Cooperation in Stochastic Games: A Prisoner's Dilemma Experiment

Andrew Kloosterman*

University of Virginia
PO Box 400182
Charlottesville, VA 22901

Keywords: Stochastic games, Experimental Economics, Prisoner's Dilemma, Cooperation
JEL Codes: C73, C92

August 29th, 2017

Abstract

This experiment investigates cooperation in a stochastic version of the infinitely repeated prisoner's dilemma. The stochastic element introduces the importance of beliefs about the future for supporting cooperation as well as cooperation and defection on the equilibrium path. The results confirm that subjects cooperate as predicted after they gain sufficient experience. There is some evidence in favor of alternating cooperation and defection, but a maximum likelihood strategy estimation suggests that cooperation is conditioned mostly on past actions. For example, the popular repeated game strategies Grim Trigger and Tit-for-Tat are still popular here, although they are not equilibria in this environment.

1 Introduction

One of the most celebrated results of the dynamic games literature is that repeated interaction provides a mechanism for agents to cooperate in the infinitely repeated prisoner's dilemma. Infinitely repeated games, while a useful abstraction in many respects, are a drastic simplification of the real world. Rather, dynamic environments

*I wish to thank Andrew Schotter, Guillaume Fréchette, Emmanuel Vespa, and Jack Fanning for their help on this project, as well as participants at the North American ESA Conference 2012, CESS-Amsterdam Graduate Student Conference 2014, the SEA Conference 2015, and the UVA Experimental Social Science Conference 2016 for insightful comments. Also, thanks to NSF Doctoral Dissertation Research grant SES-1260840 for financial support.

usually evolve over time. This feature is captured by stochastic games, a generalization of repeated games in which the game played each period changes and is determined by an underlying stochastic process. This paper considers a stochastic version of the infinitely repeated prisoner's dilemma where a prisoner's dilemma is played each period, but the payoffs of the dilemma can change over time.

Adding a stochastic element is not just more realistic, but also important to test empirically as the environment becomes more complex in a specific way that potentially makes supporting cooperation more difficult. More precisely, the infinitely repeated prisoner's dilemma is simple in the sense that the path of the most cooperative equilibrium is constant, there is either cooperation in every period (with a threat of punishing defections that never get realized) or defection in every period, whereas the stochastic version considered here is more complex in that the path of the most cooperative equilibrium prescribes cooperation in some periods and defection in others.

This complexity presents two issues for supporting cooperation that are the focus of this paper. The first is that beliefs about which dilemmas will be played in the future determine the scope for punishing unilateral defections, and therefore are an important determinant of when cooperation can be supported. Beliefs about future dilemmas play a complex role in supporting cooperation when future play is not constant, as players must both calculate an expected value over the different payoffs that are possible in future periods, and determine which action, cooperate or defect, will be taken in each of these future periods. This, of course, requires a great deal of forward reasoning that may break down and lead to the simple, and constant, equilibrium strategy of always defecting.

Second, there are many equilibrium (and other reasonable) strategies in this environment and coordination becomes more difficult as the players must not only coordinate on cooperation or defection in period 0, but coordinate on cooperation or defection in each period. For example, as just noted, always defecting is an equilibrium strategy so after defection it may be difficult to determine if one's opponent is playing the strategy always defect or playing a strategy that will switch to cooperation once beliefs about the future permit it. Another way of saying this is that conditionally cooperative equilibrium strategies do not only condition on past actions (i.e. past cooperators are rewarded with mutual cooperation and past defectors are punished with mutual defection) as in the infinitely repeated prisoner's dilemma, but also condition on the outcomes of the stochastic process.

The game played in the experiment is the following. There are two prisoner's dilemmas, call them A and B for convenience, where, for reasons to be noted momentarily, the gain from unilateral defection against mutual cooperation is the same. One of the dilemmas is played in period 0 (which one is the first treatment variable). In period 1, A is played with probability p and B with probability $1 - p$. Then, the game realized in period 1 is infinitely repeated thereafter.[1] The important point is that cooperation can only be supported in periods greater than or equal to 1 if B is realized in period 1. This means punishment for defection in period 0 can only be executed when B is realized, as there must be defection when A is realized regardless of the outcome in period 0. Therefore, the importance of beliefs about the future is that the scope for punishment increases as p decreases. Furthermore, the gain is the same for A and B, so it does not matter whether A or B is played in period 0.

In other words, because the gains from defecting are the same, the experimental design isolates the belief parameter p as the only factor that should theoretically matter for cooperation in period 0. The second treatment variable is this probability p which is set at the values $p = .25$ or $p = .75$. As the discussion above suggests, cooperation in period 0 is theoretically possible when $p = .25$ where B is more likely in period 1.

Furthermore, defection is predicted in period 0 when $p = .75$, but cooperation is possible in period 1 if the realized period 1 game is B. Similarly, cooperation is predicted in period 0 when $p = .25$, but players must switch to defection in period 1 if the realized period 1 game is A. That is, in theory, there can be cooperation in some periods and defection in others in equilibrium. While these results are developed in the context of equilibrium, the theory section of this paper also shows that they are robust to risk-dominance based arguments that have been shown to describe behavior in the infinitely repeated prisoner's dilemma (Blonski et al. 2011, Dal Bó and Fréchette 2011).

The main result for period 0 behavior confirms the theoretical effect of beliefs about the future on cooperation in period 0, although only after subjects gain experience with the game. That is, there is a similar amount of period 0 cooperation in the two treatments where $p = .25$ and a similar amount in the two treatments where $p = .75$ and the former is greater than the latter. Interestingly, before they become experienced, subjects first focus on the period 0 game and not the probability p, the exact opposite of the theoretical prediction. Hence the main finding is that the

[1] In a stochastic game, the realization of each game is at the beginning of each period. The only uncertainty here is in period 0 where the players do not know if they will play A or B in all periods greater than 0.

setting is indeed complicated, but learning trends behavior towards the theoretical predictions.

The results show weaker evidence for shifts from cooperation to defection and from defection to cooperation that are predicted in the most cooperative equilibrium. Cooperation in period 1 does show a small shift from defection to cooperation when the theory predicts it should and vice versa. However, the differences are much smaller than the differences between when there was cooperation in period 0 and when there was defection in period 0. In other words, the evidence is much stronger for conditioning cooperation on previous actions than on the outcome of the stochastic process. The final section of the results estimates entire strategies using the method developed in Dal Bó and Fréchette 2011. The main finding from this estimation is that, to the extent that subjects cooperate at all, they use the same conditionally cooperative strategies found in the infinitely repeated prisoner's dilemma confirming that cooperation is more likely to condition on past actions than the realization of the stochastic process. It is important to stress that these strategies are not equilibrium strategies in this environment, as it is shown that every cooperative equilibrium strategy must condition on both. This finding is likely due to some amount of social preferences that the subjects exhibit, and is important to highlight as no such considerations are made in the infinitely repeated prisoner's dilemma.

As the beginning of the introduction claimed that the stochastic element adds realism to the problem, an example is warranted. Perhaps, because the prisoner's dilemma is often thought of as a reduced-form for collusion, the best example is collusion over the business cycle (Rotemberg and Saloner 1986, Haltiwanger and Harrington Jr. 1991, Kandori 1991, Bagwell and Staiger 1997). These studies explore under what conditions firms can collude when demand changes through time. The relevant conclusion from this line of research in regards to the effect of beliefs is that collusion is easier to support when future demand is expected to be large as defecting firms would forego potentially large collusive profits. Another motivation for the experiment is therefore to test this hypothesis.[2]

The closest-related experiment is Rojas 2012. Rojas considers a stochastic version of the prisoner's dilemma where the gains from defection vary each period but beliefs about the future are held constant (the opposite of the present setup). He does find some evidence for subjects cooperating and defecting over the course of the game, although perhaps it is easier in that environment, because the consequences

[2] Of course, in the experiment the game repeats from period 1 on whereas business cycles usually evolve every period. The theory section below addresses why the simplifying assumption is a good one for the experiment.

are immediate rather than through forward looking behavior. Furthermore, he does not address the issue of varying beliefs about the future that is the main treatment of the present experiment. Wilson and Vespa 2016 also considers a stochastic version of the prisoner's dilemma but they only consider one treatment with an exogenous stochastic process, and cooperation rates are low in this treatment.[3]

A few other stochastic games have been investigated in the lab (Charness and Genicot 2007, Ruffle 2013, Cabral, Ozbay, and Schotter 2014, Roy 2014, Vespa 2015, Saijo et al 2015, Kloosterman 2016). Cooperation in infinitely repeated games, mostly prisoner's dilemmas, has been studied extensively in the laboratory. Early experiments are inconclusive in regards to how subjects view the future and use it to support cooperation (i.e. Roth and Murnighan 1978, Murnighan and Roth 1983, Holt 1985, Feinberg and Husted 1993, Palfrey and Rosenthal 1994). More recent experiments have suggested that strategic cooperation does develop when cooperation is sufficiently attractive and subjects play many periods to gain experience (i.e. Engle-Warnick and Slonim 2004, 2006a, 2006b, Dal Bó 2005, Dreber et al. 2008, Duffy and Ochs 2009, Blonski et al. 2011, Dal Bó and Fréchette 2011). The main implication of recent work is that, as noted briefly above, behavior is best described by risk-dominance based arguments. This finding, and it's relevance here, is discussed thoroughly below.

