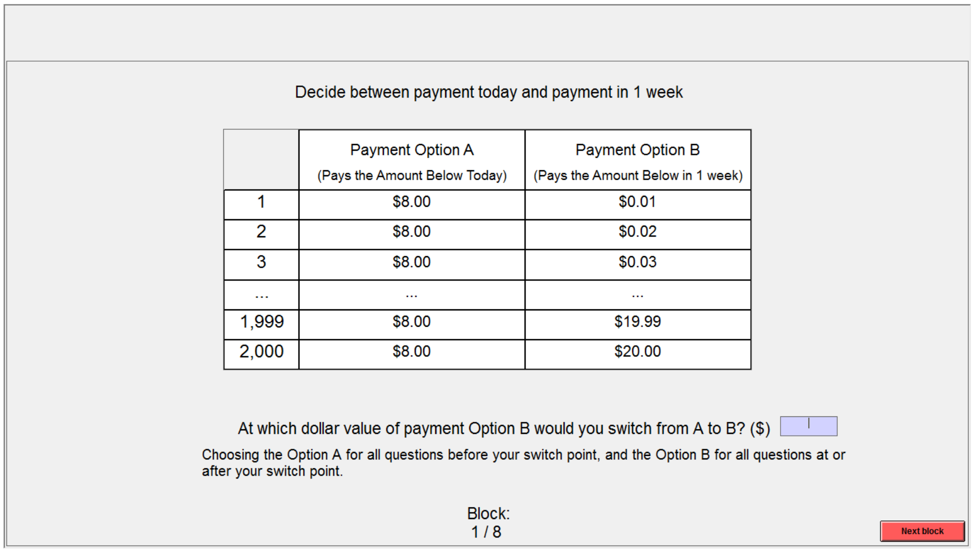


Figure 1.4: The screen shot of phase 1 (block 1)

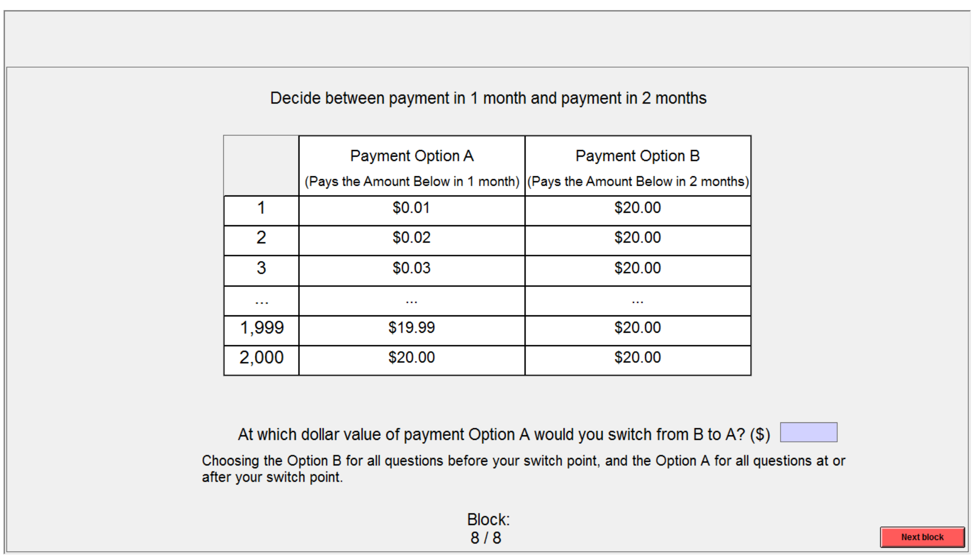


Figure 1.5: The screen shot of phase 1 (block 8)

1.6.4 Phase 2 screen shots (Delay treatment)

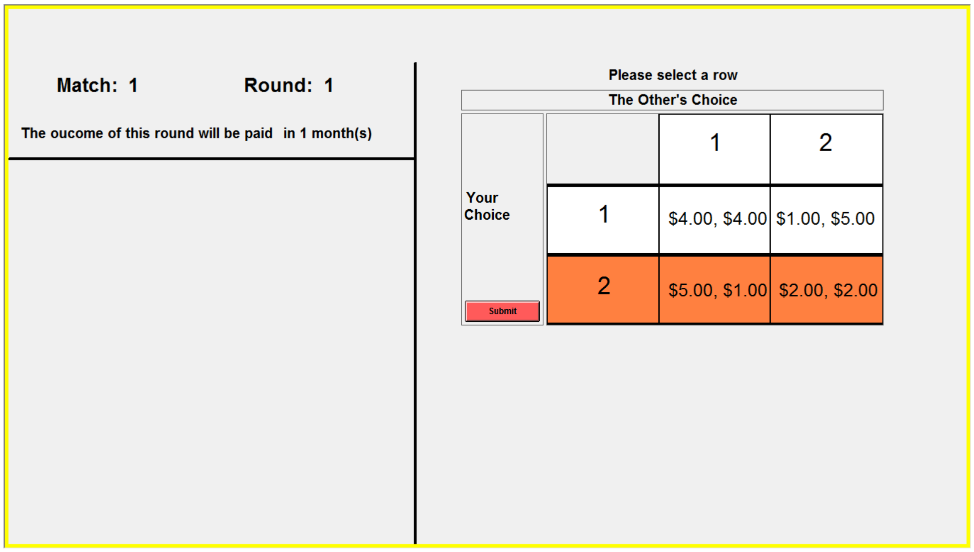


Figure 1.6: The screen shot of the round 1 decision stage

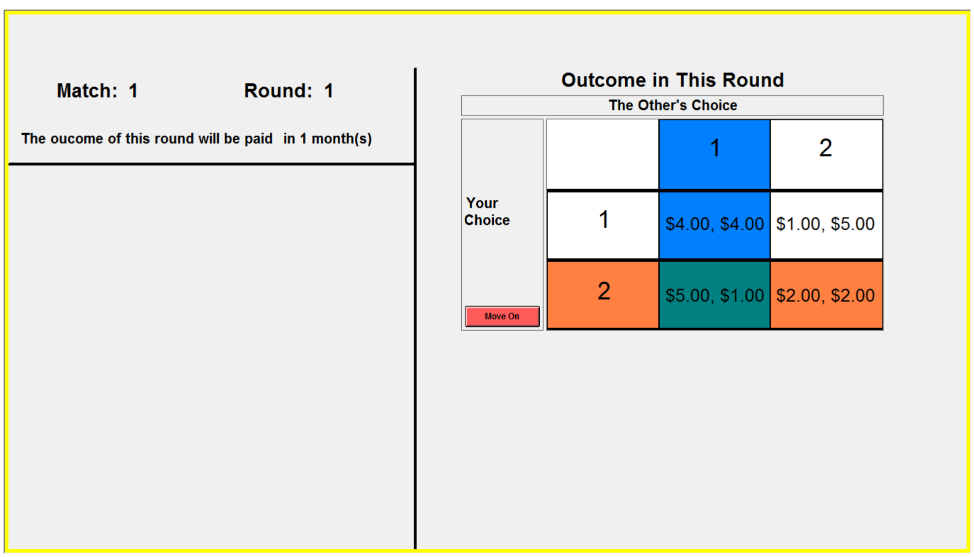


Figure 1.7: The screen shot of the round 1 feedback stage

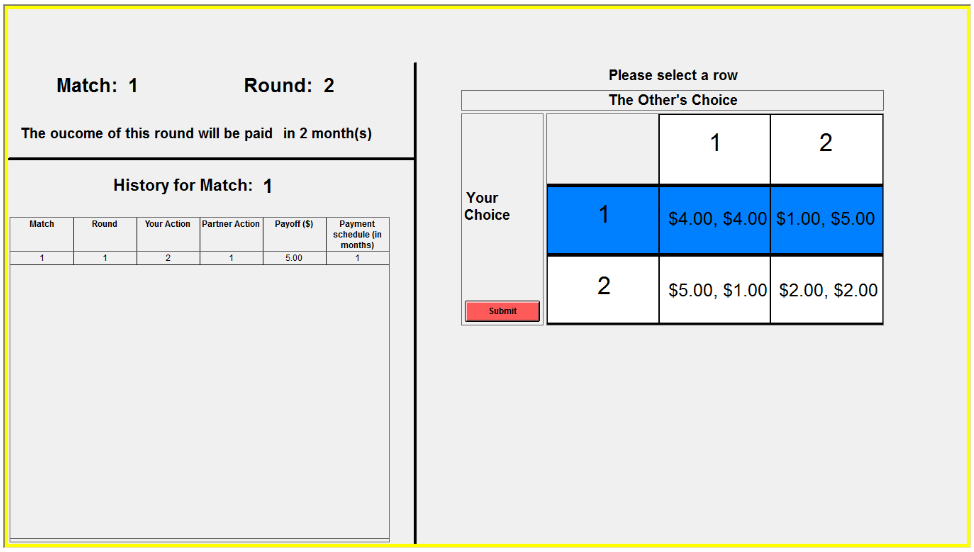


Figure 1.8: The screen shot of the round 2 decision stage

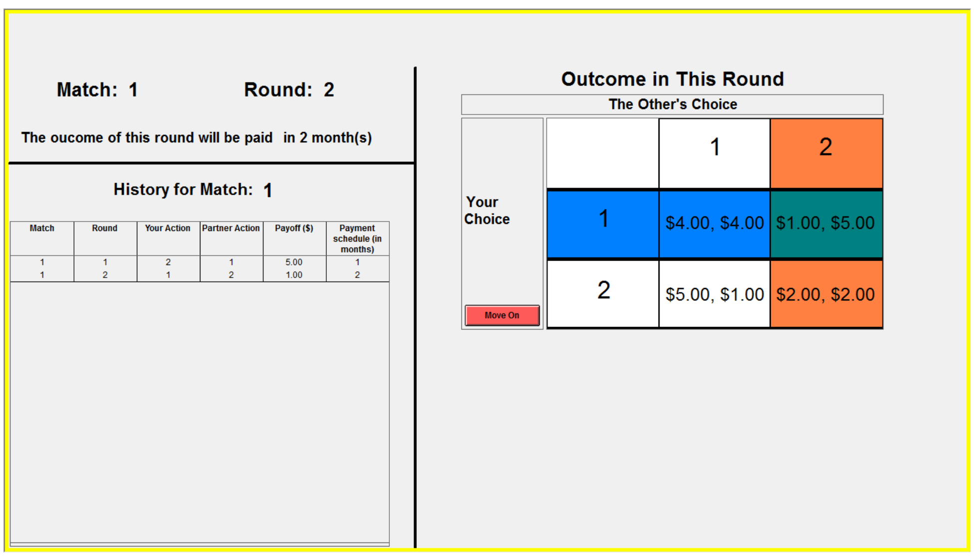


Figure 1.9: The screen shot of the round 2 feedback stage

1.6.5 Robustness checks for Table 1.7

(1) Monthly discount factor: $\delta = (\delta_w^4 + \delta_m)/2$

Table 1.8: β, δ, and Round 1 Cooperation (Probit - Marginal effects)

	Weekly			Monthly			Delay		
	(1) Match 1	(2) Match 20	(3) All matches	(4) Match 1	(5) Match 20	(6) All matches	(7) Match 1	(8) Match 20	(9) All matches
β	0.007	-0.035	-0.125	0.601	0.229	0.289	0.232	0.285	0.299**
	(0.422)	(0.190)	(0.242)	(0.635)	(0.538)	(0.264)	(0.263)	(0.223)	(0.149)
δ	0.010	0.399***	0.184**	-0.480	-0.197	0.114	0.066	0.102	0.082
	(0.149)	(0.114)	(0.094)	(0.348)	(0.234)	(0.227)	(0.077)	(0.153)	(0.167)
Obs.	63	63	1,260	67	67	1,340	76	76	1,520

Notes: Dependent variable: cooperation=1, defection=0. Clustered standard errors in parentheses. Marginal effects are taken at the mean.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

(2) Weekly discount factor: $\delta = (\delta_w + \delta_m^{\frac{1}{4}})/2$

Table 1.9: β, δ, and Round 1 Cooperation (Probit - Marginal effects)

	Weekly			Monthly			Delay		
	(1) Match 1	(2) Match 20	(3) All matches	(4) Match 1	(5) Match 20	(6) All matches	(7) Match 1	(8) Match 20	(9) All matches
β	-0.004	-0.073	-0.135	0.6017	0.205	0.270	0.265	0.300	0.314**
	(0.446)	(0.201)	(0.257)	(0.629)	(0.557)	(0.372)	(0.277)	(0.224)	(0.144)
δ	-0.024	0.969***	0.466	-1.292	-0.673	0.228	0.312	0.360	0.294
	(0.424)	(0.316)	(0.320)	(0.936)	(0.693)	(0.674)	(0.240)	(0.364)	(0.444)
Obs.	63	63	1,260	67	67	1,340	76	76	1,520

Notes: Dependent variable: cooperation=1, defection=0. Clustered standard errors in parentheses. Marginal effects are taken at the mean.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

CHAPTER 2

DISCOUNTING, DYNAMIC CONSISTENCY, AND COOPERATION IN AN INFINITELY REPEATED GAME EXPERIMENT

2.1 Introduction

Understanding cooperation in human interactions is key in economics and other social sciences. Contributions in the theory of infinitely repeated games showed that allowing for repeated interactions can make punishment for opportunistic behavior and rewards for cooperative behavior credible, and cooperation can then be supported as an equilibrium outcome. This implies that deciding to cooperate may embed intertemporal comparisons between sooner benefits from defection and later, but larger overall benefits from cooperation. In other words, time preferences (e.g. discount factor) are essential in determining cooperation in repeated games. However, whether measured time preferences of human subjects are indeed associated with cooperative behavior in a repeated game has yet to be studied.

In this paper the relationship between time preferences and cooperation in infinitely repeated prisoner's dilemma experiments is explored.[1] We elicit time preferences of human subjects by using an incentivized mechanism.[2] And measured time preferences are related to cooperative behavior in a repeated prisoner's dilemma game.[3]

[1]Throughout the paper 'time preference' refers to individual subject's β, δ, time consistency, and time invariance. Definitions and identifications of these parameters in detail will be given in the next section.

[2]A growing body of literature has studied how we can measure individuals' time preferences and whether measured time preferences are correlated with behavior over time. See Frederick et al. (2002) for a critical review of early attempts to measure time preferences, and Urminsky and Zauberman (2016) for a recent survey. Many studies show that measured time preferences are indeed related to intertemporal decision making. For example, Meier and Sprenger (2010) find that individuals with present bias are more likely to have credit card debt.

[3]Infinitely repeated games are induced in the laboratory by using the probability of continuation which determines a length of a repeated game (see Roth and Murnighan, 1978). After each stage game, there would be one more stage game with a fixed probability. Otherwise, that repeated game would be terminated. This fixed probability is known to all subjects. See Dal Bó and Fréchette (forthcoming) for

To the best of our knowledge, Davis et al. (2016) is the sole paper to investigate the relationship between time preferences and behavior in repeated prisoner's dilemma experiments in the laboratory.[4] Not surprisingly, almost no evidence of such correlation was found since there exist difficulties for examining the relationship between time preferences and cooperation in the conventional laboratory setting of repeated game experiments.[5]

First, there is no time horizon over which payoffs from each stage game can be discounted. Second and more importantly, payoffs are not separable across stage games. In other words, for payoffs to be discounted over time, subjects should receive their payoffs at the end of each stage game. Infinitely repeated game experiments in the laboratory violated these conditions in that subjects were paid all of their payoffs at the end of the experiment.[6]

To overcome these difficulties, we introduce a novel experimental design in which a repeated game is played over time - subjects play one stage game and receive associated payoffs each week. Along with the elicitation of subjects' time preferences, we implement this experimental design on Amazon's Mechanical Turk (MTurk). Subjects are recruited and asked to repeatedly participate in the experiment once a week until the session is terminated.[7]

---

a survey of this literature.

[4]Several papers investigate the relationship between other personal characteristics and cooperation in repeated games. Sabater-Grande and Georgantzis (2002), Proto et al. (2014), and Davis et al. (2016) examine whether risk aversion affects cooperation, and no robust relationship is found. Proto et al. (2014) study the relationship between intelligence and cooperation. They find that, only for a high continuation probability, a group of subjects with higher IQ test scores cooperate more than a group of subjects with lower IQ test scores as they gain experience. For social preferences, Dreber et al. (2014) show that there is no correlation between giving behavior in a dictator game and cooperation when cooperation can be supported as an equilibirum outcome.

[5]Resonating with the theoretical predictions, Dal Bó (2005) shows that different continuation probabilities affect cooperation in infinitely repeated prisoner's dilemma experiments. However, this does not necessarily imply that time preferences of human subjects would be correlated with cooperative behavior in repeated games. In this paper we fix the continuation probability as 0.75.

[6]Even if subjects are paid at the end of every stage game in the conventional laboratory setting, some problems will still remain that time horizons over which stage game payoffs can be discounted would be too short (probably for some minutes), and there would be no opportunity for payoffs from one stage game to be consumed before subjects receive payoffs from the next stage game.

