

MOTIVES BEHIND COOPERATION IN FINITELY REPEATED PRISONER’S DILEMMA

ANUJIT CHAKRABORTY*

ABSTRACT. This paper deploys a novel experiment to compare three behavioral theories that explain both selfish and non-selfish cooperation. The three theories fuse reputational cooperation (à la [Kreps et al. \[1982\]](#)) with the following three non-selfish motives respectively: caring about others (Altruism), being conscientious about cooperation (Duty), and enjoying social-efficiency (Efficiency-Seeking). We use reputational cooperation under purely Selfish preferences as a fourth theory. Our experimental design varies the decline-rate of future rewards, under which these theories predict rich patterns of behavior. Based on a Finite Mixture Model, the data is best explained if the modal subject types are Selfish and Efficiency-Seeking: We estimate that 40-49% of our subjects are Selfish, 36-45% are Efficiency-seeking, and 6-20% are Altruistic. We find little evidence for Duty players. *JEL Codes:* C72, C73, C92.

What motivates cooperation in interactions that can be modeled as Finitely Repeated Prisoner’s Dilemma (FRPD)? According to [Kreps et al. \[1982\]](#)’s reputation theory, it is the pursuit of higher payoffs under incomplete information: all players are *selfish* but they are uncertain if their partners are the standard selfish-type or a behavioral type who plays a conditionally cooperative strategy. Under this uncertainty, selfish individuals cooperate and maintain a cooperative reputation whenever it is profitable.¹ [Andreoni and Miller \[1993\]](#), [Cooper](#)

Key words and phrases. Experimental Economics, Finitely Repeated Prisoner’s Dilemma, Behavioral game theory.

* chakraborty@ucdavis.edu. Department of Economics, University of California, Davis.

** I would like to thank Yoram Halevy, Ryan Oprea, Yaroslav Rosokha, and seminar & conference attendees at UBC, IIMA, IIMB, IITK, Indiana University, Purdue University, NYU CESS, and the Rady Spring School 2017 for useful comments and suggestions. I would also like to thank Bikramaditya Datta, Jacob Schwartz and the seminar speakers at UC Davis for their valuable advice, and SSHRC (FAS # F11-04991) for financial support.

¹[Kreps et al. \[1982\]](#) say “The issue then is whether this puzzle can be resolved in the context of rational, self-interested behavior”. Thus, they consider purely selfish agents who hold wrong beliefs about the existence of a conditionally cooperative type among them.

et al. [1992] find the high levels of cooperation observed in the middle and terminal rounds of FRPD inconsistent with “reputation theory with *purely* selfish agents”. Similarly, Cox et al. [2015] run a sequential-move FRPD where they reveal the second-mover’s past actions from an earlier FRPD game to the first-mover. Even though pure reputation-building predicts complete unraveling, Cox et al. [2015] find significant cooperation, consistent with the presence of non-pecuniary motives for cooperation. Andreoni and Miller [1993], Cooper et al. [1992] propose altruism as the complementary explanation and comment “all of the alternative models of altruism can be viewed as one of, or some combination of” Pure Altruism (Dawes [1980]), Duty (Palfrey and Rosenthal [1988], Cooper et al. [1992]), and Efficiency-Seeking (Kreps et al. [1982]).

To reconcile reputation and other-regarding preferences under a simple-unified framework, this paper constructs three new behavioral models. In each model we combine the uncertainty about partner’s strategy (à la reputation) with the players themselves having one of the three above-mentioned types of altruistic preferences: (i) caring about others (Pure Altruism), (ii) being conscientious about cooperation (Duty), and (iii) enjoying the efficiency of mutual cooperation (Efficiency-Seeking). We assume that players update their beliefs every round and play actions that maximize the sum of pecuniary and non-pecuniary payoffs over the supergame. This gives us three new models of behavior alongside the fourth standard (selfish reputational) model of maximizing only pecuniary payoffs. This paper estimates the proportion of Selfish (reputational), Pure Altruism, Duty, and Efficiency-Seeking behavior that best explains FRPD cooperation.

Our experimental variation is based on the theoretical observation that these four models predict rich patterns of behavior in a perturbed FRPD that has a declining payoff-profile. We run 5-period FRPD games where the payoffs decline geometrically: All payoffs in period- t are only δ^{t-1} times the period-1 payoffs, and this is common knowledge. We have four treatments with $\delta = 1, 3/4, 3/8, 1/4$. Thus, $\delta = 1$ is the standard FRPD game and $\delta < 1$ determines how quickly payoffs diminish across periods 1 to 5. Previous FRPD experiments have either varied the stage-game payoffs across supergames, or

varied the length of the interaction for a fixed stage-game. Our treatment variation is novel: by using declining payoffs within a supergame we can derive testable predictions from the theories. In Section III, we organize the predictions based on if they address initial cooperation, terminal cooperation, or the link between initial and terminal cooperation.

Section V operationalizes our main empirical analysis, a Finite Mixture Model, that estimates the combination of Selfish, Pure Altruism, Duty and Efficiency-Seeking behavior that best explains our data. We preface this analysis with the description of the three central patterns in the data (Section IV), and when feasible, link a pattern with the behavioral model(s) from Section III that they are most consistent with. The first pattern is, *initial cooperation increases as δ increases*, and this is consistent with all three behavioral models. Second, *terminal cooperation decreases as δ increases*, and this is a pattern exclusively explained by Duty and Efficiency-Seeking models. Third, *players consistently reciprocate their partner's most recent cooperative or non-cooperative action, even if either player defected in the initial periods*. This pattern is inconsistent with the Duty model that predicts first-period cooperators would cooperate unconditionally in the final periods of the $\delta = 1/4$ supergame.² Thus, Reciprocal Cooperation is the only behavioral model consistent with all three patterns.

Next, using our Finite Mixture Model, we jointly estimate the proportion of each of the four utility types in the population, the utility parameters of Altruism/ Duty/ Efficiency-Seeking models, how beliefs evolve after each round, and how sensitive players are to utility differences. All specifications of our estimation exercise paint a similar picture: 40-49% of our subjects are selfish, 36-45% are Efficiency-seeking, 6-20% are Altruistic, and 1-4% are Duty players. Thus, we find that Efficiency-Seeking is the modal other-regarding behavior whereas there are barely any Duty players.

²We also consider two other models of reciprocal behavior, Simple Reciprocity and Sequential Reciprocity [Dufwenberg and Kirchsteiger, 2004], but they are both inconsistent with the third pattern.

In Section VI, we categorize subject behavior based on their conditional response to partner’s last-period behavior, and find results that echo the structural estimation results. Particularly, the $\delta = 1/4$ treatment allows us to characterize subjects into types. We find that around one-third of those subjects behave purely selfishly, defecting almost every time, irrespective of their partner’s action. Another third of subjects, have a highly asymmetric response to partner’s cooperation: they behave like Efficiency-Seekers for whom partner’s last period cooperation implies high probability of partner-cooperation in the upcoming period. Only 4% of subjects cooperate unconditionally, rest are difficult to characterize. In Section VII we link our results to recent studies and Section VIII concludes.

I. AN OVERVIEW OF THE LITERATURE

The experimental literature on cooperation in PD games is vast. We discuss some of the relevant papers here at the risk of inadvertently missing out on others.

Our models were motivated by the multiple papers that establish reputational cooperation (Kreps et al. [1982]) in both finite and indefinitely repeated PD games. Roth and Murnighan [1983] find that players cooperate more at higher continuation-probabilities in indefinitely-repeated PD. Bó [2005] replicates this result and additionally disentangles higher continuation-probabilities from higher expected-interactions. Andreoni and Miller [1993] test the reputation model by varying the probability that subjects interact with a computer that plays a Tit-For-Tat strategy. They find that higher beliefs about playing the computer are more conducive to higher cooperation. Bereby-Meyer and Roth [2006] find more cooperation in the first period of FRPDs than in the one-shot games. Their exercise is similar to comparing first-period cooperation rates of $\delta = 1$ and $\delta = 0$ in our setting. Dal Bó and Fréchette [2011], Embrey et al. [2018] show that sizeBAD captures the value of reputational cooperation and the strategic uncertainty of cooperating in infinitely and finitely repeated PD games respectively. Calford and Oprea [2017] and Friedman and Oprea [2012] study cooperative behavior in continuous-time versions of PD, and find that unraveling of cooperation can stop when players can react quickly.