Imperfect monitoring also induces equilibria with cooperation and defection. While there has been extensive experimental work on this topic as well (Holcomb and Nelson 1997, Feinberg and Snyder 2002, Aoyagi and Fréchette 2009, Fudenberg et al. 2012, Rojas 2012), it is difficult to compare results because the non-constant outcomes are driven by imperfect signals rather than the stochastic process.

The rest of this paper is organized as follows. Section 2 presents the stochastic prisoner's dilemma. Section 3 presents the experimental design. Section 4 presents results and section 5 concludes.

2 Theory

2.1 The Stochastic Prisoner's Dilemma

This paper investigates behavior in the following stochastic version of the infinitely repeated prisoner's dilemma. There are two players and two prisoner's dilemmas, call them A and B, with game matrices

[3]The focuses of these papers are not about the questions investigated here. The main results of Rojas 2012 consider imperfect monitoring. The main results of Wilson and Vespa 2016 consider endogenous processes where the game is determined also by the actions of the players in the previous period.

	Dilemma A	
	C	D
C	c_A, c_A	b_A, a_A
D	a_A, b_A	d_A, d_A

	Dilemma B	
	C	D
C	c_B, c_B	b_B, a_B
D	a_B, b_B	d_B, d_B

The strategies are cooperate (C) and defect (D). The games are prisoner's dilemmas which means that the payoffs satisfy $a_i > c_i > d_i > b_i$ for $i = A, B$. Additionally, it is assumed that $a_i + b_i < 2c_i$ so (C, C) is the socially efficient outcome. To simplify expressions below, the gain in game i from unilateral defection is denoted $g_i \equiv a_i - c_i$, the loss from mutual defection is denoted $l_i \equiv c_i - d_i$, and the loss from getting the worst payoff relative to defecting is denoted $h_i \equiv d_i - b_i$. An exogenous stochastic process Γ determines which of the two games is played each period of the infinite horizon $t = 0, 1, 2, \ldots$. Finally, the players discount future payoffs at the common rate δ.

One final point worth stressing is that players observe the game they play in each period before they play it. Unlike a repeated game of incomplete information, players know the game they play each period and are just uncertain about what games they will play in future periods.

The experiment focuses on investigating games that satisfy the following parameter cases (for reasons to be explained momentarily). First, the gain from unilateral defection and loss from the worst payoff are equal in the two games:

$$g \equiv g_A = g_B \text{ and } h \equiv h_A = h_B$$

Second, consider the Grim Trigger strategy that prescribes mutual cooperation until any unilateral defection and then mutual defection thereafter. Both players choosing Grim Trigger is a subgame perfect equilibrium if game i is infinitely repeated if

$$c_i \geq (1 - \delta)a_i + \delta d_i$$

which can be rearranged as $\delta \geq g/(g + l_i)$. The second parameter restriction is

$$\frac{g}{g + l_A} > \delta \geq \frac{g}{g + l_B}$$

so the Grim Trigger strategy is a subgame perfect equilibrium if B is infinitely repeated, but not if A is infinitely repeated. As Grim Trigger specifies the harshest possible punishment for a defection, it is well-known that mutual cooperation is possible in equilibrium if and only if the Grim Trigger strategy is an equilibrium strategy. So this second condition can be interpreted as mutual cooperation is possible if B is infinitely repeated, but not if A is infinitely repeated.

Third, the stochastic process specifies either A or B in period 0. Then, A is played in period 1 with some probability p and the game realized in period 1 is infinitely repeated thereafter. This means there is only uncertainty about the future in period 0 where the players do not know which game they will infinitely repeat from period 1 onwards.

2.2 The Most Cooperative Equilibrium

The Nash Equilibrium of the one-shot prisoner's dilemma is (D, D). However, just as in the infinitely repeated prisoner's dilemma, there are subgame perfect equilibria with cooperation when the discount factor is large enough. In this section, a *most cooperative equilibrium* is characterized which is the subgame perfect equilibria with the most (C, C) outcomes. It is (usually) not unique, because there can be many off-equilibrium punishments that support cooperation. However, the equilibrium-path outcomes (that is, the outcomes that occur with positive probability) are the same in every most cooperative equilibrium. For simplicity, this section focuses on the most cooperative equilibrium with the harshest possible off-equilibrium punishment as this most cooperative equilibrium, like Grim Trigger for the infinitely repeated games, determines the cutoff discount factor required for players to achieve any mutual cooperation.

The game becomes an infinitely repeated prisoner's dilemma from period 1 on. Therefore, given the second restriction above, (C, C) is possible in periods greater than 0 if and only if B is realized in period 1.

The interesting case for supporting cooperation is period 0. Unlike the infinitely repeated game, inflicting a loss on a defector is not always possible. If A is realized in period 1, then the players must choose D regardless of the period 0 outcome, and so a punishment is only possible when B is realized. That is, the incentive constraint that determines whether cooperation is possible in period 0 when game i is played is

$$(1-\delta)c_i + \delta(pd_A + (1-p)c_B) \geq (1-\delta)a_i + \delta(pd_A + (1-p)d_B)$$

or, in the rearrangement that focuses on gains and losses

$$\delta \geq \frac{g}{g + (1-p)l_B} \tag{1}$$

The equation looks similar to the equation for when Grim Trigger is an equilibrium except that the loss is multiplied by $1 - p$ here. This is capturing the fact that no punishment can be inflicted if A is realized so the loss only occurs when B is realized. Proposition 1 summarizes this most cooperative equilibrium, which will be called the most cooperative equilibrium hereafter for succinctness even though, as mentioned above, it is only the equilibrium-path outcomes that are unique[4]

Proposition 1. *The most cooperative equilibrium prescribes:*

1. *Period 0: cooperation if and only if δ satisfies Equation (1).*

2. *Periods $t \geq 1$: cooperation if B is realized until any player defects when cooperation had been prescribed and defection if A is realized.*

3. *Periods $t \geq 1$: defection forever after any player has defected when cooperation had been prescribed.*

Some features of the most cooperative equilibrium are worth noting. First, the cutoff discount factor for sustaining cooperation in period 0 is directly related to the probability p that A is realized in period 1. This is capturing the analysis that punishment is only possible when B is realized in period 1 and, therefore, the more likely B is, the easier it is to sustain cooperation in period 0. Second, the cutoff is independent of the game played in period 0. This is because the gain from unilateral defection is the same in both games. These two points will form the basis for the hypotheses below that cooperation in period 0 is determined solely by beliefs about the future game, and not at all by the present game.

Third, the equilibrium path of the most cooperative equilibrium is not constant in the sense that there is both mutual cooperation and mutual defection, with probability greater than 0, over the course of the game. Depending on whether Equation (1) is satisfied or not, either players start with cooperation but switch to defection when A is realized in period 1 or start with defection but switch to cooperation when B is realized in period 1. Testing for non-constant play is the other main empirical question investigated in the experiment. There are numerous reasons to think it may

[4] More precisely, this means that any most cooperative equilibrium satisfies points 1 and 2 in the proposition. The last point is just assumed for simplicity as the players never actually must implement it.

be difficult, perhaps most notably that a non-constant strategy requires coordination by subjects who are unable to communicate with each other.