[7]As will be clearly demonstrated, what we mean by a "session" in this paper refers to a cohort of a longitudinal experiment which includes the elicitation of time preferences (week 0) and one repeated game (week 1 and after, if applicable). 5 sessions started on 4 different dates, and each subject is allowed

We present experimental evidence of the effects of measured time preferences on various facets of cooperation. First, consistent with a model of quasi-hyperbolic discounting (e.g., Laibson, 1997) and its application to repeated games (Chade et al., 2008), the effects of two parameters of time preferences - β, the extent to which present and future biases are measured, and δ, a discount factor - are examined. We find that β is positively and significantly correlated with cooperation in week 1 and in all weeks. Surprisingly, we find no relationship between δ and cooperation.

Second, we look at the effects of time consistency and time invariance á la Halevy (2015) on cooperation. Time consistency requires that the preferred choice does not depend on the time at which decisions are made. To relate time consistency to behavior in a repeated game, we adopt the novel design of Dal Bó and Fréchette (2015) eliciting subjects' plan of action in which when subjects choose their action in week 1, they are also asked to specify their actions in week 2 for all possible contingencies. We find that subjects with time consistent preferences are significantly less likely to deviate from their plan of action in week 2 than subjects with time inconsistent preferences.

Time invariance means that subjects' preferences at different times should be identical. This may imply that subjects who exhibit mutual cooperation in week 1 are expected to cooperate in week 2 as they have identical preferences over these two weeks. We find that after mutual cooperation in week 1, subjects with time invariant preferences are more likely to cooperate than subjects with time variant preferences.

Finally, we find that subjects with higher β and δ are less likely to drop out of the experiment. They also participate in the experiment longer than subjects with lower β and δ. As such, more patient subjects are more able to maintain cooperative relationships in our experiment.

---

to participate in only one session. See Section 2 for the experimental design in detail.

2.2 Experimental Design

2.2.1 Overall Design

The experiments were conducted on Amazon's Mechanical Turk (MTurk) which is an online labor market platform provided by Amazon, where an increasing number of experiments in economics have been conducted on MTurk.[8] The novelty of this experiment is that it was run with the same subjects over several weeks. Once a week, the same subjects were invited to the MTurk page, where they participated in the task (either intertemporal choice, stage game or both) and were paid for decisions made in that week.[9] More specifically, in week 0, subjects were recruited from MTurk and participated in the incentivized task which measured their time preferences. In week 1 and after, the same subjects were invited to the MTurk page by the internal message system in MTurk, and participated in one stage game each week. The timeline of a session is presented in Figure 2.1.

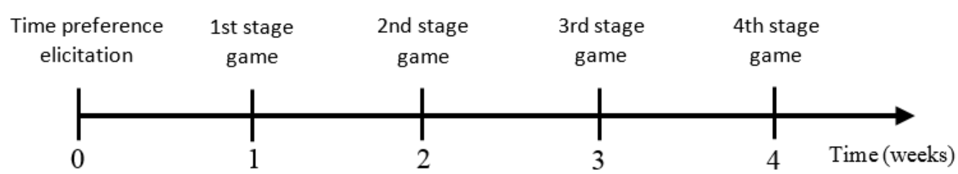


Figure 2.1: Timeline of a session

When the subjects were recruited in week 0, they were told that this experiment will last at least two weeks or longer, but they were not informed about what they will be asked to do in week 1 and after. They learned about the rules of an infinitely repeated game, including the payoff matrix and the probability of continuation in week 1. To keep consistency of the timing at which subjects were recruited (or invited) and paid, throughout the

[8]See Horton et al. (2011) for a comparison between laboratory experiments and online experiments on MTurk.

[9]Upon accepting the task on the MTurk page, subjects were given a link to a survey website provided by Qualtrics. After finishing their task, subjects were asked to return to the MTurk page and to enter a code which was provided at the end of the survey. This is one typical way of having surveys or experiments on MTurk, and all the payments were made on MTurk by transferring money from the experimenter's account to the subjects' accounts.

experiment, we recruited and sent the invitation messages to subjects only on Wednesday and made clear announcements that subjects who completed the task by 5:00 pm E.T. on Friday will be eligible for being paid before midnight E.T. on Friday in the same week.[10]

2.2.2 Measuring Time Preferences

Time Preferences Elicitation

We elicit subjects' time preference by using the experimental design proposed by Halevy (2015). This experimental design consists of two different time points at which subjects were asked to make intertemporal choices. In week 0, after being recruited, each subject was asked to make two blocks of decisions, and in week 3 one block of decisions was given to each subject. For each decision block, the multiple price list (MPL) method was used in which there are ten lists of two options: a sooner payment and a later payment, which was delayed for a week. Sooner payments always paid subjects \$0.50, and later payments ranged from \$0.50 to \$0.68 (with \$0.02 increments) and were placed from top to bottom with an increasing order.[11]

In block 1 of decisions in week 0, each subject was asked to choose between sooner payments in week 0 and later payments in week 1. Block 2 of decisions in week 0 required subjects to choose between sooner payments in week 3 and later payments in week 4. For block 3 of decisions in week 3, subjects had to decide between sooner payments in week 3 and later payments in week 4. Denote by x_1, x_2, and x_3 the switching point from a sooner payment to a later payment in block 1, block 2, and block 3, respectively. Figure 2.2 presents the timeline of the time preference elicitation, and t refers to a week on which

[10]In MTurk, workers usually do not exactly know when they will be paid by the requester to whom they submitted their work. Rather, requesters are asked to set a deadline by which they have to pay their workers and this deadline is known to workers before they decide to accept the task or not. The maximum possible duration of this deadline is 30 days. In this experiment, each week's experiment started on Wednesday and subjects were clearly informed that they will be paid on Friday in the same week to avoid any kind of uncertainty from the timing of payments being made.

[11]An example screen capture of a block of time preference tasks is presented in the Appendix.

corresponding intertemporal choices were made.

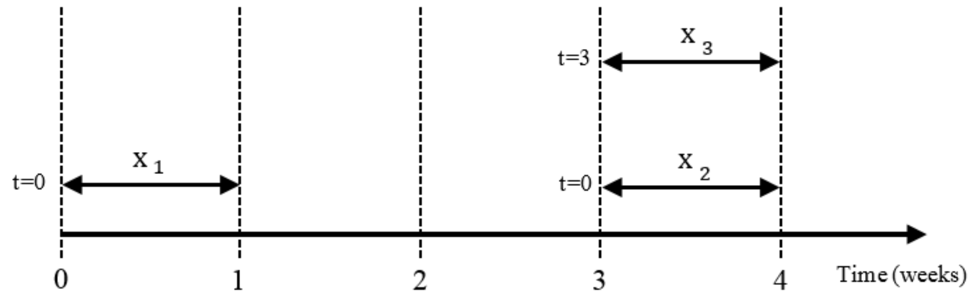


Figure 2.2: Timeline of the time preference elicitation

To pay subjects in an incentive compatible way, we use the design of the robustness treatment in Halevy (2015). When subjects made their decisions for block 1 and block 2 in week 0, they were told that only one decision for each block will be randomly chosen for their actual payment. In week 3, the decisions for block 3 are such that subjects were given an opportunity to revise their decisions for block 2 and one of these revised decisions was randomly selected for payment.[12] After all the relevant decisions had been made, subjects were notified of decisions that were randomly selected for their payments.

Identification of Time Preferences

First, the parameters of time preferences of interest are β and δ.[13] Subjects' decisions for each block may reveal their intertemporal marginal rate of substitution between two

[12]In the main treatment of Halevy (2015), when making decisions in week 0, subjects were informed that they will be asked to make decisions for the block 3 in week 3, and by tossing a coin either the block 2 or the block 3 will be implemented for their actual payment with equal probability. We decide to use the robustness design to avoid making the decision problems too complicated for our subjects to understand. This also aims not to mislead subjects' perceptions about the length of the experiment. When recruited, subjects did not know about how long the experiment will exactly continue, and telling them that they will make decisions in week 3 for sure may mislead subjects to believe that the experiment will not finish before week 3 and consequently, their behavior in a repeated game may also be influenced by this misperception. See Halevy (2015) for the discussion about incentive compatibilities and drawbacks for these two payment designs.

[13]In a model of quasi-hyperbolic discounting, a person with $\beta - \delta$ preferences evaluates a stream of payoffs with the sequence of quasi-hyperbolic discount, 1, $\beta\delta$, $\beta\delta^2$,.... If β equals 1, this model is equivalent to the standard model of exponential discounting. If $\beta <$ (>)1, it can capture the notion of present (future) bias. See Laibson (1997).

different time points. Inferring the accurate estimates of β and δ from these decisions needs us to have more structural assumptions and information on subjects' utility function and liquidity constraints.[14] However, it is important to note that our primary purpose in this paper is not to have precise point estimates of these parameters for each subject, but to investigate how heterogeneity of time preferences in our sample affects their behavior in an infinitely repeated game. Therefore, we simply define β and δ from the decisions for block 1 and 2.[15] We also assume that an individual is indifferent between two payments at the last point at which an individual prefers a sooner payment to a later payment.[16] We can then estimate δ from the decisions in block 2 and $\beta \times \delta$ from the decisions in block 1.[17] If β equals 1, then preferences are stationary.

Second, comparing switching points across blocks enables us to identify time consistency and time invariance. Time consistency requires that the preferred choice does not depend on the time at which decisions are made. In other words, once a decision maker makes a decision over temporal payments at t, he does not have an incentive to deviate from his decision at t' from his ex ante decision at t. This leads us to have the following identification: If $x_3 = x_2$, then preferences are time consistent.

Time invariance means that subjects' preferences at different times should be identical. If preferences are invariant, the decision maker's evaluation does not account for a specific date when the decision is made, but only account for the time delay between the time at which the decision is made and the time at which payments are given. Then, time invariance can be identified as follows: If $x_3 = x_1$, then preferences are time invariant.[18]

[14]See Andersen et al. (2008) and Andreoni and Sprenger (2012) for recent developments in measuring time preferences. See also Dean and Sautmann (2014) for the discussion about accounting for financial shocks on liquidity constraints.

[15]More specifically, we assume that subjects have linear utilities and their decisions for each block are narrow bracketed in the sense that their intertemporal decisions are not affected by the conditions outside the laboratory.

[16]Among 1,345 subjects, only 2.9% of subjects have multiple switching points for at least one block. For these subjects we take the first switching point as their true switching point. Exclusion of these subjects does not affect the results in this paper.

[17]With a sooner payment, x and a later payment, y between which an individual is indifferent, we can have $x = \delta y$ from decisions in block 1 and $x = \beta\delta y$ from decisions in block 2.

[18]Halevy (2015) proves that any two of the three properties: Stationarity, Time Invariance, and Time

2.2.3 Infinitely Repeated Prisoner's Dilemma

In week 1 and after, subjects participated in one infinitely repeated prisoner's dilemma game. They played one stage game and received associated payoffs each week. The probability of continuation is fixed as 0.75 and the payoff table shown to each subject is presented in Table 2.1. The same payoff structure was used in Dal Bó and Fréchette (2011) allowing us to have 56.82% of the average rate of cooperation in the first repeated game as a benchmark. Throughout the experiment, cooperation and defection are represented as action 1 and action 2 , respectively.

Table 2.1: Stage game payoffs

	The other's choice	
Your choice	1	2
1	\$0.48, \$0.48	\$0.12, \$0.50
2	\$0.50, \$0.12	\$0.25, \$0.25

To control for subjects' beliefs about attrition of other participants, subjects were told that they would be re-matched with another person if the partner did not return to the experiment in a future week. The subjects were also told that the other person's attrition would not affect their eligibility of participation until the session is terminated. Otherwise, subjects played with the same counterpart in all weeks.[19] In week 2 and after, when choosing an action, subjects were reminded of the history of actions and informed about whether they were re-matched with another person.

We conducted 4 sessions on different dates. In addition, we have one more session in which we elicit subjects' plans of action. This session adopts the novel strategy elicitation

Consistency, imply the third and shows that a substantial portion of subjects with time inconsistent preferences has stationary time preferences.

[19]The exact wording is: "In the event, which we hope is unlikely, that your counterpart fails to continue the interaction in a future week, we will arrange for you to be able to continue playing with another counterpart or will make other arrangements to minimize any impact on your predicted earnings. Unless we inform you otherwise, you will definitely be playing with the same counterpart at each future stage."

method used in Dal Bó and Fréchette (2015). In week 1, subjects are asked to specify a plan of action by answering one question for their action in week 1 and four questions regarding their action in week 2, which include all possible combinations of actions chosen by the subject and the other person in the first week.[20] Without being reminded of or having the option to review their plan of action, subjects are asked to choose their action in the second stage game, regardless of their specified plan of action in the first stage game. This allows us to study whether time inconsistent subjects are more likely to deviate from their plan.

2.2.4 Mechanisms to Prevent Attrition

It becomes important to prevent attrition of subjects as the same subjects are asked to repeatedly participate in the experiment for several weeks. Two mechanisms were used to achieve that goal. First, when being recruited in week 0, subjects were promised to be paid the completion bonus of $3 at the end of the session if they took1 part in all weeks. This completion bonus of $3 was given on top of the earnings from the experiment. Second, subjects were told that if they did not complete their task for a week, they would not be invited to the future tasks, nor would they be eligible for the completion bonus.[21] These two mechanisms were clearly announced in week 0 when we recruited subjects.

[20]The exact wording is: "In addition to your choice above, you are asked to specify a plan of action. A plan of action is specified by answering 4 questions: After this week, if the experiment continues for one more week and
(1) I last selected 1 and the other selected 1, then I will choose...
(2) I last selected 1 and the other selected 2, then I will choose...
(3) I last selected 2 and the other selected 1, then I will choose...
(4) I last selected 2 and the other selected 2, then I will choose..."

[21]Even if a MTurk worker who was not invited by our MTurk message visited our MTurk page, he was not able to participate in our experiment as we prevented such workers from doing our task.

2.3 Research Questions

2.3.1 Discounting and Cooperation

Theoretical contributions in infinitely repeated games have shown that discount factors are essential in determining cooperation supported as an equilibrium outcome. Fudenberg and Maskin (1986) show that if players have sufficiently a high discount factor, δ, individually rational payoffs can be supported in a subgame perfect equilibrium. Abreu et al. (1990) show that the set of subgame perfect equilibrium payoffs expands in δ.[22] Lehrer and Pauzner (1999) study a model of repeated games with differential discount factors under complete information. They show that the set of feasible payoffs can be larger than the convex hull of the stage game payoffs as players can be better off by trading payoffs over time.[23]

Chade et al. (2008) study infinitely repeated games under quasi-hyperbolic discounting. They show that an intuitive idea that the set of equilibrium payoffs increase in δ and β may not hold in general. However, they prove that for a class of games including the prisoner's dilemma in which the minimax point of the stage game coincides with a Nash equilibrium, with δ (β) fixed, the set of equilibrium outcomes expands as β (δ) increases.[24] Taken together and even when the environment we study may differ from the one studied by theory, these justify us to have the following question.

Question 5. *Do higher β and higher δ promote higher cooperation?*

[22] Under the assumptions of complete information about payoffs and a common discount factor, the threshold of a weekly discount factor, δ, over which cooperation can be supported as an equilibrium outcome is 0.17 in this experiment.

[23] The environment of this experiment embeds incomplete information about differential discount factors as subjects don't know the discount factor of the other person paired with. The characterization of the set of equilibrium payoffs under such incomplete information is still in question.

[24] Chade et al. (2008) assume $\beta \leq 1$. Whether the same results can hold when there is no restriction on β needs to be studied.

2.3.2 Time Consistency and Cooperation

Time consistency states that a decision maker's evaluation of intertemporal choices should not depend on the time at which he makes such decisions. When a subject begins to play an infinitely repeated game, he may think that he will cooperate in week 2 if he and the other person cooperate in week 1. However, he may change his mind to defect in week 2 if his preferences in week 2 are different from those in week 1. In other words, if a subject has time consistent preferences, he is more likely to commit to his *ex ante* plan of action than a subject with time inconsistent preferences. Eliciting subjects' plan of action in week 1 and comparing it to their actual behavior in week 2 will enable us to test for the following question.

Question 6 (Time consistency). *In week 2, is a subject who exhibits time consistency less likely to deviate from his/her plan of action specified in week 1 than a subject who exhibit/s time inconsistency?*

2.3.3 Time Invariance and Cooperation

One interesting feature of a repeated game is its recursive structure. For instance, a subgame that starts in week 2 has the identical structure with that of an original repeated game that begins in week 1. After week 1, if there is no drastic change in something that may affect cooperation (e.g., beliefs about the other person's action), choosing an action in week 2 can be regarded as the same intertemporal decision making that subjects faced in week 1. This implies that under some conditions, changes in subjects' action in week 2 may be due to changes in their time preferences over week 1 and 2.

A subject with time invariant preferences should have identical time preferences at different time points. That is, if everything else is equal, decisions made by subjects with time invariant preferences do not depend on specific dates. Then, it may be natural to expect that after mutual cooperation occurs in week 1, subjects are more likely to cooperate in week 2 if they have time invariant preferences. This leads us to have the following question.

Question 7 (**Time invariance**). *After mutual cooperation in week 1, are subjects with time invariant preferences less likely to defect than subjects with time variant preferences in week 2?*

2.4 Results

4 sessions were conducted between September 2015 and February 2016.[25] In addition, one more session was conducted on October 14th to test for the time consistency hypothesis. Subjects who reside in U.S. were eligible for participating in the experiment, and a total of 1,355 subjects participated in the experiment.[26] Table 2.2 represents the information of each session.