A complementary literature calls for unifying reputation with with other behavioral forces. [Cooper et al. \[1992\]](#) find evidence of both reputation-building and altruism. They conclude that neither can explain all the features of the data on its own. [Mengel \[2018\]](#) disentangles the role of ‘risk’ (to co-operate unilaterally) and ‘temptation’ (to defect against a co-operator) to find that they are good predictors of cooperation in one-shot and finitely repeated games respectively.

Payoff discounting, used in this paper, has previously been used by [Fréchette and Yuksel \[2017\]](#) as one of four different implementations of infinitely repeated games in the laboratory. In [Fréchette and Yuksel \[2017\]](#)’s implementation, a fixed (known) number of rounds were played with certainty, and payoffs in these rounds were discounted at a known rate $\delta \in (0, 1)$. After the rounds with certainty, the supergame continues with known probability δ for every additional round, and payoffs in these rounds are no longer discounted. This procedure was first introduced by [Sabater-Grande and Georgantzis \[2002\]](#) and has since also been used in [Cabral et al. \[2014\]](#) and [Vespa \[2020\]](#).

Our paper merges the literature on Prisoner’s Dilemma with the literature on other-regarding preferences ([Fehr and Schmidt \[1999\]](#), [Charness and Rabin \[2002\]](#)) to explain terminal cooperation. Just as we show the importance of Efficiency-seeking motives in FRPD games, [Engelmann and Strobel \[2004\]](#) do the same using one-shot distribution tasks. A parallel theoretical literature uses bounded rationality to explain FRPD cooperation. For example, [Radner \[1986\]](#) shows that initial cooperation can be sustained as part of an equilibrium if agents are willing to deviate slightly from best responses. [Mengel \[2014\]](#) and [Heller \[2015\]](#)³ use limited foresight to explain cooperation, the end-game effect, and the restart effect in finitely repeated PD. Similarly, [Jehiel \[2001\]](#) shows that an equilibrium concept based on limited foresight can sometimes single out cooperation in the infinitely repeated PD as a unique prediction. These theories provide an alternative explanation of the initial cooperation we observe in the current and other FRPD experiments. But, all theories of limited foresight that assume perfect foresight in the terminal period of FRPD,

³[Heller \[2015\]](#) shows that a FRPD game admits an evolutionarily stable strategy in which bounded ability agents who look one step ahead and three steps ahead coexist.

would predict defection at the terminal round, and hence cannot explain why we observe terminal cooperation increasing as δ decreases. Thus, when it comes to addressing different features of FRPD data, limited foresight is more of a complement than a perfect substitute of other-regarding preferences.

II. EXPERIMENTAL DESIGN

A total of 90 subjects participated in our experimental sessions between November 2015 to April 2017. All participants were undergraduate students at the University of British Columbia. Subjects received 6 CAD (Canadian dollar) as show-up fee. Lab currency was converted to cash payments at the exchange rate of 300 lab currency=1 CAD. Subjects, on average, were paid 12 CAD and 20 CAD for the $\delta = 1/4$ and $\delta = 3/4$ sessions respectively. The experimental instructions were read aloud while the subjects saw a written version of the same on their screen. They also received a printout of the same instructions. The subject instructions are included in the Appendix.

There were a total of 4 sessions, run in a Between design, two each for the $\delta = 3/4$ and $\delta = 1/4$ treatments. Each session, subjects were randomly divided into two groups. Each supergame, subjects from the first group were matched with a new subject from the second group using turn-pike matching⁴ [Bó, 2005]. In any experiment where pairs of players play multiple times, player A may behave differently with B to influence what B does with C next, in the hope of changing how C behaves once matched with A. Turn-pike matching rules this out, but at the disadvantage of allowing no more than n matches in a session of $2n$ subjects.⁵

The sessions took around 75 minutes to run. At the beginning of each session, we described the game, the interface, and the re-matching protocol to the subjects. Then they played eight Finitely Repeated Prisoner’s Dilemma (FRPD) supergames under the same δ treatment. Each supergame lasted for 5 periods.

⁴Subjects in both groups G1 and G2 are first enumerated as $1, \dots, n$. Then in supergame $s \leq n$, player i from G1 faces off against player $(i + s - 1) \bmod n$ from G2. Upto n supergames are run.

⁵This makes it difficult to run sessions with 20 or 30 supergames, as Embrey et al. [2018] do. More supergames allow subjects more chances at learning.

		Column Player	
		Defect	Cooperate
Row Player	Defect	$1200\delta^{t-1}, 1200\delta^{t-1}$	$2600\delta^{t-1}, 200\delta^{t-1}$
	Cooperate	$200\delta^{t-1}, 2600\delta^{t-1}$	$2000\delta^{t-1}, 2000\delta^{t-1}$

TABLE 1. Stage-game payoffs in period t for δ treatment

All δ treatments had the same period-1 payoffs. All payoffs in period- t were only δ^{t-1} times the period-1 payoffs, and this was common knowledge. Period $t \in \{1, 2, 3, 4, 5\}$ stage-game for a δ -treatment is given in Table 1. For a treatment with lower δ , payoffs declined faster after period-1. With $(3/4)^4 = .3164$ and $(1/4)^4 = .0039$, the period-5 stage-payoffs of the $\delta = 3/4$ treatment were around 80 times those of $\delta = 1/4$.

The subjects could see the payoffs for all current and future stage-games on their instruction sheet. They also saw the current stage-game on their screen as they made each decision. Subjects could see the past actions of their partners only from the current supergame of 5-rounds. Given they could not see partner's actions from previous games, there was no reputation across supergames. At the end of each session, subjects were paid their total earnings from 5 rounds of a randomly chosen supergame they had participated in. The experimental currency was converted to cash payments at the pre-announced rate. We asked the subjects four belief/ prediction questions at the start of each supergame with a new partner. The belief questions were not incentivized and thus could not influence the incentives of the FRPD play in any way. We describe the belief-related questions in Section A.

We also ran robustness sessions with 132 more subjects in a Within design with $\delta = 1, 3/4, 3/8, 1/4$. We describe the design details as well as the results in Section B in the Appendix.

III. THEORIES

The following three behavioral models support rich patterns of comparative statics across the δ -treatments:

i) **Pure Altruism:** Players also care about their partner's pecuniary payoffs. The period-utility of player i is $u_i = p_i + \alpha_i^A p_j$, where $\alpha_i^A \in [0, 1)$. p_i is the pecuniary payoff of player i .

ii) **Duty:** Players consider cooperation as the conscientious action. The period-utility of player i is $u_i = p_i + \alpha_i^D$, where $\alpha_i^D \in [0, \infty)$ if player i cooperates and $\alpha_i^D = 0$ otherwise.

iii) **Efficiency-Seeking:** Players enjoy extra utility at the efficient and equitable outcome. The period-utility of player i is $u_i = p_i + \alpha_i^E$, where $\alpha_i^E \in [0, \infty)$ for the (Cooperate, Cooperate) outcome and $\alpha_i^E = 0$ otherwise.

Each non-selfish utility function in (i)-(iii) maps the actions and payoffs to the player's utility for that period. Players start with a prior belief over opponent's supergame strategy, update those beliefs every period, and play the action that maximizes their total (pecuniary and non-pecuniary) expected utility *over all periods of the supergame*. As the fourth model, we consider the purely selfish model where the player only cares about her own payoffs and hence $u_i = p_i$. Under all four models, we allow the players to be uncertain about their partner's strategy (à la [Kreps et al. \[1982\]](#)). If she assigns a higher enough belief⁶ that her partner would play a conditionally cooperative strategy, then cooperation can be rationalized for her as long as the total (pecuniary and non-pecuniary) gains from sustaining cooperation are large enough.