Sometimes defecting is a property of cooperative equilibria that is actually very robust in this setting. There are many equilibrium strategies in these games, not just the most cooperative equilibrium described here. However, because players must defect if game A is realized in period 1, every equilibrium strategy (other than always defect) prescribes both cooperation and defection on the equilibrium path. The results section compares a number of non-equilibrium strategies where the outcomes do not condition on the realization of the period 1 game to equilibrium strategies as a more robust test of whether play corresponds to equilibrium or not.

The feature of the design that the dilemma can only change between periods 0 and 1 and then the game reverts to an infinitely repeated game is not the most natural design, especially if one thinks of the experiment as representing collusion over the business cycle in which case the game should be allowed to change every period, but it is quite useful for the laboratory experiment for many reasons. First, it exaggerates the importance of the parameter p, making the effect of beliefs easier to ascertain, because the realization determines not just the likelihood of the next game but all future games. Second, it implies that the number of different realized games in the experiment is manageable so subjects may have an easier time understanding an already complex environment. Finally, most importantly, it still captures the interesting hypotheses of the more general model, just only in the first two periods. This is not very wasteful as the data in the early periods is the most useful anyways in infinite horizon games, because history dependence (on the outcomes of previous periods) makes it more and more difficult to compare choices further into the game anyways.

2.3 Risk Dominance

Recent experimental evidence on infinitely repeated prisoner's dilemmas has shown that a discount factor for which cooperation is an equilibrium is necessary, but not sufficient, for cooperation to prevail. Two closely related better predictors of cooperation are the size of the basin of attraction (Dal Bó and Fréchette 2011) and the difference between the actual discount factor and the discount factor required for cooperation to be risk dominant (Blonski et al 2011). In the stochastic version here, the size of the basin of attraction can be characterized as a cutoff probability q^* such that the most cooperative equilibrium is a best response when one's opponent plays the most cooperative equilibrium with probability at least q^* and always defects with probability at most $1-q^*$. The values of initial cooperation and defection when one's opponent chooses the most cooperative equilibrium with probability q are

$$\begin{aligned} &\text{C:} \quad q[(1-\delta)c_i + \delta(pd_A + (1-p)c_B)] + (1-q)[(1-\delta)b_i + \delta(pd_A + (1-p)d_B)] \\ &\text{D:} \quad q[(1-\delta)a_i + \delta(pd_A + (1-p)d_B)] + (1-q)[(1-\delta)d_i + \delta(pd_A + (1-p)d_B)] \end{aligned}$$

Setting these expressions equal and solving for the cutoff q, the size of the basin of attraction is

$$q^* = \frac{(1-\delta)h}{(1-\delta)(h-g) + \delta(1-p)l_B}$$

The discount factor required for cooperation to be risk dominant, δ^{RD}, is determined when one's beliefs are $q = 1/2$. Plugging $q = 1/2$ into the expression above and solving for δ results in

$$\delta^{RD} = \frac{g+h}{g+h+(1-p)l_B}$$

The argument of these risk-based approaches is that always defect is also an equilibrium strategy so if it is believed that some players choose cooperative strategies and others always defect, then the value of cooperation decreases and only those subjects who place a high enough belief on facing a cooperative opponent will choose to cooperate. It has been shown in the infinitely repeated prisoner's dilemma that when cooperation is risk dominant ($q < .5$ or $\delta - \delta^{RD} > 0$), there will be significant cooperation, and cooperation rates increase as cooperation become more risk dominant (see the survey paper Dal Bó and Fréchette 2016 for a complete analysis).

Just as for the most cooperative equilibrium, the risk-based approaches predict that period 0 cooperation is more likely when B is more likely in period 1 and independent of the period 0 game. The goal of this paper is to test these predictions, and that is the reason for the parameter restrictions suggested above which mean that the predictions can be tested robustly regardless of which theory better describes behavior. Still, the results will offer evidence showing that these risk-based arguments seem to describe observed cooperation rates almost exactly as they do for infinitely repeated prisoner's dilemmas.

3 Experimental Design

3.1 Treatments

The two treatment variables in the experiment were the game played in period 0 and the parameter p of the stochastic process that determines the period 1 game. The variable p took on the values .75 (A is more likely in period 1) and .25 (B is more likely in period 1). Each of two possible period 0 games and values for p were paired together to make 4 treatments in a standard 2×2 experimental design.

For all four treatments, the payoffs were

Dilemma A

	C	D
C	$32, 32$	$10, 52$
D	$52, 10$	$24, 24$

Dilemma B

	C	D
C	$62, 62$	$10, 82$
D	$82, 10$	$24, 24$

The discount factor was fixed at $\delta = 2/3$ for all four treatments. It was implemented with random termination so the game was technically an indefinitely repeated game that ended after each period with probability 1/3 rather than an infinitely repeated game.[5]

Eighteen sessions with a total of 224 students (10-14 students per session) from the undergraduate population at the University of Virginia were run at the VeconLab at UVA. Each session consisted of play of just one of the 4 treatments and no subject participated in more than one session. Table 1 summarizes the four treatments. For ease of reference, the treatments are named by two letters, the first is the period 0 game and the second is the more likely game for period 1.

Table 1: Treatments

Treatment Name	Period 0	p	Subjects
AA	A	.75	48
AB	A	.25	60
BA	B	.75	54
BB	B	.25	62

Each session consisted of subjects playing the indefinitely repeated stochastic prisoner's dilemma 50 times, each of which will be called a match. Subjects played the entirety of each match with the same other participant from the room, but were randomly rematched with a new participant between matches.

[5] Standard arguments show that these are equivalent if players are risk neutral. This is the most common way to implement an infinite horizon in the laboratory.

Realizations for the random variables of the experiment, specifically match lengths and period 1 dilemmas, were pre-generated before the sessions. Five sets of realizations were generated and each set was used in one of the five sessions of each treatment.[6] Obviously, the realized period 1 dilemmas cannot be the same in treatments where A is more likely in period 1 (Treatments AA and BA) and treatments where B is more likely in period 1 (Treatments AB and BB). However, because A is 75% likely in the first case and B is 75% likely in the second, one set of realizations for period 1 game was generated with a more likely game realized with probability .75 and a less likely game with probability .25. In Treatments AA and BA, the more likely game was set to A and in Treatments AB and BB, the more likely game was set to B.

This design choice maximizes the number of generated realizations such that results comparing behavior across treatments cannot be clouded by different realizations of random variables.[7] The following table provides some descriptive statistics for the five sets of realizations (note that there are 50 matches but the total number of realized period 1 games is less than 50 because some matches only lasted one period).[8]

Table 2: Randomization Statistics

Set	Average Match Length	Longest Match Length	# of More Likely Period 1 Games	# of Less Likely Period 1 Games
1	3.12	10	29	4
2	3.10	9	28	8
3	3.36	18	21	12
4	2.92	12	22	8
5	3.12	10	24	9

Subjects had all the outcomes and payoffs for previous periods of their current match on their screen when making their decision. They also could look up this information for any previous match (although they could only look at one previous match at a time).

All terms were neutrally framed and the experiment was programmed and run in z-tree (Fischbacher 2007). Instructions for Treatment AA are in Appendix A and instructions for other treatments were identical except for the differing parameter values.

[6]Except for the last set which was used only for the fifth session of Treatments AB and BB. There were only four sessions for Treatments AA and BA because there was little variation in behavior there.

[7]A result first found by Engle-Warnick and Slonim 2006a.

[8]The shortest match length is 1 for all three sets.

Subjects earned points corresponding to the payoffs in the matrices above for each period of every match. The points were converted into dollars at the rate of 1/2 cent per point. They also got \$6 for participating. Average earnings were \$33.43.

Finally, an earlier paper, Kloosterman 2013, analyzed data from a separate data set where the payoffs in A and B differed.[9] The main results are qualitatively the same, but game B in that data set did not satisfy the criterion $a_B + b_B < 2c_B$ so there was an incentive to maximize payoffs by switching back and forth between (C, D) and (D, C). Therefore, only the new data described here is analyzed in the results section.

3.2 Predictions

Consider cooperation in period 0 first. The theory section noted that cooperation requires less patience the more likely B is in period 1. The parameters in the experiment make this explicit. Cooperation is possible in period 0 when B is likely ($p = .25$) but not when A is likely ($p = .75$). Table 3 presents choices that are theoretically predicted if players choose the most cooperative equilibrium. Note that Treatments AA and BA (Treatments AB and BB) do not only have D (C) as the equilibrium prediction for period 0 cooperation, but this prediction stems from identical incentive constraints. For periods $t \geq 1$, behavior depends on the realized game in period 1. As explained in the theory section and summarized in Table 3, cooperation is possible when B is realized. Perhaps most interestingly, the path is not constant as the players must switch from cooperation in period 0 to defection in some treatments and vice versa in others.