Table 2.2: Session information

	Starting date	Subjects	Length: session / repeated game (weeks)	Strategy elicitation
Session 1	Sep. 30th	270	6 / 5	No
Session 2	Oct. 14th	277	5 / 4	No
Session 3	Oct. 14th	265	5 / 4	Yes
Session 4	Oct. 21st	268	4 / 2	No
Session 5	Nov. 11th	275	11 / 10	No

[25]The 4 sessions were conducted on September 30th, October 14th, October 21st, and November 11th. These dates were chosen for the time preference elicitation task that for all of the intertemporal choices, the two options, a sooner and a later payment, were located in the same month on the calendar which was shown to subject's monitor. This effort is to avoid possible biases that if a sooner and a later payments were in different months of the calendar, subjects might perceive a sooner payment even closer to the timing on which the corresponding decision was made.

[26]MTurk provides an option that a requester can set up eligibility conditions for participation such as country of residence and workers' reputation regarding their performance in the previous tasks. For the experiment we restricted to subjects who reside in U.S.. Even with this eligibility condition, we could find some subjects whose ip addresses are outside U.S. by checking subject's ip address after the experiment. In the analysis below, we exclude subjects whose ip addresses are outside U.S.. This criterion results in 1,344 subjects. Inclusion of such subjects, however, does not affect any of the results qualitatively.

2.4.1 General Description of Behavior

We first describe the overall behavior of participation and cooperation in the experiment. The sessions were successfully conducted as a substantial proportion of subjects took part in the experiment over time. The left panel in Figure 2.3 shows the number of subjects in each week and the right panel shows the rate of subjects that returned relative to subjects who participated in the previous week. Starting with around 270 subjects for each session, there is decay of participation over time until the experiment ends. After about attrition of 15% of subjects between week 0 and week 1 in all sessions, the rate of subjects that returned to the next week is in general increasing.[27]

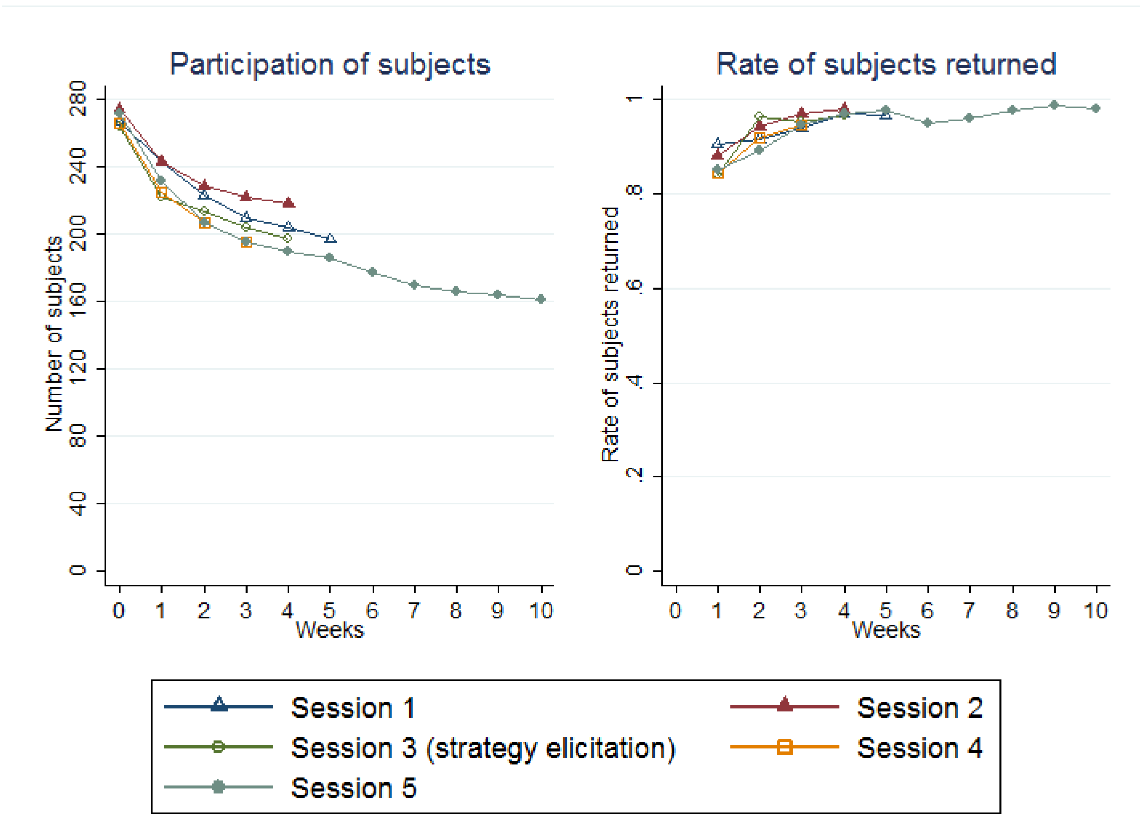


Figure 2.3: Subjects' participation over time

Figure 2.4 shows the average rate of cooperation of all subjects over time for each

[27]The exception is week 6 of the session 5 which was the week of Christmas. Even in this week, 94.8% of subjects returned.

session. In week 1 the average rate of cooperation is 78.1%, and after week 1 the rate is slightly decaying until week 5, resulting in the overall cooperation rate of 71.1% for all weeks. These rates of cooperation are higher than the rates of cooperation in the benchmark paper, Dal Bó and Fréchette (2011). In their experiments the average rate of cooperation for the first stage game of the first repeated game with the same parameters was 56.8% and the average rate of cooperation for the first repeated game was 56.1%. Note that there were two sessions which started on October 14th, with and without elicitation of a plan of action. As in Dal Bó and Fréchette (2015), cooperative behavior in theses sessions are not significantly different in week 1 and in all weeks (p-values: 0.191 and 0.999, respectively).[28] Therefore, all 5 sessions are included the analyses below.

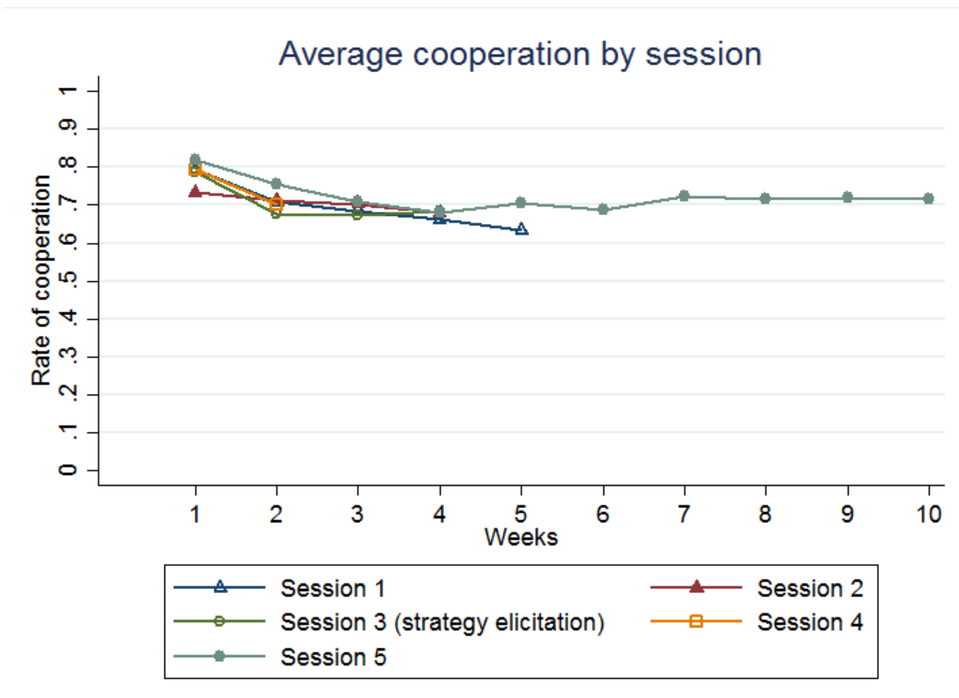


Figure 2.4: Average cooperation over time

[28]P-values are assessed by probit regressions. In all regressions throughout the paper, unless specified otherwise, standard errors are clustered at the level of pairs.

2.4.2 Distribution of Time Preferences

Before examining the effects of time preferences, we present the distribution of time preferences. First, Figure 2.5 shows the distribution of β and δ. There is a negative relationship between δ and β by construction. As subjects' choices in block 1 and 2 are highly correlated (Spearman's rank correlation coefficient: 0.755 with p-value<0.001), 71.1% subjects have β equal to 1, i.e., stationary time preferences. 16.4% and 12.5% subjects have present bias ($\beta < 1$) and future bias ($\beta > 1$), respectively. In an aggregate level, the average of δ and β are 0.906 and 0.995, respectively.[29]

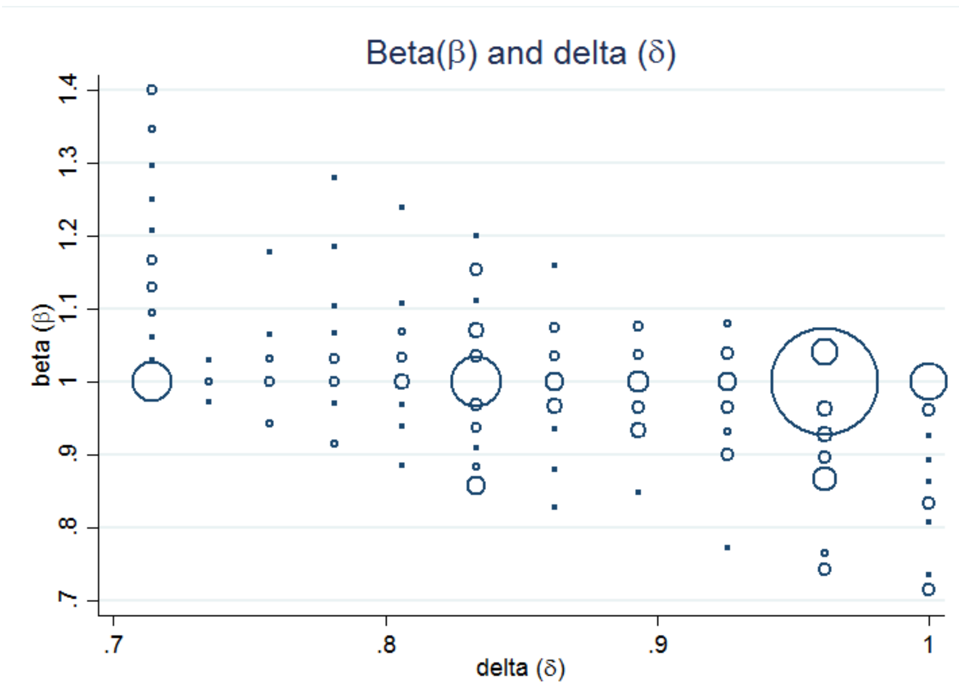


Figure 2.5: Distribution of β and δ

We also look at the distribution of three properties of time preferences á la Halevy (2015). Table 2.3 compares the classification of subjects in Halevy (2015) and this exper-

[29]These numbers are quite consistent with the numbers in previous experiments. For instance, Halevy (2015) used two different amounts of money to measure subjects' time preferences - \$10 and \$100. The average of estimated β are 0.994 and 1.00 for \$10 and \$100, respectively, and are consistent with the average of β in this paper. The average of estimated δ are 0.944 and 0.963 for \$10 and \$100, respectively, and are higher than the average of δ in this paper. These differences may be due to the "Magnitude effect" that small outcomes are usually discounted more than large ones. See Frederick et al. (2002) for more experimental evidence.

iment.[30] It seems that the distributions of time preferences in both experiments are very similar. Interestingly, the proportion of subjects who have stationary, time consistent, and time invariant preferences is higher in this experiment than in Halevy (2015) which used undergraduate students as subjects. Confirming these facts helps us validate the quality of the data in this experiment because, compared to laboratory experiments, it is harder to have controls over subjects and environments in online experiments.

Table 2.3: Classification of subjects

	This experiment	Halevy (2015) - $10	Halevy (2015) - $100
	%	%	%
Time Invariant Stationary ($x_1 = x_2 = x_3$)	45.95	38.07	35.80
Time Invariant Non-Stationary ($x_2 \neq x_1 = x_3$)	5.88 (3.47)	7.95 (5.68)	10.23 (4.55)
Time Varying Stationary ($x_1 = x_2 \neq x_3$)	25.14	24.43	21.02
Non-Stationary consistent ($x_1 \neq x_2 = x_3$)	9.06 (4.91)	13.64 (5.68)	10.23 (5.11)
Time Varying Non-Stationary Inconsistent	13.97 (7.71)	15.91 (5.68)	22.72 (11.36)
Total	100	100	100

Note: Present bias is identified if $x_2 < x_1$. Numbers in parentheses indicate the proportion of subjects who have preferences consistent with present bias.

2.4.3 β, δ, and Cooperation

We first examine the effects of β and δ on cooperation. Table 2.4 presents results from probit regressions and corresponding marginal effects in which the effects of β and δ on

[30] Halevy (2015) used two different treatments for eliciting time preferences. Since the results from the two treatments are quite similar, we report the total number of subjects over the two treatments in table 2.3. See Halevy (2015) for more details.

cooperation are tested. In week 1 and all weeks, we find that β is the only predictor of cooperation in an infinitely prisoner's dilemma. In column 1 and 3, β has marginally positive impacts on cooperation in week 1, and column 2 and 4 reveal that subjects with higher β significantly cooperate more in greater weeks. Surprisingly, δ is found not to be significantly correlated with cooperation.

Table 2.4: β, δ, and cooperation

	Probit		Marginal effects	
	(1) Week 1	(2) All weeks	(3) Week 1	(4) All weeks
β	1.134*	1.310**	0.330*	0.443**
	(0.675)	(0.525)	(0.196)	(0.177)
δ	-0.419	0.109	-0.122	0.037
	(0.527)	(0.422)	(0.153)	(0.143)
Constant	0.0439	-0.832		
	(0.910)	(0.703)		
Observations	1,165	5,107	1,165	5,107

Notes: Dependent variable: cooperation=1, defection=0. Clustered standard errors in parentheses. Marginal effects are taken at the mean.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

There may be some reasons of why we find no significant relationship between δ and cooperation. First, as shown in Dal Bó and Fréchette (2011), subjects learn about playing an infinitely repeated game as whey gain experience. For instance, in the treatment with the same continuation probability and payoffs used in this experiment, the average rate of cooperation starts at around 50% in the first repeated game and converges to the full level of cooperation as subjects gain experience. This may indicate that having multiple repeated games can allow subjects to learn about their optimal behavior in a repeated game, leading to improve the relationship between δ and cooperation.

Second, the multiplicity of equilibrium may make it harder for δ to be associated to cooperation. Note that under the range of δ that could be measured in this experiment, both cooperation and defection can be supported as equilibrium outcomes. That is, rather than being a threshold over which cooperation can be an equilibrium outcome, δ may be expected to guide subjects with higher δ to coordinate on cooperative outcomes. However, as shown in Dal Bó and Fréchette (2011), determining criteria for equilibrium selection in repeated games is a challenging problem. Relatedly, the weekly time horizon for a repeated game may be too short to draw meaningful implications of δ on cooperation. This implies that heterogeneity in discount factors over a short time horizon may not be salient enough so that a discount factor may not work as a primary determinant of cooperation.[31]

2.4.4 Time Consistency and Deviation from the Plan of Action

To study whether time consistency affects cooperative behavior, we compare subjects' behavior in week 2 with their plan of action specified in week 1. Table 2.5 represents results from probit estimations in which the session with elicitation of a plan of action is considered. Column 1 and 4 reveal that in week 2, subjects with time consistent preferences are significantly less likely to deviate from their plan of action than subjects with time inconsistent preferences. Looking separately at deviations by the different paths of history allows us to see the explanatory power of time consistency in a clearer manner. Column 2 and 5 only deal with deviations from a plan of action in which mutual cooperation occurred in week 1, and we find that the deviations are highly significantly correlated with time inconsistency. In other words, among 132 subjects who observe mutual cooperation in week 1, 97.2% of the subjects originally have planned to cooperate after mutual cooperation in week 1, and most of subjects who deviate from this plan of action are time inconsistent. For other paths of histories, 37.0%, 17.2%, and 6.2$ of subjects planned to cooperate after CD, DC, and DD, respectively, and after DC and DD, time consistent subjects are less likely deviate from their plan of action, but these correlations are not significant. This may be due to a lack of observations for each path

[31]Note that identification of β does not depend on the length of time horizons. This may support the results of this paper that we only find significant relationship between β and cooperation.

of histories (27, 29, and 16 observations for CD, DC, and DD, respectively).