We derive three classes of predictions from these theories: about initial cooperation, about terminal cooperation, and about the link between initial and terminal cooperation within a supergame. Below we intuitively describe the predictions and provide hyperlinks to the more technical analysis in the Appendix.

Initial Cooperation: In terms of pecuniary payoffs, cooperation is strictly dominated by defection in every stage-game of FRPD. Thus, initial cooperation comes at an opportunity cost in the current period, but it can provide future benefits if it helps sustain cooperation in the future periods. As δ decreases, the pecuniary benefit of sustained future cooperation relatively falls, and with it falls the net (total pecuniary and non-pecuniary) benefit of cooperating initially. This is the reason why the selfish model and the three behavioral models

⁶Beliefs are probability distributions over partner's supergame strategies.

unanimously predict that initial cooperation should decrease as δ decreases. We use and adapt the sizeBAD measure [Dal Bó and Fréchette, 2011, Blonski and Spagnolo, 2015, Embrey et al., 2018] to formally derive these predictions in Propositions 2 and 7 in Section D.⁷

Terminal Cooperation: At the terminal period of FRPD, reputation does not matter anymore, and cooperation can only be justified when its non-pecuniary benefits are larger than the pecuniary loss from paying a payoff-dominated action. Obviously, the Selfish model predicts no terminal cooperation at all, at any δ . Under the assumption that the non-pecuniary utility parameters (α_i^D , α_i^E , α_i^A) stay constant even when payoffs diminish, Duty and Efficiency-Seeking models uniquely predict that *terminal-period cooperation* would be *higher* in the low δ treatments (see Proposition 2). Under these two models, pecuniary utility dictates behavior in the early rounds when payoffs are larger, and non-pecuniary utility (α_i^D or α_i^E) dictates behavior at the terminal rounds when payoffs have diminished significantly. Thus, smaller is δ , the higher is the relative impact of non-pecuniary utility in the terminal rounds.

Under the Pure Altruism model, even if α_i^A stays constant across periods, the non-pecuniary utility $\alpha_i^A p_j$ scales proportionally to the payoffs, and hence this model does not predict the same behavior. More generally, all behavioral theories where the non-pecuniary utility scales proportionally with the pecuniary utility are inconsistent with terminal cooperation increasing as δ decreases.

Link between early and late cooperation: Early and late cooperation within a supergame are linked through two channels: First, partner’s defection in the early rounds makes it unlikely that she is a conditionally cooperative type, and thus, reduces expected reputational gains from cooperating. This is true for all four models. Second, in models of reciprocity, partner’s early defection makes players feel negatively about the partner, making cooperation less appealing.

We design our FRPD game with quickly diminishing payoffs and an aptly chosen payoff matrix to provide three additional predictions. First, only under

⁷Additionally, reputational cooperation with purely selfish players is inconsistent as it predicts no cooperation in any period of $\delta = 1/4$ treatment (Proposition 1).

the Duty model, first-period cooperators under $\delta = 1/4$ must cooperate in the last 2 periods of the supergame. Under $\delta = 1/4$, the reputational gain from potential future cooperation is insignificant, and hence first period cooperation can only be justified when it provides a large enough non-pecuniary utility. As we formally demonstrate in Proposition 4, first-period cooperation under $\delta = 1/4$ requires a Duty parameter α_i^D so large, that it makes cooperation the strictly dominant action when the payoffs have greatly diminished. Contrary to the other models, the Duty model is unique in having a non-pecuniary utility of cooperation that is independent of partner's actions, which plays a crucial role in implying unconditional terminal cooperation.

For the next two predictions, we introduce and use two other models: *Simple Reciprocity* and *Sequential Reciprocity* (Dufwenberg and Kirchsteiger [2004], Rabin [1993]). These two models provide additional insight (over and above models i-iii) into the reciprocity patterns in the data.⁸ First, under *Simple Reciprocity*, i 's reciprocal attitude towards j depends on partner j 's past defections. To fix ideas, suppose η_t is the percentage of times i 's partner has defected in periods 1, 2, .. $t - 1$. At every history in period t , player i evaluates the supergame outcomes as

$$u_{i,t}(p_i, p_j, \eta_t) = p_i + \beta(\eta_t)p_j, \quad \beta \in (-1, 1)$$

where $\frac{d\beta}{d\eta_t} \leq 0$. This nests the case of Pure Altruism under $\frac{d\beta}{d\eta_t} = 0$, but it also allows i 's altruism to dwindle and then turn to spite under $\frac{d\beta}{d\eta_t} < 0$. At period s , player i has to foresee how her current decision could change j 's future behavior at $s + 1$, which in turn would also change i 's own future preferences at $s + 2$ (through $\beta(\eta_{s+2})$) and beyond.⁹ But the analysis of i 's actions simplifies at period $t = 5$, as no future periods remain. Under this model, lower is initial cooperation, the more selfish/ spiteful subjects should become, and hence lower should be terminal cooperation (see Proposition 5).

⁸Hopefully, this also convinces the reader that we have not selectively ignored other models of reciprocity in our discussion.

⁹This can easily turn the model intractable, and this is perhaps one reason why models of social preferences [Fehr and Schmidt, 1999, Charness and Rabin, 2002] are mostly conceived for and applied to (simultaneous and sequential) one-shot games.

Finally, under *Sequential Reciprocity*, player i 's reciprocal attitude towards j depends not only on j 's past actions but also on i 's beliefs about j 's future actions. Further, instead of reciprocating directly to partner j 's actions, a Sequentially Reciprocal i reciprocates to how partner j 's actions affected i 's payoffs. This model is generally hard to refute as any action might become justifiable for particular beliefs i can hold about j 's future actions. But, in a $\delta = 1/4$ supergame, partner's defection in early rounds has exponentially higher payoff consequences than partner's defection in later rounds. For example, j 's defection in the first-period of $\delta = 1/4$, when the stage-game payoffs are largest, reduces i 's supergame payoffs drastically. There are no beliefs about j 's future actions that would redeem j after her earlier unfair action because of the sharp decline of payoffs thereafter. Hence, Sequential Reciprocity *uniquely* predicts (see Proposition 6) that after j 's first-period Defection, i has a dominant strategy to unconditionally defect thereafter, *irrespective of i 's belief about j 's behavior later in the game.*

IV. DATA PATTERNS

Before we estimate what combination of Selfish, Altruism, Duty and Efficiency-Seeking models best explains our data through a structural model, we present the main data patterns through tables and graphs, and link those to the theory-predictions from Section III. This exercise gives us a rough idea of what models might be favored by the structural analysis and why. All of the patterns are replicated for the robustness sessions data in Section B of the Appendix.

IV.1. Initial and terminal cooperation: *As δ decreases, initial cooperation decreases and terminal cooperation increases.* In Table 2, we compare the first and last-period cooperation rates. Period 1 cooperation is significantly higher in the $\delta = 3/4$ treatment than that in the $\delta = 1/4$ treatment. Similarly, Period 5 cooperation is significantly lower in the $\delta = 3/4$ treatment than that in $\delta = 1/4$ treatment. Dropping the first four supergames to allow for subject-learning does not change these results.¹⁰

¹⁰This opposite pattern in initial and terminal cooperation is inconsistent with Simple Reciprocity.

TABLE 2. Average cooperation rates (%) and standard error (in parenthesis) in Periods 1 and 5. Left panel is for the main data. Right panel is for pooled data from main and robustness sessions. Lower panel reports p-values from F test against null $H_0 : \delta_2 = \delta_4$.