Table 3: Most Cooperative Equilibrium

Treatment	Period 0	Period 1 on	
		A Realized	B Realized
AA	D	D	C
AB	C	D	C
BA	D	D	C
BB	C	D	C

While the most cooperative equilibrium provides the predictions for the coming hypotheses it should be stressed that the the experiment is not simply a test of this equilibrium. First, as explained in the theory section, every equilibrium (other than

[9] This was a job market paper containing several experiments that has now been separated into two papers; this one and a second paper (Kloosterman 2016).

always defecting) prescribes some cooperation and some defection so the experiment can robustly test for equilibrium play.

Second, in regards to the risk-dominance based arguments for cooperation given above, for Treatments AB and BB, $q^* = .27$ so cooperation is a best response as long as a subject believes their opponent will play cooperatively with probability at least .27 and $\delta^{RD} = .54$ so the difference between the actual discount factor and the discount factor required for cooperation to be risk dominant is $2/3 - \delta^{RD} = .13$. For the infinitely repeated prisoner's dilemma, similar values for q^* and $\delta - \delta^{RD}$ yield significant cooperation which coincides with the prediction of the most cooperative equilibrium. As mentioned above, the point of the experiment is not to address whether these predictors are better than the equilibrium in this environment, but rather to address the question of beliefs about the future and how, or even whether, they determine behavior. Therefore, in order to give beliefs the best chance, the parameters were chosen to compare a case where beliefs are good for cooperation among all measures to a situation where beliefs are bad among all measures.

To summarize, the predictions in Table 3 are robust to different arguments for why cooperation may exist in this game. The goal of the experiment is to determine whether the forward-looking behavior that supports cooperation with threats of punishment is present in this complex environment where such schemes include cooperation and defection and are independent of the period 0 game. With that said, the following three hypotheses formalize the predictions for behavior.

Hypothesis 1. *Period 0 cooperation will be more frequent when B is more likely than A in period 1. More precisely, the amount of period 0 cooperation will be larger in Treatments AB and BB than in Treatments AA and BA.*

Hypothesis 2. *Period 0 cooperation will not depend on the period 0 game. More precisely, the amount of period 0 cooperation will be the same in Treatments AA and BA and in Treatments BB and AB.*

Hypothesis 3. *Period 1 cooperation will be more frequent when B is realized in period 1 than when A is realized in period 1 in all treatments. This means that behavior will not be constant. More precisely, there will be cooperation in period 1 after (D, D) in period 0 in Treatments AA and BA when B is realized in period 1. There will be defection in period 1 after (C, C) in period 0 in Treatments BB and AB when A is realized in period 1.*

4 Results

4.1 Period 0 Behavior

Table 4 displays the cooperation rates (the proportion of C choices) for period 0 behavior for match 1 and the last half of matches.

Table 4: Cooperation Rates for Period 0

Treatment	Match 1	Matches 26-50
AA	.44	.21
AB	.47	.48
BA	.63	.23
BB	.65	.58

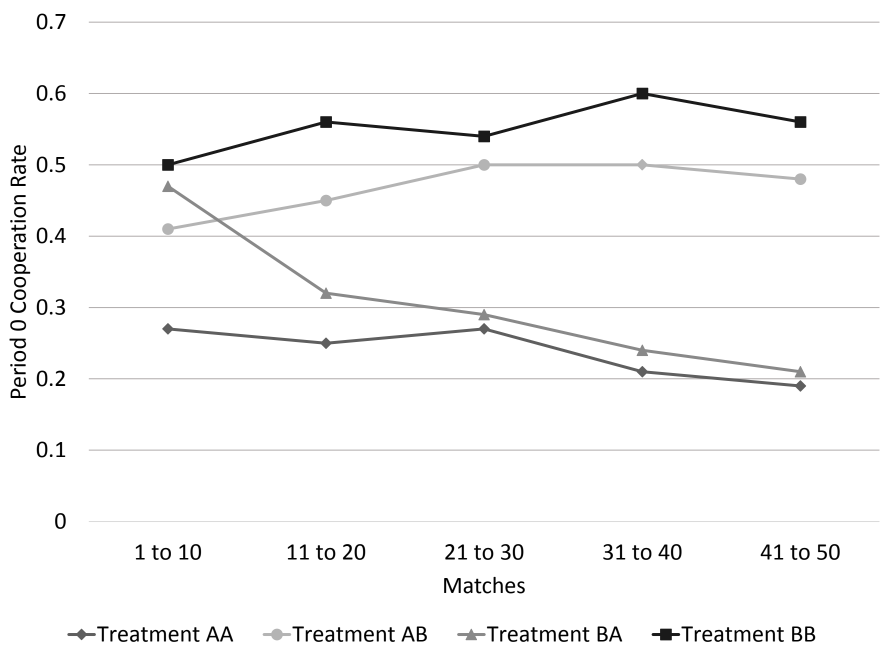
To evaluate Hypotheses 1 and 2, and give theory the best chance, consider the latter column first which incorporates time for subjects to gain experience and learn what choices they want to make. In the last half of matches, there is about 30% more cooperation when the period 1 game is likely to be B lending support for Hypothesis 1. All four such comparisons are statistically significant ($.48 >^{***} .21$, $.48 >^{***} .23$, $.58 >^{**} .21$, $.58 >^{**} .23$).[10] Additionally, the period 0 game is mostly irrelevant which supports Hypothesis 2. There is only 10% more cooperation in Treatment BB than in Treatment AB and 2% more in Treatment BA than in AA and neither difference is statistically significant.

Result 1. *Once subject gain experience, the evidence supports Hypotheses 1 and 2. There is more cooperation when B is more likely in period 1 and cooperation does not depend on the period 0 game.*

On the other hand, Match 1 cooperation appears to reject both Hypotheses 1 and 2. At the beginning of sessions, cooperation is apparently determined by the period 0 game rather than by beliefs about the future. There is about 20% more cooperation when the period 0 game is B. Three of the differences are statistically significant and the last is marginally significant ($.63 >^{**} .44$, $.63 >^{**} .47$, $.65 >^{**} .44$, $.65 >^{*} .47$). Additionally, it is the more likely period 1 game that seems irrelevant. There is just 3% more cooperation in Treatment AB than in Treatment AA and 2% more in Treatment BB than in BA and neither difference is statistically significant.

[10]Significance is calculated with a probit regression on a dummy for one of the two treatments in the comparison (and controls for randomization sets) with standard errors clustered at the session level. As usual, * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level.

Figure 1: Cooperation Rate Across Matches



In order to understand the impact of experience better, Figure 1 disaggregates the cooperation rates into five bins; matches 1-10, matches 11-20, matches 21-30, matches 31-40, and matches 41-50.

Table 4 and Figure 1 are telling a clear story. Theoretically, the experiment is designed so that only beliefs about the future matter. However, when subjects first enter the lab, it seems that they are behaving in exactly the opposite way. Subjects focus on the game they are playing in period 0, and cooperate more in game B. This makes sense in light of the fact that the payoffs to cooperation are larger in game B. However, they learn over the course of the experiment that it is the future game that matters. Subjects switch from some myopic decision rule based on the payoffs in the period 0 game to the more complicated dynamic incentive schemes that actually support cooperation.

Result 2. *Initially, subjects cooperate more when they play game B in period 0 in stark contrast to the predictions of Hypotheses 1 and 2. They learn that when the period 1 game is likely to be A, cooperation cannot be supported in period 0 and that* *the period 0 game is irrelevant.*

This section concludes with two more analyses that go beyond the scope of the hypotheses.

First, while the comparative static result that there is more cooperation when B is likely in period 1 seems to hold, if all subjects adhered to the most cooperative equilibrium, then the cooperation rates would be 1 when B is likely and 0 when A is likely. There is very little cooperation when A is likely as predicted, but the cooperation rate when B is likely still falls well below the upper bound of 1. This result is consistent with the arguments of risk dominance and basin of attraction. In fact, the cooperation rate of about 50% is very similar to past results for the infinitely repeated prisoner's dilemma with similar discount factor differences $\delta - \delta^{RD}$ or basin of attraction cutoffs q^* (see the survey paper Dal Bó and Fréchette 2016 for examples of infinitely repeated prisoner's dilemmas with similar risk parameters). Since there is only one set of risk parameters where cooperation is supported, it's hard to conclusively favor the risk-based argument over equilibrium, but the results do seem to be surprisingly consistent with infinitely repeated prisoner's dilemmas in this respect.