Table 2.5: Time consistency and deviation from the plan of action

	Probit			Marginal effects		
	(1) All histories	(2) After CC	(3) Other	(4) All histories	(5) After CC	(6) Other
Time consistency (=1)	-0.475**	-0.999***	-0.138	-0.134**	-0.179***	-0.052
	(0.184)	(0.299)	(0.259)	(0.053)	(0.058)	(0.098)
Constant	-0.587***	-0.787***	-0.272			
	(0.145)	(0.196)	(0.222)			
Observations	204	132	72	204	132	72

Notes: Dependent variable: deviation from a plan of action=1, otherwise=0. Clustered standard errors in parentheses. Marginal effects are taken at discrete change from 0 to 1 for dummy variables.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

This result may provide an insight on why people fail to maintain cooperative relationships. Deviations from an *ex ante* plan of action can be viewed as a conflict between dual selves: long-run self vs. short-run self. For example, Fudenberg and Levine (2006) and many other papers propose a dual-self model in which the conflict between long-run and short-run selves may result in time inconsistent behavior. One interesting question would be whether time inconsistent subjects are sophisticated enough that they demand commitment devices to hole up their decisions for the future in strategic interactions over time.

2.4.5 Time Invariance and Breaking the Cooperative Relationship

In this subsection we test the idea of whether time invariant subjects are less likely to deviate from mutual cooperation than time variant subjects in week 2. Table 2.6 shows the results from probit estimations in which behavior in week 2 is examined. Column 1 and 3 consider behavior in week 2 after all histories and column 2 and 4 consider behavior in week 2 only after mutual cooperation in week 1. Column 1 and 3 show that after all histories, time invariance is not significantly correlated with deviations from subjects' action chosen in week 1. However, in column 2 and 4, it is clear that time invariance is significantly and negatively correlated with subjects' deviation from mutual cooperation that happened in week 1.[32]

Table 2.6: Time invariance and deviation from mutual cooperation in week 1

	Probit		Marginal effects	
	(1) All histories	(2) After CC	(3) All histories	(4) After CC
Time invariance (=1)	-0.123	-0.598***	-0.034	-0.065***
	(0.086)	(0.163)	(0.024)	(0.018)
Constant	-0.799***	-1.333***		
	(0.064)	(0.099)		
Observations	1,028	644	1,028	644

Notes: Dependent variable: deviation from an action in week 1=1, otherwise=0. Clustered standard errors in parentheses. Marginal effects taken at discrete change from 0 to 1 for dummy variables.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Looking at mutual cooperation that lasted longer than the first two weeks, we find no effect of time invariance on deviations from mutual cooperation. It seems that observing mutual cooperation in the first two weeks works as an almost sufficient condition for cooperation in later weeks. After week 2, more than 97.5% of subjects who observe no defection in previous weeks keep cooperating in week 3 and after, suggesting that trust established between matched subjects may outweigh the effects of time variant preferences on cooperation.

[32]We also check whether time invariance may affect cooperation after mutual defection in week 1 and find no significant correlation. This may imply that changes in time preferences over week 1 and 2 may not be the main reason of converting an action from defection to cooperation after mutual defection in week 1.

2.5 Discussion

Since we find no relationship between δ and cooperation, it is important to investigate whether β and δ are correlated with another intertemporal decision making to validate what we elicited from subjects. Participation of subjects in the experiment over time can be interpreted as intertemporal decision making because there is a tension between immediate benefits of attrition vs. later, but larger overall benefits of participation. Therefore, in this subsection we examine whether β and δ measured at the beginning of the experiment can predict attrition and the length of participation of subjects in later weeks.[33] If β and δ measured subjects' time preferences, it is natural to expect that subjects with higher β and higher δ are (1) less likely to drop, and relatedly, (2) more likely to participate longer in the experiment.

Table 2.7: Time preferences, attrition, and the length of participation

	Probit - marginal effects			Tobit	OLS
	(1)	(2)	(3)	(4)	(5)
	All weeks	After week 0	After week 1	Up to 4 weeks	All weeks
β	-0.491***	-0.369***	-0.090	4.292**	3.445**
	(0.564)	(0.138)	(0.122)	(1.836)	(1.432)
δ	-0.401***	-0.165	-0.228**	3.600**	2.396**
	(0.151)	(0.114)	(0.088)	(1.465)	(1.064)
Constant				-1.205	-1.390
				(2.537)	1.928)
Observations	1,345	1,345	1,164	1,345	1,345
R-squared					0.006

Notes: Dependent variable: for probit regressions, attrition=1, otherwise=0. For tobit and OLS, the dependent variable is a number of weeks in which subjects participated. Standard errors in parentheses are clustered at the level of individuals. Marginal effects are taken at discrete change from 0 to 1 for dummy variables.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 2.7 presents the results from probit estimations which test for these hypotheses.

[33]It is less likely that subjects' time constraints would affect their participation. We allow subjects to have at least 48 hours to complete their tasks upon receiving the invitation email, and subjects know that it takes less than 10 minutes for completing each week's task.

Column 1-3 are the marginal effects from probit regressions in which the dependent variable equals 1 if attrition happens before the session has finished and 0 otherwise. Column 1 takes attrition of subjects for all weeks of the experiment. Column 2 and 3 look at attrition in a more subtle sense in that we focus on the subjects who leave the experiment after week 0 and week 1, respectively. Overall subjects with higher δ and β are significantly less likely to drop off of the experiment. Looking at attrition after week 0 and after week 1 separately allows us to distinguish different effects of β and δ on attrition of subjects. Attrition after week 0 is significantly and negatively correlated with β, but not δ, and attrition after week 1 can only be explained by δ.

Based on these results we further investigate whether β and δ can predict how long subjects will take part in the experiment over time. It is important to note that as a dependent variable, a number of weeks in which subjects participated is censored from above that subjects could have participated longer unless the session had finished. Therefore, it is appropriate to use a Tobit regression to account for the characteristic of the censored data. In addition, we take the data only up to week 4 because the shortest session had lasted only for 4 weeks. In column 4 and 5 we present the result from Tobit regressions as well as that of OLS which takes the dependent variable not censored from above. In both regressions, we find that both β and δ are significant predictors of participation of subjects. Subjects with higher β and δ are significantly more likely to participate in the experiment for more weeks, implying that such subjects are more able to maintain cooperative relationships in our experiment.

Taken together, this result may suggest that δ can reflect subjects' discount factor. Otherwise, it is less likely to happen δ to predict participation of subjects over time. Given that we find no significant relationship between δ and cooperation, assessing individual's time preferences as a determinant of cooperation in repeated games over time would be a challenging problem. Not only time preferences matter, but also beliefs about the counterpart's cooperativeness would be important in coordinating on cooperative outcomes.

Relatedly, if a time horizon over which payoffs can be discounted is too short, the effects of time preferences on cooperation may be outweighed by the effects of other factors on cooperation. At the same time, subjects may need to gain experience for realizing intertemporal consequences of cooperation and defection. One possible remedy would be to extend time horizons that subjects have lower and more diversified discount factors and allow subjects to play many repeated games. We leave this possibility for future research.

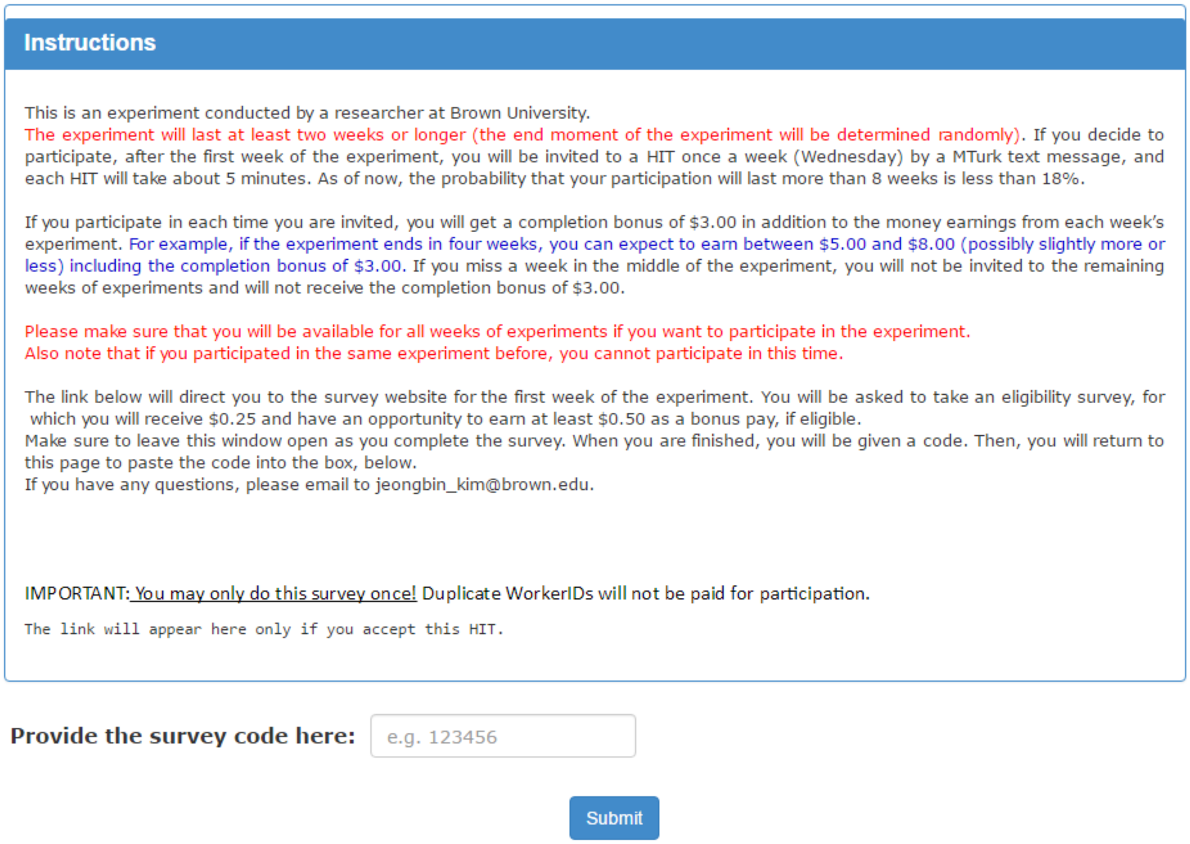
2.6 Conclusions

In this paper we measure and relate subjects' time preferences to their behavior in a prisoner's dilemma game. We implement a novel experimental design for a repeated game in which subjects play one stage game and receive associated payoffs each week. This allows us to examine the effects of time preferences on cooperation. First, we find that the degree of present bias is negatively correlated with cooperation in week 1 and all weeks. Surprisingly, we find no significant relationship between a discount factor and cooperation. Second, subjects with time consistent preferences are less likely to deviate from their plan of action. Third, we find that subjects with time varying preferences are more likely to break cooperative relationships. Finally, the degree of present bias and the discount factor measured at the beginning of the experiment can predict attrition and the length of participation of subjects in later weeks.

In the literature many previous works found that measured time preferences are related to individuals' decisions over time - credit card borrowing (Meier and Sprenger, 2010), adolescents' alcohol consumption and obedience to the school's code (Sutter et al., 2013), and smoking (Harrison et al, 2010). To the best of my knowledge, this paper is the first attempt to explore the role of time preferences on strategic interactions over time. Studying strategic interactions over time is important because people usually interact with others repeatedly over time in many real world situations. For instance, in firms and other organizations, many projects are conducted based on the unit of a team or a partnership which may be maintained for months or years. The novel experimental design suggested in this paper may well proximate such environments. Therefore, the experimental evidence provided here can shed light on the question from personnel economics and management: who is going to work (or shirk) on a team based task? We hope this paper can be one small step to enhance the validity of experiments and our understanding toward the real world.

2.7 Appendix

2.7.1 The recruitment screen of the MTurk page



Instructions

This is an experiment conducted by a researcher at Brown University.
The experiment will last at least two weeks or longer (the end moment of the experiment will be determined randomly). If you decide to participate, after the first week of the experiment, you will be invited to a HIT once a week (Wednesday) by a MTurk text message, and each HIT will take about 5 minutes. As of now, the probability that your participation will last more than 8 weeks is less than 18%.

If you participate in each time you are invited, you will get a completion bonus of $3.00 in addition to the money earnings from each week's experiment. For example, if the experiment ends in four weeks, you can expect to earn between $5.00 and $8.00 (possibly slightly more or less) including the completion bonus of $3.00. If you miss a week in the middle of the experiment, you will not be invited to the remaining weeks of experiments and will not receive the completion bonus of $3.00.

Please make sure that you will be available for all weeks of experiments if you want to participate in the experiment.
Also note that if you participated in the same experiment before, you cannot participate in this time.

The link below will direct you to the survey website for the first week of the experiment. You will be asked to take an eligibility survey, for which you will receive $0.25 and have an opportunity to earn at least $0.50 as a bonus pay, if eligible.
Make sure to leave this window open as you complete the survey. When you are finished, you will be given a code. Then, you will return to this page to paste the code into the box, below.
If you have any questions, please email to jeongbin_kim@brown.edu.

IMPORTANT: You may only do this survey once! Duplicate WorkerIDs will not be paid for participation.

The link will appear here only if you accept this HIT.

Provide the survey code here: e.g. 123456

Submit

Figure 2.6: The recruitment screen of the MTurk page

2.7.2 The instructions for week 0 (Qualtrics)

Instructions

1. The task for this week's experiment is that you choose (1) the amount of and (2) the timing of a bonus that you will receive for completing the survey.
Please indicate for each of the two blocks of 10 decisions, whether you would prefer the payment closer to today or the payment in the future.

2. For each row, you can choose one payment: Option A, the sooner payment or Option B, the later payment.
For each block of 10 decisions, one decision will be randomly selected and paid to you as a bonus on the promised day.

3. When you finish this survey, you will be notified of which two decisions are selected for your payment.
In this page, the first block of 10 decisions will be shown below.

November 2015

S	M	T	W	T	F	S
1	2	3	4	5	6	7
8	9	10	11 (HIT started)	12	13 (Option A)	14
15	16	17	18	19	20 (Option B)	21
22	23	24	25	26	27	28
29	30					

December 2015

S	M	T	W	T	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

4. The following is the first block of 10 decisions.
In this block, you're asked to choose between an option which will pay you on November 11th and an option which will pay you on November 20th.
In the calendar above, you can see when this HIT begins and when each option will pay you on each promised day.
Remember that one of these 10 decisions will be randomly selected and paid to you as a bonus.

	Option A	Option B	
$0.50 guaranteed on November 11th	○	○	$0.50 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.52 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.54 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.56 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.58 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.60 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.62 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.64 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.66 guaranteed on November 20th
$0.50 guaranteed on November 11th	○	○	$0.68 guaranteed on November 20th

Figure 2.7: The instructions for block 1 in week 0

Instructions

1. The task for this week experiment is that you choose (1) the amount of and (2) the timing of a bonus that you will receive for completing the survey.
Please indicate whether you would prefer the payment closer to today or the payment in the future.

2. For each row, you can choose one payment: Option A, the sooner payment or Option B, the later payment.
For each block of 10 decisions, one decision will be randomly selected and paid to you as a bonus on the promised day.

3. When you finish this survey, you will be notified of which two decisions are selected for your payment.
In this page, the second block of 10 decisions will be shown below.

November 2015

S	M	T	W	T	F	S
1	2	3	4	5	6	7
8	9	10	11 (HIT started)	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

December 2015

S	M	T	W	T	F	S
		1	2	3	4 (Option A)	5
6	7	8	9	10	11 (Option B)	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

4. The following is the second block of 10 decisions.
In this block, you're asked to choose between an option which will pay you on December 4th and an option which will pay you on December 11th.
In the calendar above, you can see when this HIT is started and when two options will pay you on each promised day.
Remember that one of these 10 decisions will be randomly selected and paid to you as a bonus.

	Option A	Option B	
$0.50 guaranteed on December 4th	○	○	$0.50 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.52 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.54 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.56 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.58 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.60 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.62 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.64 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.66 guaranteed on December 11th
$0.50 guaranteed on December 4th	○	○	$0.68 guaranteed on December 11th

Figure 2.8: The instructions for block 2 in week 0

2.7.3 The instructions for a stage game in week 1 (Qualtrics)

Instructions

1. Welcome to the 2nd week of the experiment.
In this experiment you will be asked to make one decision for each HIT per week until the experiment ends.

2. The length of the experiment is randomly determined. After each week's HIT, there is a **75%** probability that the experiment will continue for at least one more week. This is as if we would draw a ball from a jar which contains 75 white balls and 25 red balls. If a white ball is selected, the experiment continues for at least for one more week and so on. Otherwise, the experiment will be finished. You will be notified of this result when you will be paid the bonus each week.

3. In the course of the experiment, you will be paired with the same person. But, you will not be provided with any information about the person you are matched with, and vice versa. In other words, all decisions will be made in an anonymous manner.

4. You will be asked to choose either action "1" or action "2". The choices and the payoffs in each task are as follows:

	The other's choice	
Your choice	1	2
1	\$0.48, \$0.48	\$0.12, \$0.50
2	\$0.50, \$0.12	\$0.25, \$0.25

The first entry in each cell represents your payoff, while the second entry represents the payoff of the person you are matched with.
As you can see, this shows the payoff associated with each choice. Once you and the person you are paired with make choices, those choices will determine your and the other's payoffs.

That is, if:

(1) You select 1 and the other selects 1, each receives \$0.48.

(2) You select 1 and the other selects 2, you receive \$0.12 while the other receives \$0.50.