	Main data				Main+Robustness data	
	Games 1-8		Games 5-8		(Pooled data)	
	P1	P5	P1	P5	P1	P5
$\delta_2 = 3/4$	40.77 (5.95)	11.90 (2.54)	39.29 (6.46)	8.33 (2.89)	34.33 (5.46)	9.33 (1.83)
$\delta_4 = 1/4$	26.82 (4.57)	23.18 (3.96)	16.67 (4.71)	19.27 (4.53)	22.99 (3.65)	20.52 (3.38)
$H_0 : \delta_2 = \delta_4$	.066	.019	.006	.045	.08	.001
Clustering	Subject level				Session level, bootstrapped	
N	90				8	

In the right panel, we pool data from both the main sessions and robustness sessions for statistical power. We cluster bootstrapped standard errors at the session level. The patterns are still statistically significant. Figure 1 describes how cooperation evolves through Periods 1 to 5.

Selfish, Duty, Altruism, and Efficiency-Seeking, all models are consistent with initial cooperation decreasing with decreasing δ (Propositions 2, 7). Only Duty and Efficiency-Seeking are consistent with final cooperation increasing with decreasing δ . Table 3 shows that the comparative statics ($\delta = 3/4$ vs $\delta = 1/4$) and the terminal cooperation persist even after subjects gain significant experience.

IV.2. Connection between first and terminal period behavior: 4.2.1: *Cooperation is systematically reciprocal throughout, even in low-payoff periods.*

In Table 4, we compare the relative frequencies of cooperation as a function of partner's last period behavior. The left panel hints why the Duty model

FIGURE 1. Percentage Cooperation across Periods 1-5

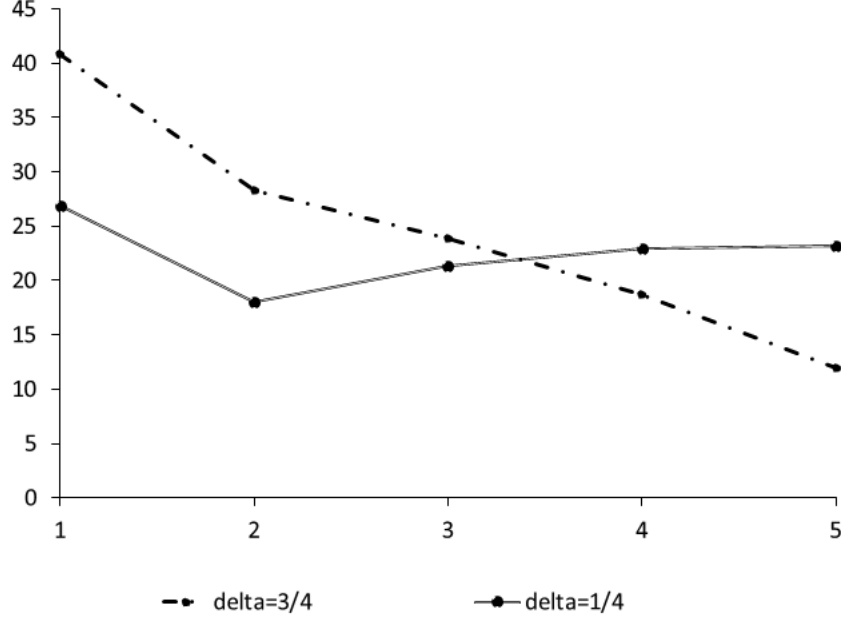


TABLE 3. Persistence with learning: Cooperation rates (%) in first and terminal periods across supergames

Supergames		1	2	3	4	5	6	7	8
Period 1	$\delta_2 = 3/4$	42.86	45.24	38.10	42.86	45.24	38.10	35.71	38.10
	$\delta_4 = 1/4$	47.92	41.67	31.25	27.08	25.00	16.67	14.58	10.41
Period 5	$\delta_2 = 3/4$	11.90	16.67	19.05	14.29	9.52	9.52	9.52	4.76
	$\delta_4 = 1/4$	33.33	33.33	27.08	14.58	16.67	20.83	14.58	25.00

might fit badly: first-period cooperators cooperate conditionally, and not unconditionally, even in low δ conditions. The right panel shows that even when imposed discounting has dissipated the pecuniary payoffs greatly in periods 4-5 under $\delta = 1/4$, subjects reciprocate their partner's last period action with a high probability: cooperation is met with cooperation and defection is met

with defection. This part of the data also suggests that even under discounted payoffs, subjects pay close attention to what their partners are playing, and pay close attention to how they respond to it. Further, given partner’s last-period behavior is almost perfectly predictive of a player’s current behavior, from an outside observer’s perspective, this data is highly systematic instead of having unexplained variability, i.e, noise.

	Periods 2-5		Periods 4-5
	First period Cooperators		All subjects
	$\delta = 1/4$	$\delta = 3/4$	$\delta = 1/4$
Partner cooperated in previous period	$\frac{158}{193}$ (83%)	$\frac{167}{240}$ (70%)	$\frac{103}{170}$ (60.59%)
Partner defected in previous period	$\frac{38}{219}$ (18%)	$\frac{28}{280}$ (9%)	$\frac{74}{598}$ (12.37%)

TABLE 4. Conditional cooperation instead of Duty: Frequency of Cooperation in Periods 2-5 by First Period Cooperators

4.2.2: *Subjects cooperate in Periods 2 – 5 of a $\delta = 1/4$ supergame even if the partner has Defected in the payoff-heavy Period 1.*

Subjects keep cooperating even after being defected-on by their partner in Period 1. 42% (138 out of the 328) of total instances of Cooperation in periods 2-5 come from subjects whose partners Defected in Period 1.¹¹ This pattern is inconsistent with the Sequential Reciprocity model that predicts after partner’s first period defection, defection becomes a dominant strategy in Periods 4-5.

To look at the same pattern differently, consider the two following hypothetical situations where the partner j has Defected for the same number of previous periods. *i ’s payoffs have been harmed exponentially more in the situation where one of those defections happened in the first period.* In Table 5,

¹¹Similarly, in the Within sessions, 61% (94 out of the 153) of total instances of Cooperation in periods 2-5 come from subjects whose partners Defected in Period 1.

$\delta = 1/4$		
Treatment	(1) Period 4	(2) Period 5
$Coop_{\tau=1}$	-2.64 (.76)	-2.42 (.79)
$PCoop_{\tau=1}$	-1.85 (.98)	-1.79 (.85)
$\sum_{\tau=1}^{\tau=t-1} Coop_{\tau}$	1.98 (.42)	1.58 (.29)
$\sum_{\tau=1}^{\tau=t-1} PCoop_{\tau}$	1.07 (.39)	.67 (.23)
Constant	-2.42 (.31)	-2.58 (.32)
Observations	384	384
Standard errors, clustered at the subject level are reported in parentheses		

TABLE 5. Logistic regression of cooperation in Period 4 or Period 5 on behavior from period $\tau = 1$ and total previous cooperation.

we run a non-causal logistic regression of Period $t \in \{4, 5\}$ cooperation on own and partner's Period 1 cooperation ($Coop_{\tau=1}$, $PCoop_{\tau=1}$ respectively), and total instances of own and partner cooperation till Period $t - 1$ ($\sum_{\tau=1}^{\tau=t-1} Coop_{\tau}$, $\sum_{\tau=1}^{\tau=t-1} PCoop_{\tau}$ respectively). The dependent variable, $Coop_t$ is binary and takes a value of 1 is the subject cooperated in period $t \in \{4, 5\}$. Sequential reciprocity predicts that the coefficient on partner's Period 1 cooperation, $PCoop_{\tau=1}$, should be positive: For a fixed $\sum_{\tau=1}^{\tau=t-1} PCoop_{\tau}$, cooperation is more likely if partner cooperated in payoff-heavy Period 1 than in a later period. Instead, the coefficients on partner's period 1 cooperation are significantly negative. The negative coefficients persist even if one removes the variables $Coop_{\tau=1}$ or $\sum_{\tau=1}^{t-1} Coop_{\tau}$ from the regression.

In Table 7, we summarize the patterns, and theories inconsistent with it. In Appendix F, we discuss how the current experiment could inform existing theory.