Second, cooperation rates change over the course of the experiment which indicates that learning occurs. It is therefore worthwhile to investigate some of the possible determinants for cooperation that have been studied in past experiments on the infinitely repeated prisoner's dilemma.

There are a number of things that may be affecting subjects' decisions. For each of the match bins defined above for Figure 1, Table 5 presents results from a probit estimation of period 0 cooperation on a number of potential determinants of behavior. The regressors are cooperation in period 1 of match 1 (match1coop), the period 0 game is A (period0A), the likely period 1 game is A (period1A), the opponent cooperated in the first round of the previous match (lastothercoop), the number of periods in the previous match (lastmatchper), and the realized period 1 game in the last match (lastperiod1A or lastperiod1B).[11]

[11] Lastperiod1A and lastperiod1B are not collinear, because some matches end in one period.

Table 5: Determinants of Period 0 Cooperation

	Matches					
	All	2-10	11-20	21-30	31-40	41-50
match1coop	0.987***	1.172***	0.916***	0.841***	0.869***	1.008***
	(0.122)	(0.114)	(0.157)	(0.133)	(0.137)	(0.151)
period0A	-0.0150	-0.111	-0.0344	0.105	-0.0300	-0.00804
	(0.195)	(0.165)	(0.187)	(0.219)	(0.235)	(0.222)
period1A	-0.541***	-0.151	-0.599***	-0.491**	-0.773***	-0.759***
	(0.184)	(0.165)	(0.186)	(0.218)	(0.233)	(0.191)
lastothercoop	0.621***	0.693***	0.615***	0.630***	0.626***	0.543***
	(0.133)	(0.0985)	(0.120)	(0.160)	(0.204)	(0.174)
lastmatchper	0.0118**	0.0465	0.0305	0.00205	0.0281***	-0.00714
	(0.00601)	(0.0317)	(0.0207)	(0.00988)	(0.0108)	(0.0203)
lastperiod1A	0.00388	0.0398	0.106	0.0216	0.0137	0.00454
	(0.0515)	(0.134)	(0.121)	(0.0926)	(0.110)	(0.0552)
lastperiod1B	0.0536	0.161	-0.0632	0.163	0.0286	0.113
	(0.0427)	(0.0994)	(0.117)	(0.112)	(0.130)	(0.0794)
Constant	-0.902***	-1.331***	-0.872***	-0.880***	-0.782***	-0.863***
	(0.223)	(0.212)	(0.225)	(0.269)	(0.294)	(0.231)
Observations	11200	2016	2240	2240	2240	2240

Dependent variable 1 if cooperate

Standard errors in parentheses

* if $p < 0.1$, ** if $p < 0.05$, *** if $p < 0.01$

The table clearly indicates that if one started the experiment by cooperating or if one's opponent cooperated last match, then cooperation is more likely. The latter finding is consistent with a learning story in which subjects are trying to figure out the proportion of cooperators in the room. They then use these beliefs to determine whether to play a cooperative strategy or not along the lines of the argument outlined in the section on risk dominance. Also, the main result noted above carries over in the form of decreased cooperation when A is likely once experience sets in.[12] Finally, there is some evidence for the length of the previous match determining cooperation (the p-values are less than .01 for one match bin and near .1 for a few others, and less than .05 for all matches) suggesting that maybe subjects are learning a little bit about the expected length of matches. However, there is not evidence that subjects are learning about the probability of the period 1 game which seems reasonable given that this statistic is given to them directly (as opposed to the expected length of matches which subjects must calculate from the given continuation probability).

4.2 Period 1 Behavior

Table 6 displays the average cooperation rates in period 1 for the last 25 matches of each treatment for each of the two possible period 1 games.[13] Much of the interest is in how period 0 behavior affects period 1 choices so the last four columns of the table disaggregate the cooperation rates by the outcome of the period 0 game.[14]

Table 6: Cooperation Rates for Period 1, Last 25 Matches

Treatment	Period 1 Game	Period 0 Outcome				
		All	(C,C)	(C,D)	(D,C)	(D,D)
AA	A	.18	.94	.27	.35	.05
	B	.36	.83	.49	.38	.30
AB	A	.34	.86	.33	.36	.07
	B	.49	.99	.44	.41	.19
BA	A	.23	.85	.19	.32	.14
	B	.30	.84	.40	.47	.23
BB	A	.28	.69	.21	.18	.04
	B	.56	.98	.24	.22	.11

Addressing the first statement of Hypothesis 3, Table 6 indicates that there is more total cooperation when B is realized in period 1. Three of the differences are statistically significant and the last is marginally statistically significant ($.36 >^{***} .18$, $.49 >^{***} .34$, $.30 >^{*} .23$, $.56 >^{***} .28$).

[12] There is no complementary significant result for period0A in the early matches. This is probably because, as Figure 1 shows, learning occurs quite quickly.

[13] The analysis is similar if all matches are considered. The last half is presented here to be consistent with period 0 results.

[14] The case where the subject cooperated in period 0 while his/her opponent defected is denoted (C,D) while the case where the subject defected while his/her opponent cooperated is denoted (D,C). In other words, all subjects are treated as Player 1.

There is also some evidence for the non-constant play predicted by the most cooperative equilibrium. In Treatments AA and BA, after the period 0 outcome (D, D), there is indeed more cooperation when B is realized in period 1 ($.30 >^{***} .05$ and $.23 >^{**} .14$). And in Treatments AB and BB, after the period 0 outcome (C, C), there is more defection when A is realized in period 1 ($.99 >^{**} .86$ and $.98 >^{**} .69$).

Result 3. *Subjects cooperate more when B is realized in period 1. There is some evidence for subjects choosing to cooperate after (D, D) in period 0 when B is realized and for choosing to defect after (C, C) in period 0 when A is realized.*

Still, while statistically significant, the differences suggesting some non-constant play are small in magnitude when compared to another clear conclusion evident in the table. Namely, cooperation in period 1 is very much conditional on cooperation in period 0. There is far more cooperation after (C, C) than after any other period 0 outcome. All 24 such comparisons are statistically significant and most are very large in magnitude.[15]

Finally, recall that there is never period 1 cooperation in any equilibrium when A is realized while the finding here is for substantial cooperation when A is realized after (C, C) in period 0. So the results robustly reject equilibrium behavior. Cooperative equilibria are conditionally cooperative, where the conditioning is on both the realized period 1 game and the outcome in period 0. The evidence suggests that subjects do condition their choice on both in the direction predicted by theory, but the results are much stronger for the case of conditioning cooperation on the period 0 outcome. The final section of the results, which will consider fully estimated strategies, will provide further evidence that it is the period 0 outcome that is determining behavior.[16]

4.3 Strategy Estimation

The results so far have focused on cooperation rates. In this final section, the analysis will be extended to try to investigate entire strategies. This can be a difficult task as strategies are plans for all possible contingencies whereas the data contains only the

[15] The results look reminiscent of the Semi-Grim strategy proposed by Breitmoser (2015) where subjects randomize after the outcomes (C, D) and (D, C). While it may look like randomization in aggregate here, the evidence for individual strategies does not look like randomization at all. The total number of subjects (out of 224) that chose cooperate and defect at least once each in the last 25 matches is 16 for period 1 game A after (C, D), 22 for period 1 game B after (C, D), 22 for period 1 game A after (D, C), and 15 for period 1 game B after (D, C). And of these already less than 10% of subjects, just 2, 6, 3, and 5 of them chose each strategy at least twice for the 4 respective cases.

[16] As a robustness test of the findings here, Table 9 in the appendix considers cooperation rates for periods 1 and greater. One might suspect that perhaps cooperation rates decrease considerably over the course of the game when A is realized after (C, C) as cooperation can not be supported there. The table shows that this is not the case, and the results noted in this section are robust to considering all periods after period 0.

contingencies that actually occur. To overcome this obstacle, Dal Bó and Fréchette 2011 developed a maximum likelihood estimation technique that can be used in this setting as well. The basic idea is that a number of possible strategies are posited, and then maximum likelihood is used to estimate the most likely frequency with which each of these strategies is used by subjects.