(3) You select 2 and the other selects 1, you receive \$0.50 while the other receives \$0.12.

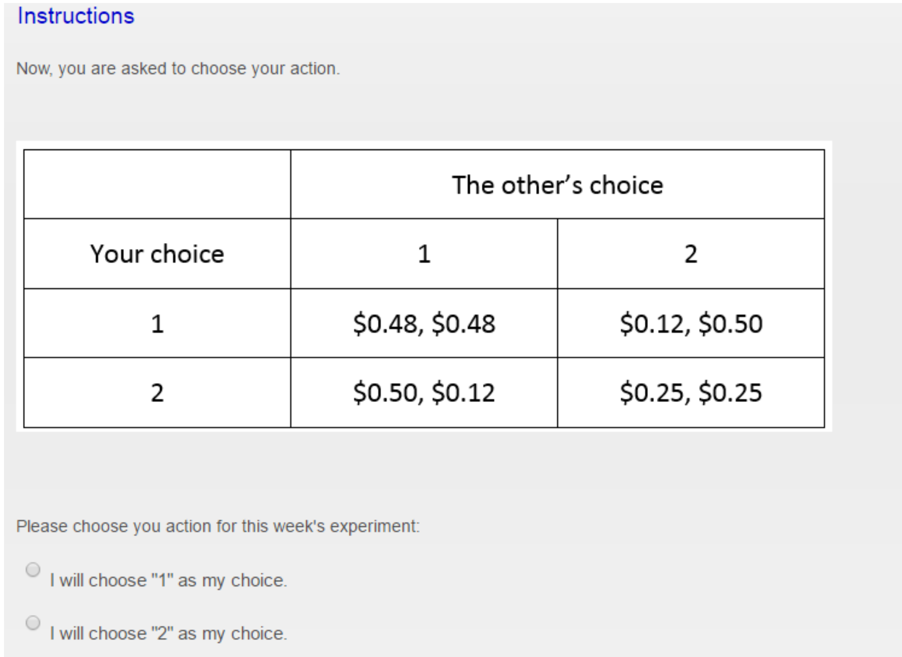
(4) You select 2 and the other selects 2, each receives \$0.25.

Before you start, let me remind you that:

1. If you do not complete the HIT in three days after you receive the invitation MTurk text message, you will not be allowed to participate in the remaining weeks of HITs. Of course, in this case, the completion bonus will not be paid to you.

2. In the event, which we hope is unlikely, that your counterpart fails to continue the interaction in a future week, we will arrange for you to be able to continue playing with another counterpart or will make other arrangements to minimize any impact on your predicted earnings. Unless we inform you otherwise, you will definitely be playing with the same counterpart at each future stage.

Figure 2.9: The instructions for a stage game in week 1



Instructions

Now, you are asked to choose your action.

	The other's choice	
Your choice	1	2
1	\$0.48, \$0.48	\$0.12, \$0.50
2	\$0.50, \$0.12	\$0.25, \$0.25

Please choose you action for this week's experiment:

I will choose "1" as my choice.

I will choose "2" as my choice.

Figure 2.10: The decision screen for a stage game in week 1

Figure 2.11: The screen for specifying the plan of action in week 1

CHAPTER 3

TRUST AND COOPERATION: AN EXPERIMENTAL ANALYSIS

3.1 Introduction

Trust has been regarded as an important factor to affect various aspects of economic prosperity, one of which is growth (Knack and Keefer, 1997; La Porta et al., 1997). In particular, Zak and Knack (2001) empirically show that trust affects growth: trust relies on the social, economic and institutional environments which accrue transactions, and higher trust reduces transaction cost, which in turn engenders higher investment rate and faster economic growth. Their conclusion echoes Arrow's (1972) argument that "Virtually every commercial transaction has within itself an element of trust,. . . , much of the economic backwardness in the world can be explained by the lack of mutual confidence", since a prerequisite for the successful development of market economies is to enlarge interactions to anonymous others (Algan and Cahuc, 2010).

In many situations where self-interest might otherwise lead to free-riding, as suggested by Fukuyama (1995), cooperation is a key to enhance efficiency in the real world: voluntary provision of local public goods, cooperation among partners of enterprises facing profit-sharing schemes, and effort to establish better institutions with less theft and corruption. Cooperation in these domains is an important contributor to overall economic efficiency and thus growth.[1]

However, an empirical question which remains to be answered is to reveal specific mechanisms behind the way in which trust affects cooperation. One plausible story for explaining the associations between trust and cooperative outcomes originates from beliefs: people cooperate because they "believe" other society members will also cooperate

[1] There has been a wide range of related discussion in the literature. For instance, see Ostrom (2010).

or other society members have expectations of high cooperation among themselves.[2] In other words, members in a society with high trust may have succeeded to have optimistic and shared beliefs about others' behavior or beliefs, and consequently to have high cooperation.

The main problem of this approach is hard to identify the effect of beliefs on cooperation in the real world since societies or groups have been formed endogenously. Although many papers posit the importance of trust by highlighting its effects on economic growth, proving specific mechanisms in which trust promotes prosocial behavior through the lens of beliefs can be difficult with observational data. That is, it is unclear whether beliefs lead to cooperation or vice versa.

In this paper we present a laboratory experiment to shed light on the role of beliefs as a channel through which cooperation on the provision of public goods can be promoted or deterred. Subjects are first asked to play a trust game as both roles - the first and second mover. Then, they move to the next phase to play a voluntary contribution mechanism (hereafter VCM) or linear public goods game. In order to manipulate beliefs about the other members in a group, we use trust game behavior as the baseline to generate five different environments with different levels of lab-measured trust and lab-measured trustworthiness: a group in which people are randomly matched, groups in which the average level of trust is high/low, and groups in which the average level of trustworthiness is high/low, respectively. Using a strategy method for VCM allows us to randomly assign subjects to these environments with different levels of trust/trustworthiness, and this is a key of our experimental design to manipulate subject' beliefs about others. Each subject plays a one-shot VCM consecutively in those five environments, and their first-order and second-order beliefs are elicited.

These two games are chosen with careful consideration of what we can infer from

[2] That is, we study two different kinds of beliefs: beliefs about others' cooperation (first-order belief) and beliefs about others' beliefs about cooperation (second-order belief). Formal definitions of these beliefs will be discussed later.

subjects' behavior. We use a trust game in the first phase as there is ample evidence that behavior as the first and second mover in this game can well approximate different aspects that are embedded in the various definitions of trust. VCM in the second phase is of particular interest because it presents a multi-person social dilemma which resembles many situations in the real world where full cooperation leads to efficiency.[3]

With elicited subject's beliefs about others and their willingness to cooperate in each environment, we show that people positively associate both trust and trustworthiness with cooperation, and that they are approximately equally more cooperative when in a high trusting as when in a highly trustworthy environment. By looking at the effects of the first-order and the second-order beliefs on cooperation separately, we also find that the effects of the first-order beliefs outweigh those of the second-order beliefs, implying that "reciprocity" is a main channel to lead to the high level of public goods provision.[4]

The results of this paper can bring over some important implications outside the laboratory. This paper tries to identify one channel of how lab-measured trust and trustworthiness, each representing a different aspect of the general trust question, can affect the level of cooperation through the beliefs about others' cooperation. From a policy perspective, it implies that for any policy that aims to enhance social capital as a mean to stimulate economic growth, it needs to ensure widely-accepted beliefs about others that one's own vulnerability will be not exploited. In other words, without a common

[3]Contribution behavior in this game has been thoroughly studied in the literature (see the survey paper by Chaudhari, 2011), and one of the notable findings is the presence of conditional cooperators (Fischbacher et al., 2001). Making up the majority of the population (Page et al., 2005; Keser and van Winden, 2000), conditional cooperator's contribution to the VCM is positively correlated with their ex ante beliefs about the contributions by their group members in repeated games. Because people hold different beliefs in different environments in our study, conditional cooperators choose to contribute more in high trust or trustworthiness group.

[4]What we mean by reciprocity in this paper refers to conditional cooperation which represents positive associations between cooperation and the first-order beliefs about the other group members' contributions. Theoretically speaking, conditional cooperation arises if participants are reciprocal (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006), inequality averse (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), or both. To distinguish different motivations behind conditional cooperation is not our purpose. This paper aims to show (1) how the different group level of trust and trustworthiness can affect beliefs about others and (2) how these trust- and trustworthiness-driven beliefs can consequently affect cooperation in VCM.

knowledge that members in a society are prone to trust others and/or are trustworthy, implementing a policy that calls for collective efforts may confront great difficulties.

3.2 Related literature

3.2.1 Measuring trust

The endeavor to measure trust in order to prove how trust can be related to various aspects of economic growth has a long tradition. Many studies in applied works approximate trust by using the answers to the standard World Values Surveys (WVS) trust question ("Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"). Knack and Keefer (1997) find that higher trust is conducive to growth for a sample of 29 market economies. La Porta et al. (1997) demonstrate similar evidence of trust on civic participation, and Guiso et al. (2009) corroborate that trust is positively related to volume of international trade, both of which lead to higher productivity and thus faster growth of a society (Putnam, 1995; Wagner, 2007; Lee, 1995).

At the same time, vagueness and lack of agreement about what the survey measure of trust truly captures evoke one stream of research in the experimental economics literature.[5] In early attempts, many papers adopt the canonical trust game introduced by Berg, Dickhaut, and McCabe (1995).[6] Assuming preferences such as altruism (Cox, 2004; Ashraf et al., 2006) are not the sheer motivations behind the act of the first mover, the sending behavior encapsulates trust. The central element of the trust decision is the tradeoff between exposing oneself to the risk of being "exploited" and achieving more efficient outcomes (Thoni, 2015). Gauging risk is equivalent to forming beliefs about the second

[5]Empirically, such a measure could only pick up the underlying influences of some fixed societal features such as quality of institutions (Acemoglu et al, 2001) and the extent of fractionalization (Alesina et al, 2003).

[6]In this trust game, the first mover can transfer money to a second mover, and the amounts get tripled upon reaching the second mover, after which the second mover can choose to send money back to the first mover.

mover's trustworthiness; as a result, the trusting behavior in this game encompasses two distinct components: beliefs in others' trustworthiness and specific preferences of the sender (Sapienza et al., 2013).

Using controlled laboratory experiments, researchers investigate how the survey measures of trust are related to specific behavior in experimental games, and have reached mixed conclusions. For example, Glaeser et al. (2000) and Lazzarini et al. (2003) show that the answers to the WVS trust question are not significantly correlated with first movers' sending behavior, but pertinent to second movers' returning behavior in the trust game. In contrast, Fehr et al. (2003) and Bellemare and Kroeger (2007) have opposite results that while first movers' sending behavior correlates with the answers to WVS-like questions, the relationship fails to hold between second movers' returning decisions and those survey answers. The contradictory findings suggest that survey-measured trust could be significantly correlated with both lab-measured trust and lab-measured trustworthiness, which seems to imply that not only do trust and trustworthiness, as measured in the game, appear to be non-separable (Fehr, 2009), but people tend not to distinguish them in real life. Such "non-separability" comes from the fact that beliefs in trustworthiness of others plays a significant role in explaining why sending varies (Thoni et al., 2012) and that players extrapolate their opponent's behavior from their own (Sapienza et al., 2013): the belief in the trustworthiness of others is often correlated with one's own trustworthiness as it is usually obtained by introspection (Varian, 1993). Nevertheless, the fact that lab-measured trust and trustworthiness correlates positively to WVS results establishes a linkage between laboratory findings and real world implications.

3.2.2 Trust and Conditional Cooperation in VCM

The VCM game is of particular interest to us because it presents a social dilemma where full cooperation, which is against self-interest incentives, leads to efficiency. Moreover, the procedure of the VCM goes as if each and every player makes decision as both trustor and trustee simultaneously (Thoni, 2015), compared to the asymmetric strategic interaction between these two roles in the trust game. On such connections, a number of existing papers investigated whether survey measure of trust or lab-measured trust is associated with contributing behavior in VCM. Using 630 subjects in rural and urban Russia, Gachter et al. (2004) show that whereas answers to the WVS trust question are not correlated with behavior in a one-shot VCM, subjects who believe that most others are fair or helpful are more likely to contribute in the VCM. Thoni et al. (2012) delve into this problem in a great detail by using a representative sample in Denmark. Subjects in their study are asked to play a VCM that elicits both yet separately their preferences to cooperate (independent of beliefs) and beliefs about others' cooperation. They find that responses to the trust questions are correlated to preferences rather than beliefs. Kocher et al. (2015) paper has a design in which a trust game followed by a VCM, and they show that in general trusting behavior correlates with the unconditional contribution in VCM.[7]

Theoretically speaking, conditional cooperation arises when participants are reciprocal (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006), inequality averse (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), or both, and these social preferences coincide with those driving the returning decision, or trustworthiness in the trust game. Consequently, it is natural to explore the cross-game correlation between trust/trustworthiness and conditional cooperation. This paper is complementary to this line of research by showing how the first and the second mover behavior in a trust game can be informative about others' contributions and different orders of beliefs in public goods game, especially for conditional cooperators.[8]

[7]Reciprocity (see survey paper by Camerer, 2003), altruism (Andreoni and Miller, 2002), inequality aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) and other preferences such as quasi-maximin preferences (Charness and Rabin, 2002) can explain the âĂIJunconditional kindnessâĂİ (Ashraf et al., 2006) of returning money to the sender in the trust game.

[8]There exists other papers investigating similar cross-game behavior associations. Dariel and Nikiforakis (2014) perform a within-subject analysis in a VCM and a gift-exchange game, and they find that conditional cooperating participants tend to reciprocate higher wages in the gift-exchange game with high levels of effort, while non-cooperators do not exhibit such tendency. Blanco et al. (2011) have their subjects play four different types of games, including VCM, to test the inequality aversion across these games, and they find a strong predictive power across games at aggregate level.

3.2.3 Manipulating beliefs

This paper is also related to the literature of measuring and manipulating beliefs in strategic interactions. Charness and Dufwenberg (2006) report a correlation between behavior and second-order beliefs, which may be affected by the second mover's pre-play communication in a modified trust game. To avoid consensus effects and achieve exogenous variations in beliefs, several attempts have been made. Dufwenberg et al. (2011) use the framing of a game as a work horse to manipulate subjects' beliefs. They show that depending on the framing (Community game or Market game) of a game, subjects have systematically different first- and second-order beliefs, which lead to changes in cooperation. Ellingsen et al. (2010) try to manipulate subjects' second-order beliefs by disclosing the first-order belief of a person who are paired with. A recent paper by Khalmetski (2017) manipulate the second-order belief by changing the probability of a game to be played where the true state of the world is only known to the sender of a message. Along with this line of research, this paper introduces another way of varying beliefs in that being informed about other group members' behavior in the previous phase can systematically vary subjects' first- and second-order beliefs, which may lead to change behavior in different environments.

3.3 Experimental Design and Procedure

Our experiment consists of two consecutive phases: the first phase involves a trust game, and the second phase includes a voluntary contribution mechanism (hereafter VCM) as well as its relevant belief elicitation in different environments.[9] Decisions in the first phase

[9] A strand of literature that motivates our design is that with multi-phase experiments where participants are shuffled and rematched in later phases based on their behavior in the first phases. For example, Gunnthorsdottir et al. (2007) show that when high contributors are grouped by the experimenter with other high contributors, their contributions in the VCM are sustained at high levels. Ones and Putterman (2004) extend this approach by varying group composition after five "diagnostic" periods in the two treatment groups, and demonstrate that differences in the inclination to cooperate have considerable persistence in the remaining periods of the repeated VCM. Similarly, Gatcher and Thoni (2010) have the subjects play a one-shot VCM to sort people's contribution levels, based on which they group people in

determine group formation in different environments in the second phase. Nonetheless, the two phases are independent in the sense that the instructions in the second phase are distributed only after the end of the first phase, to avoid strategic response and contamination of the first phase decisions. Decisions in both phases are incentivized and monetarily rewarded in the end. There is no treatment or control group, and all participants make each and every decision simultaneously.

Everything described in the following sub-sections is common knowledge among all subjects. The instructions are included in the Appendix. A total of six experimental sessions were conducted between November and December 2015 at Brown University Social Science Experimental Laboratory (BUSSEL) using the experimental software z-tree (Fischbacher, 2007). Each session consisted of 20 subjects from the diverse undergraduate student body (including numerous international students) of Brown University, where they are representative of a comprehensive set of majors. Each session lasted around 1.5 hours, and the individual average earning was $18.2, including the $5 show-up fee.

3.3.1 The Trust Game (First Phase)

We used a slightly modified version of the original trust game designed by Berg et al. (1995) in which each subject played both roles as sender and receiver. Our subjects knew they would make decisions as a sender first and then, without learning a counterpart's decision, as a recipient. They were explicitly told that the two decisions are independent, and only one of these decisions will be randomly realized and matched with a counterpart decision of another subject in the room, for payment purpose in the end at a conversion rate of $0.10 per experimental token. No feedback was given with respect to the decisions of others throughout this phase.