V. STRUCTURAL ESTIMATION

In this section, we operationalize a Finite Mixture Model to estimate what combination of Selfish (S), Duty (D), Efficiency-seeking (E) and Altruism (A) best explains the data.¹² To simplify our analysis, we assume the following:

(1) Choices: Every round t , player i matched with player j chooses action $a_{i,t}^j \in \{C, D\}$.

(2) Supergame payoffs: Player i approximates the expected utility of current actions by assuming that if both players cooperate in the current round ($a_{i,t}^j = a_{j,t}^i = C$), then they both cooperate in all future rounds $s > t$ of the supergame; if either player defects, then they both defect in all future rounds $s > t$ of the supergame.

(3) Beliefs: In the $\delta \in \{1/4, 3/4\}$ treatment, i 's belief in period $t \geq 2$ about their partner j 's Cooperation can take four values $\beta_{CC,\delta}, \beta_{CD,\delta}, \beta_{DC,\delta}, \beta_{DD,\delta} \in [0, 1]$, depending on the outcome $o \in \{CC, CD, DC, DD\}$ of the last period.¹³

(4) Initial Beliefs: For period $t = 1$, there are no previous actions, but we directly use the Period 1 belief data we had collected about partner's cooperation from that supergame (see Appendix A for details of our belief data).

Let $\Pi_{CC}^t = \sum_{s=t}^{s=5} \delta^{s-1} 2000$ be the total payoff from sustaining the (C,C) outcome from Period t to Period 5. Similarly, let $\Pi_{DD}^t = \sum_{s=t}^{s=5} \delta^{s-1} 1200$ be the same for the (D,D) outcome. Given assumptions (1)-(4), Subject i is modeled as a random utility maximizer whose expected utility in Period t from taking action $a_{i,t}^j \in \{C, D\}$ depends on her type $\tau(i) \in \{S, D, E, A\}$ and her beliefs which in turn depend on the last period's outcome $o \in \{CC, CD, DC, DD\}$.

¹²Unfortunately, the [Dufwenberg and Kirchsteiger \[2004\]](#) model is not tractable enough to use in this exercise.

¹³It would perhaps be better to also let the beliefs depend on the period t , but that would make the number of parameters to be estimated intractable.

$$\begin{aligned}
\Pi^t(a_{i,t}^j = C | \tau(i) = S, o, \delta) &= \beta_{o,\delta} \Pi_{CC}^t + (1 - \beta_{o,\delta}) \left(\underbrace{\delta^{t-1} 200 + \Pi_{DD}^{t+1}}_{(C,D) \text{ in } t, (D,D) \text{ for rest}} \right) \\
\Pi^t(a_{i,t}^j = C | \tau(i) = D, o, \delta) &= \beta_{o,\delta} \left(\Pi_{CC}^t + \alpha_D \underbrace{(6-t)}_{\text{remaining periods}} \right) \\
&\quad + (1 - \beta_{o,\delta}) (\delta^{t-1} 200 + \alpha_D + \Pi_{DD}^{t+1}) \\
\Pi^t(a_{i,t}^j = C | \tau(i) = E, o, \delta) &= \beta_{o,\delta} (\Pi_{CC}^t + \alpha_E (6-t)) + (1 - \beta_{o,\delta}) (\delta^{t-1} 200 + \Pi_{DD}^{t+1}) \\
\Pi^t(a_{i,t}^j = C | \tau(i) = A, o, \delta) &= \beta_{o,\delta} (1 + \alpha_A) \Pi_{CC}^t \\
&\quad + (1 - \beta_{o,\delta}) (\delta^{t-1} 200 + \alpha_A \delta^{t-1} 2600 + (1 + \alpha_A) \Pi_{DD}^{t+1}) \\
\Pi^t(a_{i,t}^j = D | \tau(i) \neq A, o, \delta) &= \beta_{o,\delta} (\delta^{t-1} 2600 + \Pi_{DD}^{t+1}) + (1 - \beta_{o,\delta}) \Pi_{DD}^t \\
\Pi^t(a_{i,t}^j = D | \tau(i) = A, o, \delta) &= \beta_{o,\delta} (\delta^{t-1} 2600 + \alpha_A \delta^{t-1} 200 + (1 + \alpha_A) \Pi_{DD}^{t+1}) + (1 - \beta_{o,\delta}) (1 + \alpha_A) \Pi_{DD}^t
\end{aligned}$$

where $\alpha_D, \alpha_E \geq 0$ and $\alpha_A \in [0, 1]$. Under the usual assumptions on idiosyncratic errors, this gives rise to the logistic form of probabilities,

$$\begin{aligned}
(1) \quad P(a_{i,t}^j = C | \tau, o, \delta) &= \frac{\exp(\Pi^t(a_{i,t}^j = C | \tau, o, \delta) / \lambda_{t,\delta})}{\exp(\Pi^t(a_{i,t}^j = C | \tau, o, \delta) / \lambda_{t,\delta}) + \exp(\Pi^t(a_{i,t}^j = D | \tau, o, \delta) / \lambda_{t,\delta})} \\
&= \frac{1}{1 + \exp((\Pi^t(a_{i,t}^j = D | \tau, o, \delta) - \Pi^t(a_{i,t}^j = C | \tau, o, \delta)) / \lambda_{t,\delta})}
\end{aligned}$$

where $\lambda_{t,\delta}$ is a scaling parameter that measures how well the subject best-responds to her expected utility in Round t of treatment δ .

V.1. Stochastic choice under diminishing payoffs: For any model τ , and for a fixed payoff difference

$$\Delta_\delta^t = |(\Pi^t(a_{i,t}^j = D | \tau, o, \delta) - \Pi^t(a_{i,t}^j = C | \tau, o, \delta))|$$

in equation (1) between defection and cooperation, when $\lambda_{t,\delta}$ increases, the utility maximizing choice is made less often. In the limit, as $\lambda_{t,\delta} \rightarrow \infty$, choices become uniformly random and each action is taken with 50% chance. Similarly, *if $\lambda_{t,\delta}$ is constant across rounds t and treatments δ , but Δ_δ^t decreases* due to discounting in the later rounds, choices should become uniformly random under any stochastic choice model. Equipped with this stochastic choice feature, just the selfish utility model ($\tau(i) = S$) with a constant $\lambda_{t,\delta}$ could also explain the second data-feature from Table 7: As δ decreases, Δ_δ^t decreases faster for

lower δ , and thus, under a constant $\lambda_{t,\delta}$, Defect (the selfish player's utility-maximizing choice) should be observed less often.

Does the stochastic selfish model explain all of the observed terminal cooperation? Is $\lambda_{t,\delta}$ constant across rounds t and treatments $\delta \in \{1/4, 3/4\}$? These are empirical questions that we allow the data to answer. We model $\lambda_{t,\delta}$ as

$$\lambda_{t,\delta} = \lambda_{1,\delta} \times (r_\delta)^{t-1}$$

with r_δ being the the per-period rate of change in responsiveness for treatment δ .

V.2. Data and likelihood: Each subject i 's data consists of their action from every period t and for every match j , and i 's first period belief about partner's cooperation that we had collected. To estimate the proportion of the population who abide by each utility model, we assume that each subject is drawn independently from a distribution (p_S, p_D, p_E, p_A) over the 4 utility types and maximize the logged likelihood of the observed data

$$(2) \quad LL = \sum_i \log \left(\sum_\tau \left(\prod_{t,j} \left(P(a_{i,t}^j = C | \tau, o, \delta) \times \mathbb{1}\{a_{j,t-1}^i = C\} + P(a_{i,t}^j = D | \tau, o, \delta) \times \mathbb{1}\{a_{j,t-1}^i = D\} \right) p_\tau \right) \right)$$

over the set of parameters

$$(p_S, p_D, p_E, p_A, (\beta_{\{C,D\}^2 \times \{3/4, 1/4\}}), (\lambda_{1,\delta}, r_\delta)_{\delta \in \{3/4, 1/4\}}, \alpha_D, \alpha_E, \alpha_A) \in [0, 1]^{15} \times \mathbb{R}_+^4$$

In equation 2, the inner product of probabilities is over all periods t and partners j , and the outer sums are over all utility-types τ and all players i respectively.