The set of possible strategies considered here consists of several of the classic strategies that have been identified in the infinitely repeated prisoner's dilemma. Then variants of the strategies are added for reasons to be noted momentarily. Table 7 explains the classic strategies and then the variants.

The way the variants work is to modify the classic Grim Trigger and Tit-for-Tat strategies. Defect and Suspicious are similar, but differ in that the former ignores period 0 behavior while the latter punishes period 0 defections by one's opponent. Stochastic is called such, because it conditions behavior on the outcome of the stochastic process in the way theory predicts. Also, Defect and Suspicious are mutually exclusive while Stochastic can be added to either the original version or in addition to either of these two variants. In total, this creates a set of 12 strategies.[17]

Table 7: Strategies

Strategy	Short-Hand	Explanation
Always Defect	AD	always choose D
Always Cooperate	AC	always choose C
Grim Trigger	Grim	start with C and then continue with C until any player defects after which choose D forever after
Tit-for-Tat	Tft	start with C and then choose the action that the other player chose in the previous period
Variant	**Prefix**	**Modification**
Defect	D	start with D then start strategy in period 1 ignoring period 0
Suspicious	Susp	start with D rather than C
Stochastic	Stoch	choose D in all periods from 1 on if A is realized in period 1

The reason for the variants Defect and Stochastic is that they are the variants that

[17]SuspGrim is not included, because the player pulls the trigger on their own initial defection so it is observationally identical to AD.

create the most cooperative equilibrium strategy. In Treatments AA and BA, the most cooperative equilibrium is Stochastic Defect Grim Trigger while in Treatments AB and BB, it is Stochastic Grim Trigger. Suspicious is also included, because it is similar to Defect and has been found to be prevalent in previous strategy estimations on the infinitely repeated prisoner's dilemma.

Recall once more that there are many equilibrium strategies (not just the most cooperative equilibrium). And, importantly, a common theme among these strategies is that the players must defect when A is realized so every equilibrium strategy is Stochastic (or AD which also, of course, defects when A is realized). Table 8 presents the results of the maximum likelihood estimates on these 12 strategies for each of the four treatments.[18]

Consider Treatments AA and BA first. Again, the most cooperative equilibrium strategy for these two treatments is Stochastic Defect Grim Trigger. There is essentially no evidence for the use of this strategy.

Not surprisingly, given the low cooperation rate for period 0 noted in Table 4, the most prominently estimated strategy is the equilibrium strategy AD. However, there are a few cooperative strategies. The only strategy that starts with cooperation with a significant estimate is Tit-for-Tat. There are also some partially cooperative strategies that start with defection and then go to cooperation; Suspicious Tit-for-Tat and, in Treatment BA, Defect Tit-for-Tat. None of these three variant strategies of Tit-for-Tat are Stochastic, and therefore none are equilibrium strategies. Rather, they are conditionally cooperative strategies that condition on past behavior only supporting the assertion from Table 6 that this type of conditioning is more prevalent than conditioning behavior on the period 1 realized game.

Now consider Treatments AB and BB. For these treatments, the most cooperative equilibrium strategy is Stochastic Grim Trigger, and while there is some mild evidence for this strategy in Treatment BB there is little evidence for it in Treatment AB. As for the previous two treatments, there is far more evidence for just conditionally cooperative strategies that condition on past behavior but are not equilibria. In this case, Grim Trigger in Treatment BB, Tit-for-Tat in both treatments (particularly in Treatment AB), and Suspicious Tit-for-Tat in Treatment AB.

Result 4. *Many subjects choose to always defect. Those that choose cooperative strategies, choose to cooperate on the condition that their opponent has cooperated in*

[18] Appendix B provides the details of how the likelihood function is formed. It also includes a second estimation where three more strategies are included. Although there seems to be minor evidence for a strategy called Tit-for-2-Tats (and one variant of it), the main takeaways do not change with the extra strategies.

previous periods but not on the condition that B is realized in period 1.

Table 8: Strategy Estimates, Last 25 Matches

Strategy	Treatment AA	Treatment AB	Treatment BA	Treatment BB
AD	0.3919***	0.2409***	0.4451***	0.3096***
	(0.0896)	(0.0819)	(0.1025)	(0.1223)
AC	0.0208	0.0655*	0.0185	0.0323
	(0.0270)	(0.0483)	(0.0209)	(0.0295)
Grim	0.0418	0	0.0489	0.2264**
	(0.0379)	(0.0638)	(0.0437)	(0.1089)
DGrim	0.017	0	0	0
	(0.0246)	(0.0123)	(0.0008)	(0.0047)
StochGrim	0	0.0806	0	0.1449*
	(0.0088)	(0.0570)	(0.0131)	(0.0925)
StochDGrim	0.0402	0.0204	0	0.0095
	(0.0625)	(0.0230)	(0.0426)	(0.0140)
Tft	0.1216**	0.352***	0.1142**	0.1161*
	(0.0528)	(0.0732)	(0.0660)	(0.0692)
DTft	0.0209	0.0061	0.0896**	0
	(0.0277)	(0.0249)	(0.0459)	(0.0087)
SuspTft	0.2215***	0.1029*	0.1903***	0.0829
	(0.0685)	(0.0504)	(0.0680)	(0.0568)
StochTft	0	0.0179	0	0.061
	(0.0165)	(0.0262)	(0.0093)	(0.0479)
StochDTft	0.0629	0.0778	0.067	0.0173
	(0.0691)	(0.0580)	(0.0515)	(0.0224)
StochSuspTft	0.0614	0.0359	0.0264	0
Gamma	0.4384***	0.4426***	0.4251***	0.3682***
	(0.0472)	(0.0387)	(0.0390)	(0.0264)
Observations	3722	4518	4128	4634

Bootstrapped standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

So why do subjects choose these conditionally cooperative strategies that are not equilibria? The most likely reason is that social preferences have a role here. For example, a subject who faces an opponent playing Grim Trigger may choose to cooperate in period 0 and then continue cooperating in period 1, even if A is realized, if they do not want to harm their opponent with whom they have built a cooperative relationship.

The question then becomes whether the equilibrium and risk dominance theories of behavior should both be rejected. In the strict sense the answer is yes, but simply rejecting these theories on the basis of cooperation in A is too simplistic. The evidence for how beliefs about the future determine cooperation is consistent with them. It is probably more the case that these theories are good at describing behavior, just not perfect due to some social preference of subjects. If behavior was purely social afterall, then the strategies would be just cooperative rather than conditionally cooperative. Furthermore, there is some weak evidence, recall Table 6, that perhaps a few subjects are trying to switch from cooperation to defection and vice versa as the equilibrium predicts.

It is important to highlight that there could be social preferences of this sort in the infinitely repeated prisoner's dilemma too, but the simple nature of that environment does not allow separate identification of equilibrium (or risk-based) play from social preference. Following the above example, a subject may want to maintain a cooperative relationship once they achieve one, but this is exactly what an equilibrium, such as Grim Trigger, also prescribes in the infinitely repeated prisoner's dilemma. So the stochastic prisoner's dilemma is actually a nice way to separately identify the behaviors, and see that yes, there does seem to be some role for social preferences in dynamic interactions.

These results open the door for future work to investigate the role for how cooperative behavior established early in a game carries over to later in the game when cooperation is not possible. This is perhaps related to Peysakhovich and Rand 2015 who investigate how cooperative behavior in a repeated prisoner's dilemma carries over to other games. They show that establishing a norm of cooperation makes people more prosocial in other games while establishing a norm of defection has the opposite effect. The results here suggest that perhaps the same norms can be established within the game too.

5 Conclusions

Understanding cooperation in infinitely repeated games is a focus of much recent experimental work. In this paper, the analysis is extended to a more general environment where new issues, beliefs about future games and non-constant play in the equilibrium, become important determinants of cooperation. The findings for how beliefs about the future impact cooperation in the present are consistent with the theoretical predictions once subjects gain experience with the game. Cooperation prevails when beliefs about the future support a large scope for punishment.

On the other hand, the most cooperative equilibrium, and in fact every cooperative equilibrium, is non-constant and the results are more mixed as to whether subjects play the games this way. There is strong evidence that behavior is conditioned on past behavior, one characteristic of cooperative equilibrium strategies, but much weaker evidence for conditioning behavior on the realized period 1 game.