The basic setting of the game is similar to its original version in Berg et al. (1995): a

order in the future rounds of play; they reach the conclusion that "like-minded" groups sustain higher cooperation levels. We adopt a two-phase design by matching people in the VCM game of the second phase according to their first-phase behaviors in the trust game.

sender is endowed with 50 tokens, and decide if and how much of the endowment to send to the receiver in multiples of 5 between 0 and 50 tokens, and this sent amount get tripled upon reaching the receiver. To minimize the impact of distributional incentives (e.g, inequality aversion) and altruistic concerns of the sender, in our experiment, we followed Berg et al. (1995) in that each receiver is also endowed with 50 tokens to start with, and any amounts returned would come only from the tripled amount received. Returning decisions of the receiver are made in a contingency table using the strategy method: the receiver decides if and how much to send back to the sender conditional on each of the 11 possible received levels. The returned amount is not tripled. Payoffs for both players are their initial endowments plus any received amount minus any sent amount, if this set of decisions were realized in the end. The unique Nash equilibrium prediction is no tokens sent between both players, were they perfectly rational and payoff maximizing: the selfish receiver will not return anything, and knowing this leads the sender to refrain from sending at the beginning. Pareto improvement, on the other hand, is feasible if the receiver returns at least one third of the tripled amount he/she has received.

As discussed previously, the sending behavior in the game is affected by sender's beliefs on recipient's reliability or likelihood of returning, as well as specific individual preferences (Sapienza et al., 2013). While natural risk preferences fail to significantly explain sending behavior (Eckel and Wilson, 2004; Kocher et al., 2015), by diminishing the impact of subjects' social preferences such as inequality aversion and altruism, we focus on the belief aspect of sending and interpret the results equivalently as willingness to bear the risk of being "exploited" by the recipient. Nevertheless, we call it trusting behavior. On the contrary, no belief component is involved in returning decision, as the recipients face no uncertainty given the contingency table. The act of returning signifies trustworthiness, yet we make no attempt to distinguish which social preference (e.g, inequality aversion or reciprocity) is the main driving force behind the act.

3.3.2 The Voluntary Contribution Mechanism and Belief Elicitation (Second Phase)

In the second phase, each subject made one set of allocation decisions and two sets of belief estimations sequentially. They first played five parallel one-shot VCMs (or linear public goods game, see Isaac et al., 1984; Fischbacher et al., 2001) in groups of five, in five corresponding environments that are independent of each other. Each environment will be elaborated in details later. After they made their contribution decisions in all environments, they are prompted to provide their first-order beliefs about what the others in that group would contribute, on average, in each and every environment. After the first set of guesses, we elicited their second-order beliefs about the average of the first-order beliefs stated by each of the four other members in every single environment. In other words, each person made five contribution decisions first followed by five first-order guesses, and lastly gave five second-order estimates. Instead of letting them proceed in the order of contribution, first-order belief, and second-order belief environment by environment, this design is chosen to prevent contribution decisions from possible contamination of previously elicited beliefs. Group members remained unchanged for each set of contribution decision and first and second order belief elicitation per environment, but could be different across environments. Only one environment was randomly selected to determine payoffs for this phase in the end. To incentivize truthful estimates, following Dufwenberg et al. (2011), any first order estimate that was within one token of the true average and any second order estimate that was within one token of the true average of the first-order estimates was rewarded five additional tokens towards total earnings. The total earnings in this phase were thus the earning from the contribution decision plus any rewarded amount from the estimates. The conversion rate at this phase is \$0.2 per experimental token.

The five environments differ in terms of the true information each person see on the screen about group composition, and no other individual information is available. The groups in the first environment are randomly formed, while group memberships in the other four environments are pre-determined by the computer program based on participants' behavior in the first phase. In each session, we rank all sending decisions and returning decisions (when being sent the highest possible amount, 50 tokens) separately from the lowest to the highest, and assign the corresponding subjects numerical ranks 1-20 in the computer. This way, each subject should have two numbers identifying him/her, one from her sending decision, which reflects the "trust" rank and the other from her returning decision, which represents the "trustworthiness" rank. Ties are broken randomly. Each person is placed in a "high trust" group where the average ranking of the sending decisions of the other four group members is above 12, as well as in a "low trust" group where the average ranking of the sending decisions of the other four group members is below 8. The "high trustworthiness" and "low trustworthiness" groups are constructed similarly based on the returning rank, and each person is also placed once in each of these two groups. The following table summarizes the above information.

Table 3.1: Difference between Environments

Name of the Environment	Brief Descriptions (Ranking is in an ascending order: the lowest rank is denoted as 1, and so on)
Random	Matching is done randomly in the computer program.
High Trust	For each participant, the average sending rank of the other four group members is above 12
Low Trust	For each participant, the average sending rank of the other group members is below 8
High Trustworthiness	For each participant, the average returning rank of the other group members is above 12
Low Trustworthiness	For each participant, the average returning rank of the other group members is below 8

We revealed this process to our subjects in details in the instructions, and started the experiment only after we made sure every participant in the room fully understood the matching mechanism. Even though we associate these environments with names such as "high trust" and "low trustworthiness" here, we never made these labels explicit to our participants. All they saw on the screen were verbal descriptions using rank information. This design is to elicit the willingness of subjects to cooperate in groups of people characterized by different degrees of trust and trustworthiness.

The game we adopt is a standard one-shot VCM, where each person in a group has an initial endowment of 20 tokens. They then decide individually how much if anything to contribute to the public account, with whatever they did not contribute going to that person's individual account. Contributions to the public account are costly to the subject but all group members would benefit equally from the contributions. Specifically, the marginal per capita return (MPCR) from each contributed unit is 0.4. The payoff function for any subject i is thus:

$$\pi_i = 20 - c_i + 0.4 \cdot \Sigma_{j=1}^{5} c_j,$$

where c_i denotes the contribution from i. This game captures a social dilemma because the overall social payoffs are maximized when all group members choose to contribute all 20 tokens, but individual dominant strategy is to free ride and contribute nothing, and the Nash equilibrium predicts that no one contributes anything.

3.3.3 Procedure and Payments

We handed out hard copies of first phase instructions at the beginning of the experiment, and distributed second phase instructions after the end of the first phase. The experimenters read out loud the instructions for both phases, and clarified all questions in private. An end-of-session survey followed up at the end of the second phase. We collected demographic information such as gender, class level, race, and major as part of the main survey questions.

For all sessions excluding the first two, we asked the participants their perceived correlations between sending and returning behavior in the trust game, on a scale of 0-4, where 0 means perfect negative correlation, while 4 means perfect positive correlation. In addition, for the last two sessions, to understand the cross-game belief correlations, we elicited our subjects' estimates on their first phase behavior. In specific, we asked them to provide four guesses for the trust game: the averages of the eight lowest and the eight highest ranked sending as well as returning decisions (when being sent all 50 tokens) in the room. This is to gauge their senses of high and low trust/trustworthiness. All of these questions were not in the instructions and were thus completely unanticipated by the subjects. To encourage truthful reporting, we used a similar incentivizing device as in the main game.

At the end of the survey, each subject was shown on the screen which of his/her roles in the first phase was realized, which environment in the second phase was realized, as well as his/her earnings. For each subject in the first four sessions, the final payoff was the sum of the earnings from the first phase and the second phase. For each subject in the last two sessions, participants got this amount, plus any rewards from correct guessing in the belief elicitation questions (about the trust game) in the end-of-session survey. They got their payments in cash in sealed envelopes and were free to leave afterwards.

3.4 Theoretical background and hypotheses

We take the similar approach used in Dufwenberg et al. (2011) to make theoretical predictions of behavior in the second phase. In this section, we mainly focus on how contributing behavior can be driven by (first order or second order) beliefs across 5 different environments, rather than characterize all (possibly multiple) equilibrium in the second phase. Based on a psychological game framework, we restrict our attention to the case in which others' behavior as a first and a second mover behavior in the first phase is informative about behavior in the second phase.

3.4.1 Reciprocity

Reciprocity in general refers to behavior that is kind to those who are kind and hostile to those who are hostile. Pioneered by Rabin (1993) in a two-person setting and extended by Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006) into a multi-person and a dynamic setting, a theory of reciprocity shows that kindness depends on beliefs about what others would do and beliefs about others' kindness play an important role to determine one's reciprocal attitude. Following the formulation of Dufwenberg et al. (2011), the utility of player i is given as:

$$u_i(a_i, a_j, b_{ij}) = 20 - a_i + (0.5)\Sigma_{i \in N} a_i + Y_i \Sigma_{j \in N \backslash i} \kappa_{ij} \lambda_{iji},$$

where N is a set of players belonging to the same environment and Y_i is a coefficient measuring player i's sensitivity to reciprocal motivation. Also, κ_{ij} is player i's kindness to j and λ_{iji} is player i's beliefs about player j's kindness to i. In general kindness depends on beliefs about others' actions. Although it is a bit arbitrary, we will assume a reference for kindness as the average of two extremes of possible contributions, full contribution (=20) and no contribution (=0). Let denote b_{ji} by player j's beliefs about player i's contribution. By replacing κ_{ij} and λ_{iji} with $a_i - 10$ and $b_{ji} - 10$, we can rewrite the equation above as follows:

$$u_i(a_i, a_j, b_{ij}) = 20 - a_i + (0.4)\Sigma_{i \in N} a_i + Y_i \Sigma_{j \in N \backslash i} (a_i - 10)(b_{ji} - 10)$$

$$= 20 - a_i + (0.4)\Sigma_{i \in N} a_i + Y_i (a_i - 10)[\Sigma_{j \in N \backslash i} b_{ji} - 40].$$

In this formulation, if player i thinks, on average, the other 4 groups members are not kind ($\Sigma_{j \in N \backslash i} b_{ji} - 40 < 0$), his utility can be maximized by having $a_i = 0$. Otherwise, his utility can be maximized by $a_i = 20$ if Y_i is high enough and by $a_i = 0$ if Y_i is low enough.

3.4.2 Guilt aversion

Guilt aversion is based on the idea that people don't want to let down others' expectations on their behavior. For guilt aversion, we also follow the simple approach used by Dufwenberg et al. (2011). Let denote c_{iji} by player i's beliefs about player j's beliefs about player i's contribution, i.e., player i's beliefs about b_{ji}. Guilt aversion in VCM can be captured by the notion that if a player contributes less than what the other 4 group members believe he would contribute, he will experience disutilities from getting down others' beliefs. Then, the utility of player i can be presented as:

$$u_i(a_i, a_j, c_{iji}) = 20 - a_i + (0.4)\Sigma_{i \in N} a_i - \gamma_i \max\{0, \frac{\Sigma_{j \in N \setminus i} c_{iji}}{4} - a_i\},$$

where $\gamma_i \geq 0$ is a coefficient measuring how player i is sensitive to disutilities from guilt aversion. If $\gamma_i < 0.6$, then $a_i = 0$ is a dominant strategy for player i. However, if $\gamma_i > 0.6$, then conforming to others' beliefs would be a best response, i.e., $a_i = \frac{\Sigma_{j \in N \setminus i} c_{iji}}{4}$.

3.4.3 Hypotheses

The primary purpose of this study is to investigate how being informed about othersâĂŹ behavior in a trust game would affect players' beliefs, which would work as a driving force to contribute in VCM. We interpret theoretical connections between behaviors in two phases based on possible correlations between beliefs and actions as given.

Being informed about other group members' average sending behavior, subjects may infer how optimistic these members' beliefs are about the second mover's returning behavior. Since both sender and receiver in our trust game are equally endowed, the first mover's sending behavior can be interpreted as an investment in second mover's returning, rather than behavior due to a distributional concern. Given this interpretation, in the second phase, if subjects are in a group where the average amount sent is relatively high (low), they may reflect that on average, their group members have optimistic (pessimistic) beliefs about others' contributions since there is a structural similarity between a trust game and VCM. At the same time, subjects who were more optimistic (pessimistic) in the first phase are more likely optimistic (pessimistic) in the second phase. Therefore, a reciprocal player in the second phase will respond to his beliefs about others' contribution by increasing (decreasing) his contributions in different environments. Therefore, we have the following hypotheses:

Hypothesis 1 (trust and the first order belief): Subject who as a first mover sent a relatively high amount in a trust game have higher first order beliefs in second phase.

Hypothesis 2 (trust and the first order belief): Subjects in high (low) trust environment in the second phase will have higher (lower) first order beliefs about other group members' contributions and these beliefs will lead to higher (lower) own contributions.

One interesting empirical question is whether subjects perceive other group members' trust and trust worthiness differently. If subjects believe strong correlations between trust and trustworthiness, we can have the following hypothesis.

Hypothesis 3 (trustworthiness and the first order belief): If subjects believe there exists a significant correlation between trust and trustworthiness, subjects in a high (low) trustworthiness group in the second phase will have higher (lower) first order beliefs about other group members' contributions and these beliefs will lead to higher (lower) own contributions.

There are two different motivations for the second mover returning behavior in the first phase. First, returning behavior may signal reciprocity: given an amount sent by a first mover, an amount returning may reveal second mover's willingness to reciprocate, which is the key to interpreting conditional cooperation in VCM. Therefore,

Hypothesis 4 (trustworthiness and conditional cooperation) Subject who returned a relatively high amount in the first phase are more likely to be conditional cooperators than free riders in VCM.

The second driving force of the second mover's returning behavior is guilt aversion. When the first mover sent a substantial amount in the first phase, a second mover may believe that the first mover has high expectations about how much the second mover would return. Thus, the second mover's returning behavior rests on the good will of not letting down other's expectation. That is,

Hypothesis 5 (trustworthiness and second order beliefs) Subjects who returned a high amount as a second mover given substantial amounts sent by the first mover are more likely to have higher second order beliefs in VCM.

3.5 Results and Analyses

3.5.1 Summery Statistics of Trust Game & Voluntary Contribution Mechanism behavior

On average, our subjects sent 19.58 tokens as senders in the first phase, with a standard deviation of 15.35 tokens and an average sending percentage of 39.17%. When being sent all 50 tokens as recipients, our participants on average returned 48.68 tokens out of the tripled 150 tokens; the standard deviation was 40.90 tokens. The average returning percentages range from 30-32% for all possible contingencies. The Spearman's correlation between the average returned proportions from the second mover and the first mover sent contingencies is 0.5968 and significant at 5% level, which suggests that a reciprocity effect as developed by Cochard et al. (2004) could be observed. Since the average second mover returning percentage tends to increase with sent amounts from the first mover, the result is consistent with Rabin's (2003) original reciprocity theory at the aggregate level (Smith, 2013).[10]

Our trust game findings in the first phase are similar to previous results (see Chaudhuri, 2008; Fehr, 2009; Johnson and Mislin, 2011 for meta-analyses): there are about 17% sender that are consistent with the standard theory who sent nothing and 23% selfish recipient who returned nothing under all contingencies. At individual level, the correlation between self sent percentage and the average returned percentage is 0.5629, and this number implies a high, positive correlation between trusting and trustworthy behavior in the game.

Figure 1 below shows that, on average, our participants contributed 8.60 tokens in random environment in the second phase, which is around 40% of their initial endowment, and the standard deviation is 6.96. In high trust and and low trust environments, the average contributions are 11.48 and 4.9 with standard deviations of 7.25 and 5.75, respectively. In high trustworthiness and low trustworthiness environments, the contributions average to 11.22 and 4.60 with standard deviations of 7.07 and 5.57, respectively. Statistically speaking, based on paired t-tests, at 95% level, our subjects contribute significantly more in both high trust and high trustworthiness environments, and contribute significantly less in both low trust and low trustworthiness environments when compared to that in random environment.

The contribution differences between high and low trust as well as high and low trustworthiness environments are striking: 7-token difference would be equivalent to a 140% increase if people were to move from a low trust/trustworthiness to a high trust/trustworthiness environment; this result is robust to paired t-test at 95% level. This leads to our first result:

Result 1: People contribute significant more in high trusting/trustworthy environments than in low trusting/trustworthy environments.

[10] A parametric linear regression of average returning percentage on sent contingencies shows a positive slope with p-value of 0.0093.