V.3. Results: We estimate the model with unrestricted parameters in model [1] of Table 6, then restrict the payoff-sensitivity $r_{\delta=3/4} = r_{\delta=1/4} = 1$ in model [2], and then further restrict the belief parameters $\beta_{o,\delta=3/4} = \beta_{o,\delta=1/4}$ for all outcomes o in model [3]. Model [4] is identical to Model [1], but restricts $p_D = 0$.¹⁴ The sample fit (log likelihood) reduces a lot under the constraints, but the estimates of the residual parameters do not change as much. Results from a likelihood ratio (LR) test comparing the constrained [2] (or [3] or [4]) and unconstrained [1] model give a p -value lower than 0.0001, implying that,

¹⁴The proportion of Duty players is not significantly different than zero in some specifications, hence we run a robustness check with $p_D = 0$.

by allowing the sensitivity parameter to decay or by allowing the beliefs to depend on the treatment being run or by allowing Duty preferences, the unconstrained model fits the data significantly better. The same is confirmed by AIC comparison across the 4 models.

Type proportions: The estimated proportions of Selfish, Duty, Efficiency-Seeking, and Altruistic subjects are 40%, 4%, 36% and 20% respectively in the unconstrained model [1], 47%, 2%, 45%, 6% in the constrained model [2], 49%, 1%, 36%, 14% in the constrained model [3]. Thus, under both models, the modal subject is Selfish¹⁵. Around 40% of subjects are estimated as Efficiency Seeking, making it the second most common preference type, despite the competition from stochastic choice under the Selfish model. This suggests that the cooperation observed in the data is best interpreted as that from a mixture of selfish subjects and other-regarding subjects.

Payoff-responsiveness: In model [1], we estimate $r_{\delta=3/4} = .64$ and $r_{\delta=1/4} = .26$. This can be interpreted as if the subjects became more sensitive to payoff-differences as the payoffs declined. Especially, at $\delta = 1/4$, the 95% confidence interval of r_{δ} contains $1/4$. Thus, based on the estimates from model [1], we do not find any evidence for choices becoming less sensitive to utility-differences at terminal rounds of $\delta = 1/4$. But, what if one restricts $r_{\delta=3/4} = r_{\delta=1/4} = 1$? A similar pattern emerges when we compare the sensitivity parameters across the treatments from model [2]: In the $\delta = 1/4$ treatment, where payments were smaller, the subjects were more sensitive to the utility differences (the scaling parameter is 86.79 as compared to 237.52).

Beliefs: Conditional on (D,D) being played in the last round, both models estimate the beliefs $\beta_{DD,\delta}$ about partner's current cooperation as approximately 0. Similarly, $\beta_{CC,1/4}$ was estimated as close to 1, which should be interpreted as players assign close to a belief of 1 on partner's cooperation in the next round after (C,C) outcome. Finally, if only one of the players cooperated in the last round (outcomes (D,C) or (C,D)), then the belief about partner's current cooperation is weakly higher if it was the partner who had cooperated, i.e, $\beta_{DC,\delta} \geq \beta_{CD,\delta}$.

¹⁵The high proportion of selfish subjects is unsurprising as around 30 percent of our subjects always defect, irrespective of their partner's actions.

Utility parameters: Across all specifications [1] to [3], we get similar estimates of the Efficiency-Seeking (α_i^E), Altruism (α_i^A), and Duty parameter (α_i^D). The Efficiency-Seeking parameter (α_i^E) is about half of the difference between the temptation payoff (2600) and the mutually cooperative outcome (2000) from period 1.

Using the estimated unconstrained Finite Mixture Model in [1] and based on how a subject played, we can assign each subject a posterior type-probability of her belonging to each of the four utility types. For example, a subject could be assigned ($p_S = 1, p_D = 0, p_E = 0, p_A = 0$) implying that the model is “confident” that she is a selfish player. Similarly, a subject could be assigned ($p_S = p_D = p_E = p_A = .25$) implying that the model assigns equal chance of her being of any type.

For each subject, we also calculate their total count of cooperation over 8 supergames. This number runs between 0 to 40. We use a jittered scatter plot of posterior type-probabilities (Y-axis) against this count (X-axis) for the $\delta = 1/4$ treatment, in Figure 2.¹⁶ Each subject is represented by four markers, one each for the four type-probabilities. For example, to the subject who cooperated for all 40 rounds, the model assigns a posterior probability of approximately 1 of being the Duty type (blue diamond at $p = 1, c = 40$), and probabilities of 0 of being the other three types (three markers clustered at $p = 0, c = 40$). Similarly, there were two subjects who cooperated 28 times, and were both assigned a posterior probability of 1 for being Efficiency-Seeking (two gray rectangles coincide at $p = 1, c = 28$, with the other markers clustered around $p = 0, c = 28$). Most points lie on or very close to the horizontal axes $p = 0$ and $p = 1$, indicating that it is only for a small number of subjects that FMM is incapable of detecting the best-fit model with confidence.

There is no subject who is assigned a posterior probability of more than 80% of being selfish, and who cooperates more than 3 times out of a total 40 times possible over the eight $\delta = 1/4$ supergames. When it comes to terminal cooperation, we get the same picture: selfish subjects seldom cooperate. To show this, we use a jittered scatter plot of posterior type-probabilities (Y-axis)

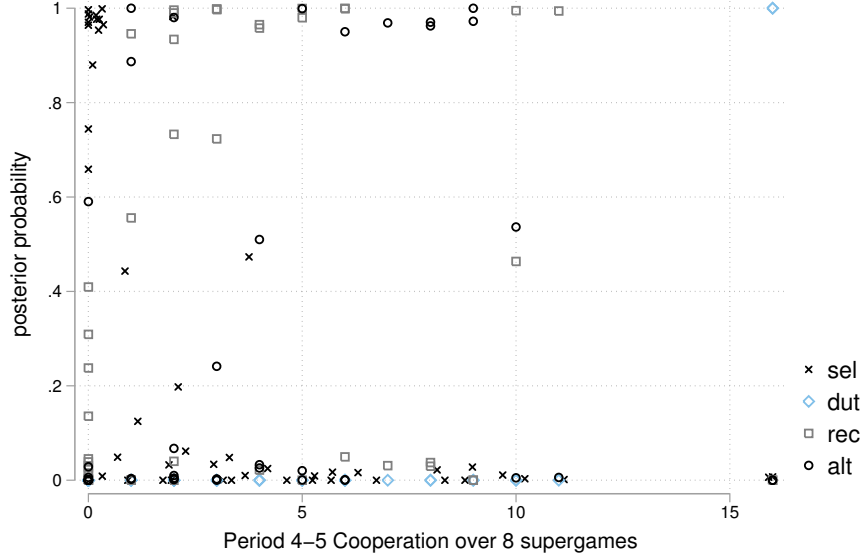
¹⁶We replicate the same figure for the $\delta = 3/4$ treatment in Figure 8 in the Appendix.

FIGURE 2. Jittered scatter plot of posterior type-probabilities against the total count of cooperation over 8 supergames. ($\delta = 1/4$ treatment) We replicate the same figure for the $\delta = 3/4$ treatment in Figure 8 in the Appendix.



against total cooperation in Periods 4 and 5 (X-axis) for the $\delta = 1/4$ treatment, in Figure 3.