This finding is potentially important for real world problems. If agents play stochastic games that start out amenable to cooperation, then perhaps cooperative relationships can carry over when the game no longer supports cooperation. In this way, perhaps outcomes can be very efficient with just a small nudge at the beginning wherein the players develop a cooperative relationship. Or on the other hand, outcomes can also be potentially inefficient when agents' relationships start out non-cooperative and they are unable to return to cooperation once the game allows it.

References

[1] Aoyagi, Masaki and Fréchette, Guillaume R. Collusion as Public Monitoring Becomes Noisy: Experimental Evidence. *Journal of Economic Theory* 2009. 144(3), pp. 1135-1165.

[2] Bagwell, Kyle and Staiger, Robert W. Collusion over the Business Cycle. *RAND Journal of Economics* 1997. 28(1), pp. 82-106.

[3] Blonski, Matthias, Ockenfels, Peter, and Spagnolo, Giancarlo. Equilibrium Selection in the Repeated Prisoner's Dilemma: An Axiomatic Approach and Experimental Evidence. *American Economic Journal: Microeconomics*. 2011, 3(3), pp. 164-192.

[4] Breitmoser, Yves. Cooperation, but No Reciprocity: Individual Strategies in the Repeated Prisoner's Dilemma. *American Economic Review*. 2015, 105(9), pp. 2882-2910.

[5] Dal Bó, Pedro. Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. *American Economic Review* 2005. 95(5), pp. 1591-1604.

[6] Dal Bó, Pedro and Fréchette, Guillaume R. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. *American Economic Review* 2011. 101(1), pp. 411-429.

[7] Dal Bó, Pedro and Fréchette, Guillaume R. On the Determinants of Cooperation in Infinitely Repeated Games: A Survey. *Journal of Economic Literature* 2016. forthcoming.

[8] Dreber, Anna, Rand, David G., Fudenberg, Drew, and Nowak, Martin A. Winners Don't Punish. *Nature* 2008. 452(7185), pp. 348-351.

[9] Duffy, John and Ochs, Jack. Cooperative Behavior and the Frequency of Social Interaction. *Games and Economic Behavior* 2009. 66(2), pp. 785-812.

[10] Engle-Warnick, Jim and Slonim, Robert L. The Evolution of Strategies in a Repeated Trust Game. *Journal of Economic Behavior and Organization* 2004. 55(4), pp. 553-573.

[11] Engle-Warnick, Jim and Slonim, Robert L. Inferring Repeated-Game Strategies from Actions: Evidence from Trust Game Experiments. *Economic Theory* 2006. 28(3), pp. 603-632.

[12] Engle-Warnick, Jim and Slonim, Robert L. Learning to Trust in Indefinitely Repeated Games. *Games and Economic Behavior* 2006. 54(1), pp. 95-114.

[13] Feinberg, Robert M. and Husted, Thomas A. An Experimental Test of Discount-Rate Effects on Collusive Behavior in Duopoly Markets. *Journal of Industrial Economics* 1993. 41(2), pp. 153-160.

[14] Feinberg, Robert M. and Snyder, Christopher M. Collusion with Secret Price Cuts: An Experimental Investigation. *Economics Bulletin* 2002. 3(6), pp. 1-11.

[15] Fischbacher, Urs. z-Tree: Zurich Toolbox for Ready-made Economic Experiments *Experimental Economics* 2007, 10(2), pp. 171-178.

[16] Fudenberg, Drew, Rand, David G., and Dreber, Anna. Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World. *American Economic Review* 2012. 101(2), pp. 720-749.

[17] Haltiwanger, John and Harrington Jr., Joseph E. The Impact of Cyclical Demand Movements on Collusive Behavior. *RAND Journal of Economics* 1991. 22(1), pp. 89-106.

[18] Holcomb, James H. and Nelson, Paul S. The Role of Monitoring in Duopoly Market Outcomes. *The Journal of Socio-Economics* 1997. 26(1), pp. 79-93.

[19] Holt, Charles A. An Experimental Test of the Consistent-Conjectures Hypothesis. *American Economic Review* 1985. 75(3), pp. 314-325.

[20] Kandori, Michihiro. Correlated Demand Shocks and Price Wars During Booms. *The Review of Economic Studies* 1991. 58(1), pp. 171-180.

[21] Kloosterman, Andrew. An Experimental Study of Public Information in Markov Games. *Working Paper* 2013.

[22] Kloosterman, Andrew. An Experimental Study of Public Information in the Asymmetric Partnership Game. *Working Paper* 2016.

[23] Murnighan, J. Keith and Roth, Alvin E. Expecting Continued Play in Prisoner's Dilemma Games: A Test of Several Models. *Journal of Conflict Resolution* 1983. 27(2), pp. 279-300.

[24] Palfrey, Thomas R. and Rosenthal, Howard. Repeated Play, Cooperation and Coordination: An Experimental Study. *Review of Economic Studies* 1994. 61(3), pp. 545-565.

[25] Peysakhovic, Alex and Rand, David. Habits of Virtue: Creating Norms of Cooperation and Defection in the Laboratory. *Management Science* 2015. 62(3), pp. 631-647.

[26] Rojas, Christian. The Role of Demand Information and Monitoring in Tacit Collusion. *RAND Journal of Economics* 2012. 43(1), pp. 78-109.

[27] Rotemberg, Julio J. and Saloner, Garth. A Supergame-Theoretic Model of Price Wars duing Booms. *American Economic Review* 1986. 76(3), pp. 390-407.

[28] Roth, Alvin E. and Murnighan, J. Keith. Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma. *Journal of Mathematical Psychology* 1978. 17(2), pp. 189-198.

[29] Ruffle, Bradley J. When Do Large Buyers Pay Less? Experimental Evidence. *Journal of Industrial Economics* 2013. 61(1), pp. 108-137.

[30] Saijo, Tatsuyoshi, Sherstyuk, Katerina, Tarui, Nori, and Ravago, Majah-Leah. Games with Dynamic Externalities and Climate Change Experiments. *Journal of the Association of Environmental and Resource Economists* 2015, forthcoming.

[31] Vespa, Emanuel. An Experimental Investigation of Strategies in the Dynamic Common Pool Game. *Working Paper* 2015.

[32] Wilson, Alistair and Vespa, Emanuel. Experimenting with Equilibrium Selection in Dynamic Games. *Working Paper* 2016.

A Instructions

Welcome. This is an experiment in decision making. Various research foundations and institutions have provided funding for this experiment and you will have the opportunity to make a considerable amount of money which will be paid to you at the end. Make sure you pay close attention to the instructions because the choices you make will influence the amount of money you will take home with you today. Please ask questions if any instructions are unclear.

The Choice Problems

In this experiment you will be engaging in two Choice Problems. You will engage in each Choice Problem with one other participant, which we will call your partner. Each Choice Problem consists of both you and your partner choosing between two options, which we label A and B. Thus there are four possible outcomes; you choose A and your partner chooses A, you choose A and your partner chooses B, you choose B and your partner chooses A, and you choose B and your partner chooses B. Each of these four outcomes results in a payoff for both you and your partner. The following two tables show these payoffs.

Choice Problem 1 is given by the following table.

Your Choice	Partner's Choice	Your Payoff	Partner's Payoff
A	A	15	15
A	B	8	45
B	A	45	8
B	B	12	12

Choice Problem 2 is given by the following table.

Your Choice	Partner's Choice	Your Payoff	Partner's Payoff
A	A	65	65
A	B	8	95
B	A	95	8
B	B	12	12

To illustrate, if you and your partner are engaging in Choice Problem 1 and you choose A while your partner chooses B then you will get 8 and your partner will get 45. Or if you and your partner are engaging in Choice Problem 2 and you choose B while your partner chooses A then you will get 95 while your partner will get 8.

Matches and Rounds

The experiment consists of 50 Matches. At the beginning of each Match you will be randomly paired with another participant from the room and engage in a randomly determined number of Rounds, each of which is one of the two Choice Problems above, with this participant as your partner.

Each of the 50 Matches proceeds as follows. In Round 1 of each Match you and your partner engage in Choice Problem 1. When Round 1 is over, the computer will randomly determine whether the Match will continue to Round 2 or end. The computer is programmed to select to continue with 2/3 chance and to end with the remaining 1/3 chance. This is true for all further rounds as well. That is, at the end of each Round, the Match continues with probability 2/3 and ends with probability 1/3.