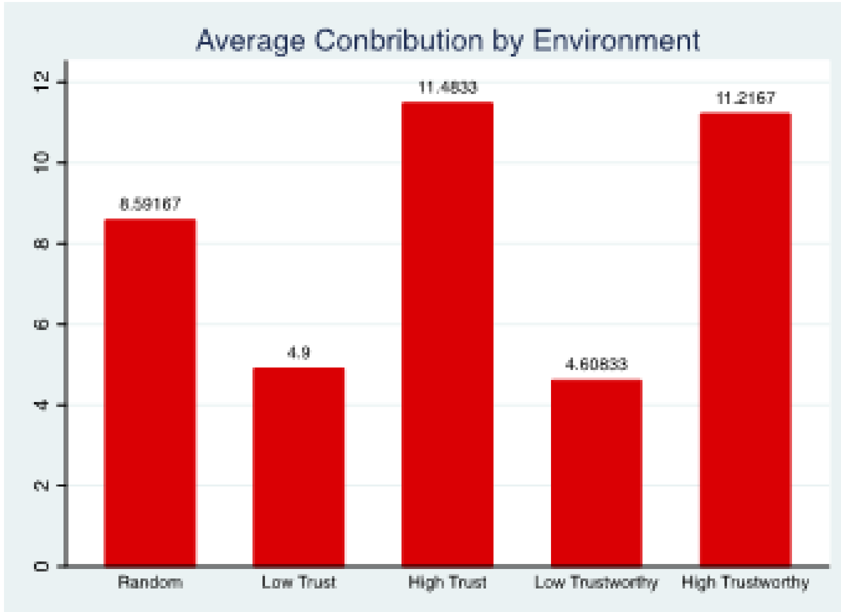


Figure 3.1: Contribution in the VCM in the second Phase

Moreover, unlike the cross-game design by Gachter and Thoni (2005) where group membership in the second phase is determined by the absolute ranking in the first phase VCM in order, group formation in the second phase of our experiment is based on the average rankings of the sending and returning behaviors in the trust game. Assuming our subjects fully understood the matching scheme, equivalently, this is to imply:

Result 2: Subjects' voluntary contribution decisions respond to average trust and average trustworthiness in their groups formed from the trust game.

owever, the contribution difference between high trust and high trustworthiness environments (0.26 tokens) is negligible and statistically insignificant, so is the difference between low trust and low trustworthiness environments (0.29 tokens). This seems to suggest that people do not distinguish trust and trustworthiness behavior of their group members when making decisions in this kind of social dilemma, and people to some extent consider trust and trustworthiness interchangeably. This is also corroborated by our survey results: out of the 80 subjects that were asked about their views of correlation between sending and returning behavior in the trust game, over 64% believed they are highly, positively correlated, suggesting one can infer the other. This makes us conclude:

Result 3: When making decisions in a social dilemma, subjects think of trust and trustworthiness of their group members interchangeably.

3.5.2 Beliefs in VCM

Figure 2 depicts the average first-order beliefs and average second-order beliefs in each environment, and the numbers track the contribution levels closely in all environments. In high trust and high trustworthiness environments, our participants hold significantly higher beliefs than in random environment, and the beliefs in low trust and low trustworthiness environments are significantly lower. This speaks to the fact that:

Result 4: In environments where average trusting and trustworthy levels differ, people's beliefs about others' behavior also differ drastically.

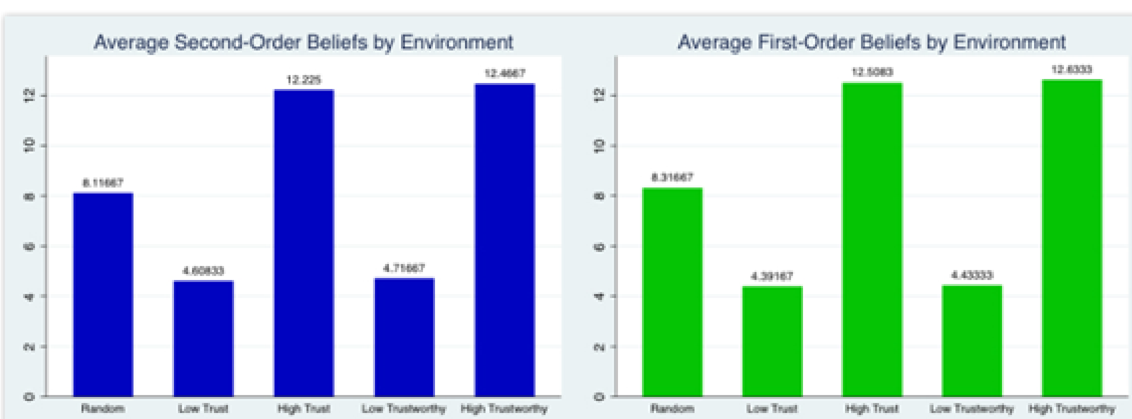


Figure 3.2: Average Beliefs in VCM in the second phase

As suggested by Dufwenburg et al. (2011), both first-order beliefs and second-order beliefs in the VCM are important determinants of contribution levels. They point out specifically that reciprocity (Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006) is the main channel through which the first-order beliefs impact contribution decisions, and guilt-aversion (Battigalli and Dufwenberg, 2007) operates independently via second-order beliefs to determine the contribution levels. However, we observe that in general the first order beliefs and the second order beliefs are highly correlated with each other in all environments, and the average correlation equals to 0.7668.

Such high, a positive correlation implies that it is inapt to regress contribution levels on first-order beliefs and on second-order beliefs separately, as Dufwenburg et al. (2011) did in their study. Instead, we try our first set of regressions by regressing contributions on both beliefs in each of the five environments, and not surprisingly, the impact of second order beliefs is absorbed: only those coefficients of the first order beliefs are statistically significant, with or without demographic controls such as race, gender, class level, and major. In this as well as in future regression tables, unless specified, none of the demographic control variables is significant so we exclude them in our report. Table 3.2 below shows the regression outcomes in each environment as well as in a pooled result. When one's belief about the average contributions of the other group members increases one token, his/her own contribution goes up accordingly by approximately one token.

Table 3.2: Contribution and Beliefs in VCM

Dependent Variable: Contributions	(1) Random	(2) Low Trust	(3) High Trust	(4) Low Trustworthy	(5) High Trustworthy	(6) Pooled
First-Order Beliefs	1.081***	1.027***	1.027***	1.156***	1.158***	0.982***
	(0.204)	(0.144)	(0.144)	(0.098)	(0.122)	(0.064)
Second-Order Beliefs	-0.0642	-0.000773	-0.000773	-0.0513	-0.117	-0.046
	(0.28)	(0.0906)	(0.0906)	(0.107)	(0.196)	(0.068)
Constant	0.119	0.394	-1.407	-0.274	-1.951	0.241
	(0.662)	(0.745)	(0.73)	(0.371)	(1.192)	(0.0354)
Observations	120	120	120	120	120	120
R-squared	0.522	0.477	0.448	0.591	0.529	0.579

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We further adopt a 2-step differencing approach by first regressing the first-order beliefs on the second-order beliefs while controlling for contributions, and we save the residuals as it captures the part of second-order beliefs that is orthogonal to the first-order beliefs. Next, in the second step we regress contribution on the first-order beliefs as well as the residuals. If the coefficient of the residuals is statistically significant, then we can conclude that the second-order beliefs separately impact contribution decisions in the VCM. Nonetheless, the results are quantitatively and qualitatively the same: the only significant coefficient is that of the first order beliefs in each environment. The second step regression outputs are shown in the Appendix. We thus report the regression results of contribution on first-order beliefs only in Table 3.3. Unlike Dufwenburg et al.'s (2011) findings, these results suggest that reciprocity, rather than guilt-aversion, plays the decisive role in explaining how beliefs impact contributions in the VCM. As a consequence, we conclude that:

Result 5: When their first-order beliefs are higher, people contribute more in the VCM game because of reciprocity.

Table 3.3: Contribution and First-Order Beliefs in VCM

Dependent Variable: Contributions	(1) Random	(2) Low Trust	(3) High Trust	(4) Low Trustworthy	(5) High Trustworthy
First-Order Beliefs	1.032***	1.026***	0.970***	1.117***	1.070***
	(0.0697)	(0.119)	(0.0575)	(0.086)	(0.0483)
Constant	0.0113	0.393	-0.644	-0.344	-2.306**
	(0.605)	(0.781)	(0.755)	(0.356)	(0.76)
Observations	120	120	120	120	120
R-squared	0.522	0.477	0.441	0.591	0.526

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Hereafter, we focus only on the first-order beliefs in our analyses. Next, we proceed one step further by demonstrating that the difference in first-order beliefs between high and low trust (trustworthiness) environments are indeed explanatory to the contribution difference in between high and low trust (trustworthiness) environments. We pool the data on environment (trust and trustworthiness) and estimate a fixed effect model by regressing the contribution difference on the belief difference between high and low environments in column 1. In column 2 we adopt an IV estimate using the estimated difference in high and low sending (returning) in the trust game as an instrument, since such estimates should be independent from contributions in the VCM game. The results are reported in Table 3.4, and the coefficients are significant.

Table 3.4: Difference in Belief drives the Difference in Contribution

Dependent Variable: Difference in Contributions	(1) Pooled	(2) IV
Difference in First-order Beliefs	0.657***	2.490*
	(0.069)	(0.458)
Constant	1.232*	-14.46
	(0.67)	(12.479)
Fixed Effect	YES	YES
Observations	240	80
R-squared	0.274	0.189

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Thus, we are able to conclude that:

Result 6: A high trusting or high trustworthy environment promotes cooperation through higher first-order beliefs.

3.5.3 Cross-Game Analysis

Trust Game Behavior and Beliefs in VCM

In regards to our hypotheses, we first check the relationship between sending behavior in the trust game in the first phase and beliefs in the second phase VCM. Table 3.5 displays the regression results when dependent variable is first order belief. We find that overall trusting behavior is weakly correlated with first order beliefs; however, in low trust (trustworthiness) environments, high senders tend to hold significantly higher beliefs in others' contribution to the collective account.

Table 3.5: Sending in the Trust game and First-Order Beliefs in VCM

Dependent Variable: First Order Beliefs in the VCM	(1) Random	(2) Low Trust	(3) High Trust	(4) Low Trustworthy	(5) High Trustworthy
Sending in the Trust Game	0.0433*	0.0753***	0.0545*	0.0651***	0.0683*
	(0.0212)	(0.0174)	(0.0282)	(0.0107)	(0.0374)
Constant	7.468***	2.917***	11.44***	3.158***	11.30***
	(0.736)	(0.349)	(0.783)	(0.315)	(0.979)
Observations	120	120	120	120	120
R-squared	0.019	0.089	0.029	0.069	0.048

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Not unexpectedly, average trustworthiness in the trust game and first order beliefs in the VCM are significantly correlated with each other in all environments, as shown in the table below. From discussion above, both returning behavior in the trust game and first order beliefs in the VCM are driven by reciprocity, which is a stable prosocial preference over time (Carlsson et al., 2014). Consequently, average trustworthiness is predictive to first order beliefs in the VCM.

Table 3.6: Average Returning Percentage in the Trust game and First-Order Beliefs in VCM

Dependent Variable: First Order Beliefs in the VCM	(1) Random	(2) Low Trust	(3) High Trust	(4) Low Trustworthy	(5) High Trustworthy
Average Returning %	3.207***	4.809***	5.056**	4.037***	5.749***
	(0.696)	(0.868)	(1.433)	(0.619)	(1.153)
Constant	7.316***	2.891***	10.93***	3.174***	10.84***
	(0.373)	(0.316)	(0.587)	(0.367)	(0.623)
Observations	120	120	120	120	120
R-squared	0.026	0.092	0.062	0.067	0.086

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Conditional Cooperation in VCM and Trustworthiness in Trust Game

Andreoni (1995) argues that about half of all cooperation in the VCM comes from subjects who choose not to free-ride out of some form of kindness. Fehr and Gachter (2000) argue that conditionally cooperative subjects reciprocate the kind contributions of other cooperators; that is, they raise their contributions when others' contributions are high in the group (Fischbacher et al., 2001) or their beliefs in others' cooperativeness is higher

(Gachter, 2006). Therefore, the higher contribution levels in high trust and high trustworthiness environments could be largely driven by the presence of conditional cooperators, as their first-order beliefs are much higher in these two environments. We are thus especially interested in learning how trust game behavior is related to VCM types. As discussed in the introduction, we hypothesize that trustworthiness and conditional cooperation are closely related, as the social preferences that motivate these behaviors largely coincide (such as reciprocity and inequality aversion).

Given our experimental design, we are unable to condition our subjects' contributions on each of the possible contribution levels of the other group members as in Fischbacher, Gachter, and Fehr (2001)'s study. Instead, we condition their contributions on their first-order beliefs, and define it in our context in Gachter's (2006) sense: a person is conditional cooperator if his/her contribution increases with his/her first-order beliefs. We are able to classify our subjects in this way, given contributions and beliefs vary greatly across environments. For each one of the five environments, we obtain one contribution-belief pair and graph each subject's five contributions against his/her corresponding beliefs. We first identify all subjects whose graph has a monotonically increasing trend, and then for this group of subjects we regress their contributions on their first-order beliefs to see if the regression coefficients are statistically significant. If the coefficient of beliefs is also significant, we label this person as a "conditional cooperator". In addition, we label all subjects whose contributions as well as beliefs are zeroes in all environments as "perfectly selfish". Based on this classification, 57 out of the 120 subjects we have are conditional cooperators, 18 are perfectly selfish, and the other 46 belong to neither and are thus grouped as a third type, "Others".

We report in Table 3.7 a set of probit regressions where the dependent indicator variable is 1 if this person is a conditional cooperator. We combine the "perfectly selfish" and "Others" types in regression (2) and (4) so they are on the full sample, and we leave out the "Others" type is regression (1), (3), and (5) so they focus on the comparison of conditional cooperators vs perfect selfish types, on a total of 63 subjects. Since the dependent variable is based on observations from all environments, the result should be interpreted in a general sense.

Table 3.7: How Trust Game Behavior Predicts VCM Type

Dependent Variable: VCM Type=1 if Conditional Cooperator	(1)	(2)	(3)	(4)	(5)
Sending in the Trust Game	0.0361**	0.00973*			0.0189
	(0.0174)	(0.00519)			(0.0177)
Returning Amount when being sent 50 tokens			0.0181***	0.00384	0.0146**
			(0.00403)	(0.00211)	(0.0055)
Constant	0.243	-0.233**	0.208	-0.23	0.0332
	(0.268)	(0.118)	(0.309)	(0.216)	(0.297)
Demographic Controls	YES	YES	YES	YES	YES
Observations	63	120	63	120	63

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As can be seen, both sending and returning behavior in the trust game can predict the VCM type. When considered separately, being more trusting and being more trustworthy both indicate high likelihood of being conditional cooperators in the VCM game; However, when sending and returning decisions are considered simultaneously, only returning behavior seems to be relevant: the higher a subject reciprocate the sender when being sent all 50 tokens, the more likely he/she is a conditional cooperator in the VCM game. This seems to corroborate that trusting behavior doesn't require a social preference, whereas trustworthiness does; hence trustworthiness and conditional cooperativeness is more closely correlated than trust with conditional cooperativeness.

Result 7: High trustworthiness is a predictor of conditional cooperation in VCM.

3.6 Conclusion and Discussions

From our experimental results, we offer a new insight into why high trusting environments are associated with high faster economic. When people's perceptions about the overall trust and trustworthiness of society change positively, they become more willing to exert efforts to contribute to local public goods and to report thefts or corruptible officers, which reduces social inefficiency and creates a healthy environment for economic takeoff.

It is worth pointing out that the incentivizing device that we use to elicit people's beliefs in the VCM (to reward subjects with 5 additional points if their guesses are within one token of the true number) might be subject to questioning. The current reward scheme to elicit stated beliefs would allow subjects to hedge by providing a guess as close to the theoretically predicted average as possible (Blanco et al., 2010), even though we try our best to control for the timing of elicitation to avoid contamination. A better procedure would be to adopt some kind of scoring rules (see the survey literature by Schotter and Trevino, 2014) to make it a dominant strategy to reveal beliefs truthfully. Nevertheless, in our experiment our subjects have to provide their guesses in high trust/trustworthiness and low trust/trustworthiness environments that are endogenous in the sense that they are determined by this particular group of people in each session, so it is nearly impossible for them to strategically manipulate their beliefs other than simply providing their true beliefs based on their living experience and the perceptions of the surrounding world. We believe that improving the belief elicitation methods will not impact the qualitative implication of our work.

Lastly, we only focus on the "positive reciprocity" aspect of the conditional cooperation in the VCM game, where the other half, that is, how the inclination of conditional cooperators to punish free-riders (Hoffman et al., 1998; Fehr and Gachter, 2000) is associated with trustworthiness, is still unexplored. Future studies could allow punishment opportunity in the VCM to investigate this. Another direction to extend this project is to vary the group size or impose different MPCR on different groups in the VCM to see how trust/trustworthiness differentially impact the contribution decisions.

3.7 Appendix

3.7.1 Experimental Instructions

Welcome

Thank you for participating in our decision-making experiment. This experiment involves a set of decision interactions among participants in two phases. Depending on your decisions and the decisions of other participants, you will be able to earn money in addition to the $5 guaranteed for your participation.

The experimental environment in the first phase will be explained shortly and the instructions for the second phase will be given after the end of the first phase. Please be assured that all other participants are actual participants in this room and any information you are given about others' decisions represents actual decisions of these participants.

No communication among participants is permitted during the experiment. Thus, you are not allowed to use your phone, tablet computer, or programs other than the designated experiment software to communicate with others, and no talking or passing of notes is permitted. If you have a question at any time, please raise your hand so that one of us can come to you to provide you with the clarification that you need.

During the experiment, we will be using a currency or unit of account we'll call **tokens**. You are initially endowed with some amount of tokens, and you can allocate these tokens as you wish in each phase. The decisions you and others make will possibly earn you tokens that will be converted to real money and paid to you when the experiment ends. The conversion rates of tokens to dollars will be specified in each phase.

I. Instruction for Phase I

I.1 General Description

The interaction in this phase involves two roles that we refer to as the ?first mover? and the ?second mover.? The participants in each role are both endowed with 50 tokens at the beginning of the interaction. The first mover chooses how many token to send to the second mover, in 5 token increments, i.e. the first mover can send 0, 5,. . . , or 50 tokens. **Denote this first mover decision as X**. Any tokens that are sent to the second mover by the first will be **tripled**. Upon receiving the tripled number of tokens, the second mover chooses how many tokens in integer amounts, from 0 to 3X, to return to the first mover. **Denote this second mover decision as Y**.