FIGURE 3. Jittered scatter plot of posterior type-probabilities against the total count of cooperation in Periods 4-5 over 8 supergames. ($\delta = 1/4$ treatment)



Here is one reason why the finite mixture model identifies significant other-regarding preferences from the cooperative subjects. The stochastic selfish model can *never* explain subjects taking a payoff-dominated action (for e.g, cooperation in the terminal round) with probability greater than 50%. This makes it a bad fit for how strongly reciprocal our cooperative subjects were: Conditional on achieving the (C,C) outcome in the 4th period, $\delta = 1/4$ players cooperated 93% of the time in the 5th period. Thus, a stochastic Selfish model that can explain the moderate aggregate rate of terminal cooperation over all rounds and matches would still fall short of explaining the high within-player and within-match correlation of cooperative behavior.¹⁷

In Section C of the Appendix we estimate a nested logit model that nests the Altruism, Duty and Efficiency Seeking models in one general model. We estimate this as an unconstrained nested logit model, and compare it to three constrained models run with $\alpha_i^A = 0$, $\alpha_i^D = 0$ and $\alpha_i^E = 0$ respectively. We

¹⁷The stochastic Selfish model and Efficiency-Seeking models, parameterized with values from model [1], that is under $\beta_{CC,1/4} = 1, \lambda_{5,1/4} = 299.58 \times (.26)^4 = 1.37$, predict 15% and 100% conditional cooperation rates in Period 5 of $\delta = 1/4$ treatment respectively.

compare the fit of each of the constrained models individually to the unconstrained model using the Likelihood Ratio test. We find that removing Efficiency-Seeking or Altruism from the unconstrained model significantly decreases the fit, but removing the Duty model does not. When we compare the unconstrained general model to a constrained purely selfish model with $\alpha_i^A = \alpha_i^D = \alpha_i^E = 0$, we find that the absence of other-regarding preferences reduces the fit significantly and by a large margin.

VI. HETEROGENEITY IN BEHAVIOR

In this section we characterize subjects into the four utility types (Altruism/ Duty/ Efficiency-Seeking/ Selfish), based on their response to partner's behavior.

In Figure 5a, we summarize the behavior of subjects by their conditional response in the latter half (Rounds 3 to 5) of all 8 supergames, against partner's previous actions.¹⁸ The x and y components correspond to a player's average empirical likelihood of cooperation against previous round's cooperation and defection.

$\delta = 1/4$ polarizes cooperative response to partner's cooperation. It has a larger proportion of subjects who are very likely ($x \in [.8, 1]$) to cooperate after partner's cooperation: 20/47 (43%) of all $\delta = 1/4$ subjects as compared to only 8/40 (20%) of all $\delta = 3/4$ subjects. Conversely, it also has a slightly larger proportion of subjects who always defect ($\max\{x, y\} \leq .2$): 15/47 (32%) of all $\delta = 1/4$ subjects versus 11/40 (27.5%) of all $\delta = 3/4$ subjects.¹⁹

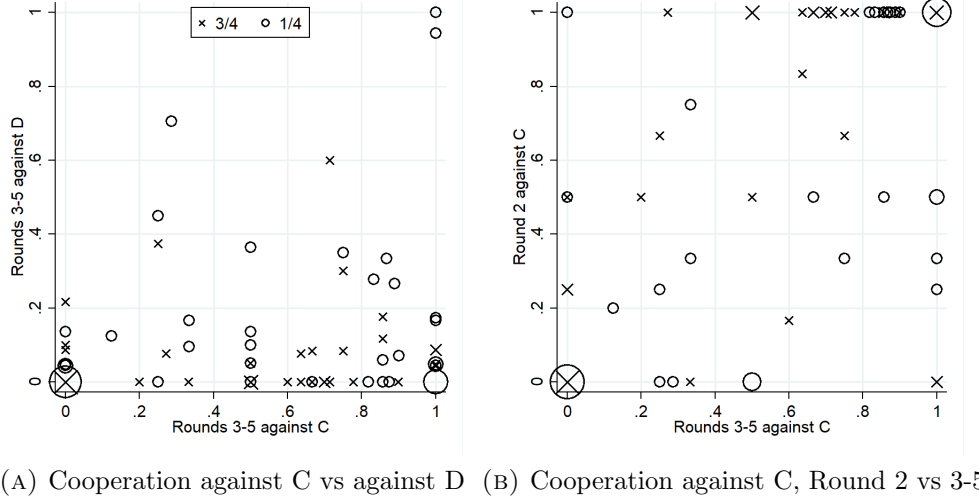
Unconditional cooperators are almost non-existent, as most subjects rarely cooperate after partner's defection in the last period. Under $\delta = 1/4$, after partner's defection, 38/47 subjects cooperate on less than 20% occasions ($y \leq .2$). Similarly, under $\delta = 3/4$, after partner's defection, 36/40 subjects cooperate on less than 20% occasions.

Proposition 4 and 6 about the $\delta = 1/4$ treatment allow us to characterize subjects into types. Type-identification is especially easier to justify for extreme

¹⁸We can define these two conditional measures for 47/48 and 40/42 subjects from the $\delta = 1/4$ and $\delta = 3/4$ sessions, whose partners have taken both actions at least once.

¹⁹This is explained by the sub-population of selfish reputational players who find it worth reciprocating only under $\delta = 3/4$.

FIGURE 4. For each subject, we calculate the average empirical frequency of cooperation against partner's last round cooperation (C) or Defection (D). On the left, we bubble-plot each subject's conditional cooperation against C vs against D. On the right, we bubble-plot how conditional cooperation against C changes from Round 2 to Rounds 3-5.



patterns. 15/47 of $\delta = 1/4$ session subjects behave selfishly ($\max\{x, y\} \leq .2$). 2/47 behave like always-cooperators ($x \approx y \approx 1$) with high Duty parameters. 16/47 other subjects, who have a highly asymmetric response to partner's cooperation ($x \geq .8, y \leq .2$) behave like Efficiency-Seekers for whom partner's last period cooperation implies high probability of partner-cooperation in the upcoming period. Rest are difficult to characterize without strong assumptions.

Many $\delta = 1/4$ subjects increase how often they cooperate against partner's cooperation from Round 2 to later rounds: see Figure (5b). Under $\delta = 3/4$, we observe the opposite pattern: subjects decrease how often they cooperate against partner's cooperation in the later rounds. Thus, Figure (5b) rejects purely Tit-for-tat behavior under which conditional behavior should not change across rounds. Figure (7-ii) in Appendix II shows that first period cooperators in $\delta = 1/4$ are no different than others in terms of conditional cooperation in rounds 3-5 (rejecting the Duty model).

VII.

For tractability, the estimation exercise has to abstract away from other important channels like bounded rationality [Radner, 1986], limited foresight [Mengel, 2014, Heller, 2015], failure of backward induction [Mantovani et al., 2014], or non-Bayesian belief-updating by FRPD subjects [Cox et al., 2015]. Thus, the utility and belief parameters should be interpreted in the usual “as if” interpretation instead of the “as is” interpretation. Still, the co-existence of a significant number of Efficiency-Seeking subjects alongside selfish subjects substantiates Andreoni and Miller [1993], Cooper et al. [1992] and fits well within the larger FRPD literature. It also explains cooperation observed in one-shot PD games: for example, Charness et al. [2016] find significant cooperation in one-shot PD games, and this cooperation increases in the mutual-cooperation payoff.

In a comprehensive study, Embrey et al. [2018] show that the parameters of a FRPD have a significant effect on initial cooperation, and this relation is captured succinctly by sizeBAD. In Section D, we modify the original sizeBAD measure to make it applicable to Efficiency-Seeking preferences and show that any variation in the FRPD parameters (for e.g, sucker’s payoff or temptation payoff) causes identical changes to the original sizeBAD for selfish subjects and the modified sizeBAD for Efficiency-Seeking subjects.²⁰ Thus, sizeBAD predicts identically for a population of fully selfish subjects or in a mixed population of selfish and Efficiency-Seeking subjects, and these predictions have already been confirmed in the Embrey et al. [2018] paper. Our treatments also provide a new test and confirmation of the predictive power of both sizeBAD measures under decreasing δ .²¹

Further, Embrey et al. [2018] find that conditional on subjects establishing cooperation, the first defection round moves earlier with experience. This

²⁰For selfish subjects, sizeBAD captures the tradeoff between the temptation to become the first defector and the potential loss from defecting too early. For Efficiency-Seeking, modified sizeBAD captures an identical tradeoff, but with more (α_i^E) to lose from defecting too early.