If the Match continues at the end of Round 1, the Round 2 Choice Problem is selected randomly by the computer. The computer is programmed to select Choice Problem 1 with 3/4 chance and Choice Problem 2 with the remaining 1/4 chance. If the Match continues again, the Round 3 Choice Problem is the problem selected randomly by the computer for Round 2. In fact, as long as the Match continues, all subsequent Rounds; fourth, fifth, sixth, etc. will be the Choice Problem determined for Round 2. For example, if the Match lasts 4 Rounds, you and your partner will either engage in Choice Problem 1 four times or Choice Problem 1 once followed by Choice Problem 2 three times.

Thus there are two important things to remember when a new Match starts. First, you get a new partner. Second, you return to Choice Problem 1 and (if the new Match continues to Round 2) you will have a randomly determined Choice Problem in Round 2 that will remain the Choice Problem for all further Rounds of the new Match.

Your Screen

Your screen is laid out as follows. In the upper left corner you will see which of the 50 Matches you are currently in. At the top in the middle you will see what Round you are in and which of the two Choice Problems you are engaging in for this Round. Below this you will have the table (from above)for this Choice Problem to remind you what payoffs each outcome produces. Below the table is the box where you make your Choice, A or B, for this Round by clicking on the corresponding button.

You can also see the outcomes and payoffs of all past Choice Problems you have engaged in. On the left side of your screen, the previous Rounds of the current Match are displayed. On the right side of your screen you can enter the Match number (then click Check) of any previous Match to see the outcomes and payoffs for that Match. The last row of each display presents cumulative payoffs. In other words, this row sums up the column for your payoff and the column for your partner's payoff.

A second screen will show up after each Round as well. It will tell you the outcome for that Round and whether the Match will continue or end. Also, if you just finished Round 1 of a Match and the Match is continuing it will tell which of the two Choice Problems the computer has selected and thus you will engage in for the rest of the Match.

Payoffs

We will add up all of your payoffs over the course of the experiment. These payoffs are denominated in points and will be converted into dollars at the rate of .5 cents per point. That is, for each 200 points you earn, you get one dollar. In addition to earnings in the experiment you will get 5 dollars just for participating.

B Cooperation in Periods Greater than Zero

Table 9: Cooperation Rates for Periods $t \geq 1$, Last 25 Matches

Treatment	Period 1 Game	Period 0 Outcome				
		All	(C, C)	(C, D)	(D, C)	(D, D)
AA	A	.16	.84	.38	.33	.03
	B	.32	.88	.48	.30	.14
AB	A	.29	.92	.35	.33	.03
	B	.42	.98	.44	.35	.05
BA	A	.18	.93	.28	.36	.05
	B	.28	.95	.43	.37	.11
BB	A	.18	.77	.28	.13	.02
	B	.46	.95	.27	.21	.05

C Likelihood and Extended Strategy Estimation

The likelihood function is constructed following the approach of Dal Bó and Fréchette 2011. Suppose subject i is following strategy s^k. Let cooperate be coded as 1 and defect as 0, and s^k_{mt} be the choice prescribed by s^k in match m and period t. The model is that subject i chooses $y_{imt} = 1$ if $\mathbb{1}_{s^k_{mt}=1} - \mathbb{1}_{s^k_{mt}=0} + \gamma\epsilon_{imt} \geq 0$ in match m and period t where ϵ_{imt} is an error term with variance parametrized by γ. Otherwise subject i chooses $y_{imt} = 0$. The error term is logistic so the probability i chooses strategy s^k given the observed choice y_{imt} is

$$\left(\frac{1}{1+e^{-s^k_{mr}/\gamma}}\right)^{y_{imt}} \left(\frac{1}{1+e^{s^k_{mr}/\gamma}}\right)^{1-y_{imt}}$$

Let p_k be the proportion of subjects that choose strategy s_k (these proportions, in addition to γ, are the parameters that will be estimated). For a set of K strategies indexed by k, the likelihood function is therefore

$$\prod_{i=1}^{I}\sum_{k=1}^{K} p_k \prod_{m=26}^{50}\prod_{t=0}^{T_m} \left(\frac{1}{1+e^{-s^k_{mr}/\gamma}}\right)^{y_{imt}} \left(\frac{1}{1+e^{s^k_{mr}/\gamma}}\right)^{1-y_{imt}}$$

where$I$ is the total number of subjects and T_m is the realized length of match m.

Standard errors are bootstrapped by resampling (with replacement) the data 1000 times. In order to control for session effects, first a number of sessions equal to the total number of sessions for a given treatment are randomly resampled. Then, a number of subjects equal to the number of subjects in the original session are resampled from each resampled session. Then, 25 matches from the last half are resampled for each resampled subject.

The final section here describes an extended strategy estimation with a few additional strategies.

Table 10: Additional Strategies

Strategies	Short-Hand	Explanation
Trigger 2	T2	start with C and then continue with C until any player defects after which choose D for two periods before returning to C
Tit-for-2 Tats	Tf2t	start with C and then stay with C unless other player chose D in both of the last 2 periods
2 Tits-for-Tat	T2ft	start with C and then stay with C unless other player chose D in either of the last 2 periods

Table 11: Augmented Strategy Estimates: Last 25 Matches

Strategy	Treatment AA	Treatment AB	Treatment BA	Treatment BB
AD	0.4017***	0.2396***	0.3892***	0.2890***
	(0.0941)	(0.0808)	(0.1051)	(0.117)
AC	0.0208	0.0169	0	0.0161
	(0.0253)	(0.0275)	(0.0086)	(0.0229)
Grim	0.0421	0	0.0501	0.2328**
	(0.0387)	(0.0628)	(0.0431)	(0.1106)
DGrim	0.0173	0	0	0
	(0.0239)	(0.0089)	(0.0008)	(0.0046)
StochGrim	0	0.0833	0	0.1441*
	(0.0053)	(0.0541)	(0.0111)	(0.0872)
StochDGrim	0.0177	0	0	0.0122
	(0.0412)	(0.0125)	(0.0302)	(0.013)
Tft	0.1005**	0.2708***	0.0989*	0.0934
	(0.0436)	(0.0677)	(0.06)	(0.0644)
DTft	0.0209	0.0096	0.0192	0
	(0.0208)	(0.0222)	(0.0285)	(0.0059)
SuspTft	0.1971***	0.0761	0.1859***	0.0638
	(0.0728)	(0.0458)	(0.0643)	(0.0481)
StochTft	0	0	0	0.0618
	(0.0162)	(0.0074)	(0.0029)	(0.0481)
StochDTft	0	0.0304	0	0
	(0.019)	(0.0315)	(0.0127)	(0.0068)
StochSuspTft	0.036	0.0367	0	0
	(0.0414)	(0.0313)	(0.0231)	(0.0144)
T2	0	0	0	0
	(0.0006)	(0.0052)	(0.0006)	(0.0006)
DT2	0	0	0.0184	0
	(0.0048)	(0.0039)	(0.0168)	(0.0001)
StochT2	0	0	0	0
	(0.0021)	(0.0011)	(0.0004)	(0.0176)
StochDT2	0	0	0	0
	(0)	(0.0015)	(0.0082)	(0.0027)
Tf2t	0.0208	0.1058**	0.0185	0.0324
	(0.0297)	(0.0571)	(0.0264)	(0.0292)
DTf2t	0	0	0.0150	0
	(0.0147)	(0.0092)	(0.0212)	(0.003)
SuspTf2t	0.0192	0	0.0394	0
	(0.0202)	(0.0142)	(0.0322)	(0.0038)

Strategy	Treatment AA	Treatment AB	Treatment BA	Treatment BB
T2ft	0	0.0055	0	0
	(0.0144)	(0.0409)	(0.0155)	(0.024)
DT2ft	0	0	0	0
	(0.0057)	(0.0011)	(0.0042)	(0.0004)
SuspT2ft	0	0.0207	0.0180	0
	(0.0204)	(0.0256)	(0.032)	(0.0236)
StochT2ft	0	0	0	0
	(0.0069)	(0.0139)	(0.0086)	(0.0333)
StochDT2ft	0	0	0	0
	(0.0093)	(0.0039)	(0.0058)	(0.0025)
StochSuspT2ft	0	0.0185	0.0679	0.0382
Gamma	0.4263***	0.4238***	0.4142***	0.3590***
	(0.0428)	(0.034)	(0.0361)	(0.0229)
Observations	3722	4518	4128	4634

Bootstrapped standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$