Based on these decisions, the earnings of the first mover will be the initial 50 tokens minus the tokens (if any) he or she sent to the second mover plus the tokens (if any) that are returned by the second mover. The earnings of the second mover will be the initial 50 tokens plus three times the tokens (if any) that are sent from the first mover minus the tokens (if any) that are sent to the first mover. Namely, the payoff functions for the first mover and the second mover are, $\pi_1 = 50 - X + Y$ and $\pi_2 = 50 + 3X - Y$, respectively.

Example:

If the first mover, A, sent 25 tokens out of his endowment of 50, then the second mover, B, would receive 25x3=75 tokens in addition to her endowment of 50. Now suppose B decides to return 15 tokens, then A will end up with 50-25+15=40 tokens, and B with 50+75-15=110 tokens.

Remember that the second mover is free to return 0; likewise, the first mover is free to send 0.

Each of you will be making decisions first as a first mover, and then as a second mover. Only one of these roles can be selected for payment, and for whichever role that is, the computer will randomly select a participant to pair you with, so you will never play the second mover part against yourself as first mover. You will not learn the identity of the individual you are paired with, nor will they learn yours. Calculation of earnings will be explained later.//

I.2 Decision as the "First Mover"

For this decision, you are the first mover. You can choose one of the 11 possible levels to send: 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50. Remember that any tokens you keep will be part of your own earnings if your choice in this role is selected to be paid off on, and every token you choose to send to the second mover is multiplied by 3.

You are allowed to experiment first by choosing one out of the 11 circles (corresponding to the 11 levels) on the screen (see below) and clicking on the "Calculate" button to see the selected numbers of tokens and see how many tokens the second mover would receive under this decision. Once you are satisfied with your choice, click "Submit" to confirm your decision.

Your decision is then final.

I.3 Decision as the "Second Mover"

For this decision, you are the second mover. Please decide on how many (if any) tokens you choose to return to the first mover under all possible contingencies (see the screen below), as you will NOT be informed of the first mover's actual choice until payoffs are reported to you later.

That is, for each of the 10 relevant sent and tripled amounts (remember that the first mover can only send multiples of 5, up to 50, and if he sent 0, there is nothing you could return), please indicate how many tokens you choose to send back to the first mover, where you can send any integer amount between 0 and the amount you received. Please type your decision conditional on each received level in each of the ten boxes. You can play around with your answers and click "Calculate" to see your earnings and the first mover's earnings resulting from your actions in each case. Once you are satisfied with your choice, click "Submit" to confirm your decision. **Your decisions are then final.**

I.4 Payoff Calculation in this Phase

Only one of the two decisions you made will be randomly selected to determine your earnings in this phase. If your first mover decision is selected for payment, your decision and the second mover decision of the participant randomly paired with you will determine your and his/her earnings in this phase. Likewise, if your second mover decision is selected for payment, your decision and that of the first mover randomly paired with you will determine your and his/her earnings in this phase. Your earnings this phase in tokens will convert to real money at the rate of $0.10 per token, which will be paid to you at the end of the experiment. (Note that your counterpart's decision and your resulting earnings in this phase will not be reported to you until after Phase 2.)

Example:

If Mr. A's first mover decision is chosen for his payment, and he got 60 tokens as the first mover in the anonymous pairing, then he will receive 60x$0.10 = $6 at the end of the experiment as his earnings for this phase.

This is the end of the Instructions for Phase I.

Please raise your hand if you have any questions.

II. Instruction for Phase II

II. Brief Introduction

In this phase, you are asked to make 3 sets of decisions. Each set needs your decisions in five different decision environments, where you have different information about the other participants matched with you. The first set of decisions are about allocating tokens, while the second and third sets are about estimates of what others will do and guesses about others' estimates. Group formation within each environment is random (although in some environments subject to relevant constraints), and no information about the other participants' identities or decisions regarding allocation and estimation will be revealed to you or others throughout the phase. At the end of the experiment, you will only be notified of others' allocation decisions for the purpose of payment, and only **one** environment will be chosen for payment. Payoff calculations in this phase will be explained later.

II.1 The Five Decision Environments

Within each set of decision, there are five different environments that differ in terms of the (true) information about the group composition.

In **Environment 1**: the other four group members are chosen randomly from among all participants.

The differences across the remaining four environments are based on differences in participants' choices in Phase 1 (the first mover/second mover decision).

To understand environments 2 and 3, suppose that all phase 1 first-mover sending decisions in this room are ranked from the lowest 1st to the highest 20th to form an ordered list. Based on such a list,

In **Environment 2**: the average level of the other four members' sending corresponds to a low rank (below rank number 8).

In **Environment 3**: the average level of the other four members' sending corresponds to a high rank (above rank number 12).

Put more intuitively (but a bit less precisely), in Environment 2 you are grouped with others who on average sent relatively small amounts as the first movers, while in Environment 3 you are grouped with others who on average sent relatively large amounts as the first movers.

To understand environments 4 and 5, suppose that all Phase 1 second-mover returning decisions, when being sent all 50 tokens by the first mover, are ranked from the lowest 1st to the highest 20th to form an ordered list. Based on such a list,

In **Environment 4**: the average level of the other four members' returning corresponds to a low rank (below rank number 8).

In **Environment 5**: the average level of the other four members' returning corresponds to a high rank (above rank number 12).

Put more intuitively (but a bit less precisely), in Environment 4 you are grouped with others who on average returned relatively small proportions as the second movers, while in Environment 5 you are grouped with others who on average returned relatively large proportions as the second movers.

<u>II.2 The Three sets of Decisions</u>

Decision 1:

In this decision set, you will be a member of a group consisting of **5 people**, yourself included, in all five different environments. As explained in the previous section, the other members of your group differ in each environment. Note that this first decision is the most payoff-relevant stage in this phase; that is, it will probably determine the largest part of your payment for the phase.

In all environments, each of you is endowed with 20 tokens at the beginning and each simultaneously makes individual decisions on how to allocate these tokens, in integer amount, between a group account and a private account. Any tokens you choose not to allocate to the group account will be automatically allocated to your private account. Everyone benefits equally from the tokens in the group account: each of you gets 0.4 tokens towards your private account per token in the group account. That is, your earnings are the number of tokens in your private account plus 0.4*the total tokens in the group account.

Example:

	CASE 1	CASE 2	CASE 3	CASE 4
a. Your Contribution	0	20	10	5
b. Other's contribution - 1	0	20	10	6
c. Other's contribution - 2	0	20	10	7
d. Other's contribution - 3	0	20	10	13
e. Other's contribution - 4	0	20	10	9
f. Tokens in your private account (= 20 – a.)	20	0	10	15
g. Tokens in the group account (= a+b+c+d+e)	0	100	50	40
h. Earnings from group account for each person (=0.4*g)	0	40	20	16
Your Total Earnings (in tokens) (= f + h)	20	40	30	31

Decision 2:

In this decision set, you will be asked to estimate, **on average**, how many tokens the other four group members have allocated to the group account in Decision 1 in each of the five environments. 5 additional tokens will be given to you if your estimate is within one token of the true average in the payoff relevant environment. No tokens will be taken away from you if your estimate is incorrect/ imprecise.

Example:

1) Suppose in Environment 1, the actual average allocation of the other four people in your Environment 1 group is 5. If this environment is selected for payment, then if your estimate of the average is between 4 and 6 tokens, you will get 5 additional tokens; otherwise you will NOT get any additional tokens

2) Suppose your guess for Environment 1 qualifies you for the 5 additional tokens, while your estimate for Environment 2 does not. If Environment 2 is chosen for payment, then you will NOT get any additional tokens for this decision.

Decision 3:

In this decision set, you will be asked to estimate the **average estimate** provided by the other four group members in Decision 2. That is, you will guess the average of the estimates that each of the other four members has given regarding the average allocation to the group account, by you and the rest, in the previous decision. Similar to Decision 2, you will receive an additional 5 tokens if your estimate comes within one token of the true average of their estimates in the payoff relevant environment. No tokens will be taken away from you if your estimate is incorrect/ imprecise.

II.3 Payoff Calculation for this Phase & the entire Experiment

For Phase 2, only one out of the five decision environments will be randomly chosen to determine your earnings for both your allocation decision (Decision 1) and your estimates (Decisions 2 and 3). For the selected environment, your earnings in tokens will be based on Decision 1 (i.e., what you put in your private account plus the earnings from the public account based on what you and the other four in that group put in the public account), plus any additional earnings from estimates in Decision 2 and Decision 3. Your token earnings will be converted to real money at a rate of **$0.20 per token**.

Example:

Suppose Environment 1 is chosen for payment, and your estimate in Decision 2 is in the correct range but that in Decision 3 is not; moreover, you contributed 10 tokens in Decision 1 and the public account ends up with 45 tokens. In the end, you would get: {(20-10) + 0.4*45} + 5 + 0 = 33 tokens, which is 33*$0.20 = $6.60 in real dollars, for this phase.

In addition to knowing the results in phase 2, you will also be given information on results in Phase 1 (refer back to page 4 of the first set of instructions). Your final earnings from the experimental decisions and outcomes are then the sum of your earnings in Phase 1 and Phase 2. Your full payment will be this sum plus the $5 show up fee.

This is the end of the instructions for Phase 2.

Please raise your hand if you have any questions.

BIBLIOGRAPHY

[1] Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990). "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica*: 1041-1063.

[2] Algan, Yann, and Pierre Cahuc (2010). "Inherited trust and growth." *American Economic Review*, 100(5): 2060-2092.

[3] Andreoni, James (1995). "Cooperation in public-goods experiments: kindness or confusion?" *American Economic Review*, 891-904.

[4] Andreoni, James, and John Miller (2002). "Giving according to GARP: An experimental test of the consistency of preferences for altruism." *Econometrica*, 70(2): 737-753.

[5] Andreoni, James, and Charles Sprenger (2012). "Estimating time preferences from convex budgets." *American Economic Review*, 102(7): 3333-3356.

[6] Andreoni, James, and Charles Sprenger (2012b). "Risk preferences are not time preferences." *American Economic Review*, 102(7): 3357-3376.

[7] Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström (2008). "Eliciting risk and time preferences." *Econometrica*, 76(3): 583-618.

[8] Ashraf, Nava, Iris Bohnet, and Nikita Piankov (2006). "Decomposing trust and trustworthiness." *Experimental Economics*, 9(3): 193-208.

[9] Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy (2016). "Incentives in experiments: A theoretical analysis." *Working paper*.

[10] Battigalli, Pierpaolo, and Martin Dufwenberg (2007). "Guilt in games." *American Economic Review*, 97(2): 170-176.

[11] Blanco, Mariana, Dirk Engelmann, and Hans Theo Normann (2011). "A within-subject analysis of other-regarding preferences." *Games and Economic Behavior*, 72(2): 321-338.

[12] Bellemare, Charles, and Sabine Kroger (2007). "On representative social capital." *European Economic Review*, 51(1): 183-202.

[13] Berg, Joyce, John Dickhaut, and Kevin McCabe (1995). "Trust, reciprocity, and social history." *Games and Economic Behavior*, 10(1): 122-142.

[14] Bolton, Gary E., and Axel Ockenfels (2000). "ERC: A theory of equity, reciprocity, and competition." *American Economic Review*, 166-193.

[15] Chade, Hector, Pavlo Prokopovych, and Lones Smith (2008). "Repeated games with present-biased preferences." *Journal of Economic Theory*, 139(1): 157-175.

[16] Charness, Gary, and Matthew Rabin (2002). "Understanding social preferences with simple tests." *Quarterly Journal of Economics*: 817-869.

[17] Chaudhuri, Ananish (2011). "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature." *Experimental Economics*, 14(1): 47-83.

[18] Cox, James C (2004). "How to identify trust and reciprocity." *Games and Economic Behavior*, 46(2): 260-281.

[19] Dal Bó, Pedro (2005). "Cooperation under the shadow of the future: experimental evidence from infinitely repeated games." *American Economic Review*, 95(5): 1591-1604.

[20] Dal Bó, Pedro and Guillaume R. Fréchette (2011). "The evolution of cooperation in infinitely repeated games: Experimental evidence." *American Economic Review*, 101(1): 411-429.

[21] Dal Bó, Pedro and Guillaume R. Fréchette (2015). "Strategy choice in the infinitely repeated prisoners' dilemma." *Available at SSRN 2292390.*

[22] Dal Bó, Pedro and Guillaume R. Fréchette (Forthcoming). "On the determinants of cooperation in infinitely repeated games: A survey." *Journal of Economic Literature.*

[23] Davis, Douglas, Asen Ivanov, and Oleg Korenok (2016). "Individual characteristics and behavior in repeated games: an experimental study." *Experimental Economics*, 19(1): 67-99.

[24] Dean, Mark, and Anja Sautmann (2014). "Credit constraints and the measurement of time preferences." *Available at SSRN 2423951.*

[25] Dreber, Anna, Drew Fudenberg, and David G. Rand (2014). "Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics." *Journal of Economic Behavior and Organization*, 98: 41-55.

[26] Dufwenberg, Martin, and Georg Kirchsteiger (2004). "A theory of sequential reciprocity." *Games and Economic Behavior*, 47(2): 268-298.

[27] Dufwenberg, Martin, Simon Gachter, and Heike Hennig-Schmidt (2011). "The framing of games and the psychology of play." *Games and Economic Behavior*, 73(2): 459-478.

[28] Falk, Armin, and Urs Fischbacher (2006). "A theory of reciprocity." *Games and Economic Behavior*, 54(2): 293-315.

[29] Fehr, Ernst, and Klaus M. Schmidt (1999). "A theory of fairness, competition, and cooperation." *Quarterly Journal of Economics*: 817-868.

[30] Fischbacher, Urs (2007). "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10(2): 171-178.

[31] Fréchette, Guillaume R., and Sevgi Yuksel (Forthcoming). ”Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination.” *Experimental Economics*.

[32] Frederick, Shane, George Loewenstein, and Ted O’donoghue (2002). “Time discounting and time preference: A critical review.” *Journal of Economic Literature*, 40(2): 351-401.

[33] Freeman, David, Paola Manzini, Marco Mariotti, and Luigi Mittone (2016). “Procedures for eliciting time preferences.” *Journal of Economic Behavior and Organization*, 126: 235-242.

[34] Fudenberg, Drew, and David K. Levine (2006). “A dual-self model of impulse control.” *American Economic Review*, 95(5): 1449-1476.

[35] Fudenberg, Drew, and Eric Maskin (1986). “The folk theorem in repeated games with discounting or with incomplete information.” *Econometrica*: 533-554.

[36] Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2009). “Cultural biases in economic exchange?” *Quarterly Journal of Economics*, 124: 1095âĂŞ1131.

[37] Halevy, Yoram (2015). “Time consistency: Stationarity and time invariance.” *Econometrica*, 83(1): 335-352.

[38] Harrison, Glenn W., Morten I. Lau, and E. Elisabet Rutström (2010). “Individual discount rates and smoking: Evidence from a field experiment in Denmark.” *Journal of Health Economics*, 29(5): 708-717.

[39] Horton, John J., David G. Rand, and Richard J. Zeckhauser (2011). “The online laboratory: Conducting experiments in a real labor market.” *Experimental Economics*, 14(3): 399-425.

[40] Knack, Stephen, and Philip Keefer (1997). "Does social capital have an economic payoff? A cross-country investigation." *Quarterly Journal of Economics*, 1251-1288.

[41] Laibson, David (1997). "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics*, 443-477.

[42] Lehrer, Ehud, and Ady Pauzner (1999). "Repeated games with differential time preferences." *Econometrica*, 67(2): 393-412.

[43] Meier, Stephan, and Charles Sprenger (2010). "Present-biased preferences and credit card borrowing." *American Economic Journal: Applied Economics*, 2(1): 193-210.

[44] Proto, Eugenio, Aldo Rustichini, and Andis Sofianos (2014). "Higher intelligence groups have higher cooperation rates in the repeated prisoner's dilemma." *Working paper.*

[45] Rabin, Matthew (1993). "Incorporating fairness into game theory and economics." *American Economic Review*, 1281-1302.

[46] Roth, Alvin E., and J. Keith Murnighan (1978). "Equilibrium behavior and repeated play of the prisoner's dilemma." *Journal of Mathematical Psychology*, 17(2): 189-198.

[47] Sabater-Grande, Gerardo, and Nikolaos Georgantzis (2002). "Accounting for risk aversion in repeated prisoners' dilemma games: An experimental test." *Journal of Economic Behavior and Organization*, 48(1): 37-50.

[48] Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler, and Stefan T. Trautmann (2013). "Impatience and uncertainty: Experimental decisions predict adolescents' field behavior." *American Economic Review*, 103(1): 510-531.

[49] Thoni, Christian, Jean-Robert Tyran, and Erik Wengstrom (2012). "Microfoundations of social capital." *Journal of Public Economics*, 96(7): 635-643.

[50] Urminsky, Oleg, and Gal Zauberman (2016). "The psychology of intertemporal preferences." *Working paper*.

[51] Vespa, Emanuel (2015). "An experimental investigation of strategies in dynamic games." *Working paper*.