²¹Intuitively, sizeBAD for selfish preferences captures the tradeoff subjects face between the temptation to become the first defector and the potential loss from defecting too early. Subjects with Efficiency-Seeking preferences face an identical tradeoff, but with more (α_i^E) to lose from defecting too early, and this is what the modified sizeBAD captures.

finding is also consistent with a population of selfish and Efficiency-Seeking subjects stopping cooperation progressively earlier in the fear of getting defected on first. Efficiency-seeking is indeed highly susceptible to such fear and unraveling, as the non-pecuniary utility is only experienced as long as the partner cooperates.²² Finally, Embrey et al. [2018] estimate the decline of cooperation in the long run by fitting a learning model to their FRPD data and simulating it for 1000 supergames. For their 8-period FRPD game which has a comparatively low temptation payoff and high sucker’s payoff, the estimated decline rate is so slow that cooperation rates for the first round of the supergame would remain above 80% even after a 1000-supergame worth of experience. The presence of Efficiency-Seeking subjects provides a potential explanation for why cooperation declines this slowly.

VIII. CONCLUSION

We incorporate theories of non-selfish motives into the reputation model, binding complementary explanations together for better predictive power. In our Finite Mixture Model, we estimate that 40-49% of our subjects are selfish, 36-45% are Efficiency-seeking, 1-4% are Duty players, and 6-20% are Altruistic. We estimate Efficiency-Seeking as the modal non-selfish behavior, and consistent with the data, it divorces itself from unconditional cooperation or unforgiving punishment. As δ decreases, it correctly predicts higher terminal cooperation and lower initial cooperation.

APPENDIX A. BELIEFS

In the main/ Between sessions, the subjects were also asked to answer the following four prediction questions at the start of each supergame with a new partner:

- How likely is your partner to cooperate on the first round of the this game?
- How likely is your partner to cooperate in the very next round if you cooperated in the previous round of the game?

²²In comparison, Altruistic or Duty subjects can experience non-pecuniary utility irrespective of partner’s action.

- How likely is your partner to cooperate in the very next round if you defected in the previous round of the game?
- How likely is your partner to cooperate on the very last (5th) round of the this game?

We used neutral language in the instructions: B and R were the names assigned to Defect, and T and L were the names assigned to Cooperate. At each question the subjects could respond on a scale of 0 to 10, and they were advised to enter a higher number the more likely they thought the event was. They were also provided the following reference points:

- A response of 0 (lowest point of the scale) would mean "never".
- 5 (midway point of the scale) would mean "as likely as getting Heads on a fair coin toss/ 50-50 odds",
- 10 (right extreme of the scale) would mean "surely".
- Events more likely than "never" and less likely than heads on a fair coin toss, should be rated between 0 and 5, and so on.

Subjects comfortable in thinking of likelihoods in terms of probabilities, were told how the 0-10 scale maps into the $[0, 1]$ probability scale. The belief questions were not incentivized and thus they are less likely to influence FRPD play in any way.²³

In Figure 5, we plot the evolution of first and terminal period beliefs (about partner's cooperation) throughout supergames 1 to 8. We also provide the corresponding 95% confidence intervals. The belief data gives us a way to test some of the assumptions we used in constructing our hypotheses.

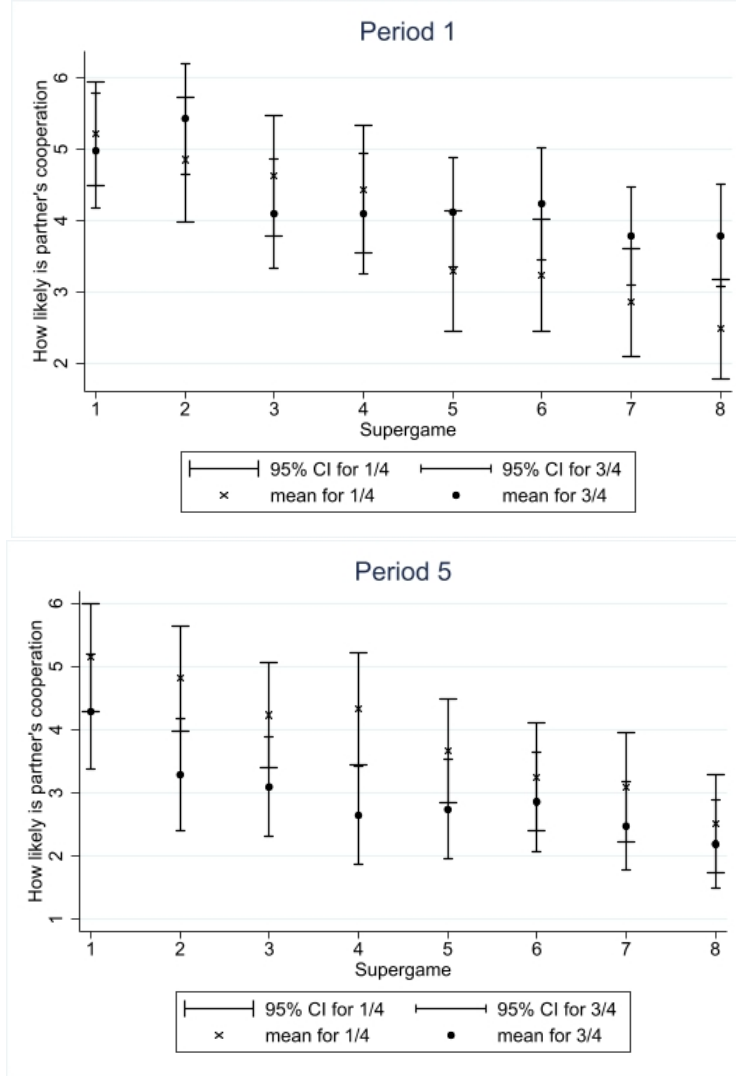
To predict initial cooperation from Propositions 2 and 2, we have assumed that Period 1 beliefs are either equally optimistic across δ or weakly more optimistic at higher δ . Initial beliefs are indeed identical at the start of the $\delta = 1/4$ and $\delta = 3/4$ session. Initial beliefs are statistically identical (for both one and two-sided t-tests) across treatments through supergames 1-4, and become more optimistic for $\delta = 3/4$ thereafter. Supergame 5 onwards, the one-sided t-test is significant at 5%.

To predict terminal cooperation from Proposition 2, we assumed that Period 5 beliefs are either equally optimistic or weakly more optimistic at lower δ .

²³One might think that the very act of asking about beliefs influences play.

Period 5 beliefs are indeed weakly more optimistic in the $\delta = 1/4$ sessions. These beliefs stay more optimistic for $\delta = 1/4$ (one-sided t-test significant at 10% or 5%) till Supergame 5, and then become statistically identical.

FIGURE 5. Reported optimism about partner's cooperation on a 0-10 scale across supergames 1-8 in the Between Sessions (for both $\delta = 1/4$ and $\delta = 3/4$)



APPENDIX B. DETAILS OF THE ROBUSTNESS SESSIONS

Design: A total of 132 subjects participated in the five robustness sessions that were run under the Within design. Matching, length and total number of

supergames, game-payoffs and payment protocol were identical to the main/ Between Treatment. No beliefs were elicited in the Within sessions. Subjects, on average, were paid 17CAD in the robustness sessions, which included a 5 CAD participation fee.

In a Within session, each subject played under all of the four treatments $\delta = 1, 3/4, 3/8, 1/4$. The two extra $\delta = 1, 3/8$ treatments²⁴ test the robustness of the empirical findings. Each session, the four treatments were repeated in two blocks, so that each subject played 2 supergames under each treatment and a total of $(4 \times 2) = 8$ supergames. The order of treatments in the Within sessions was randomized at the session level. At the beginning of each session, a coin-toss decided if the treatments would be run in a decreasing order

$$\delta = 1, 3/4, 3/8, 1/4, 1, 3/4, 3/8, 1/4$$

or in an increasing order

$$\delta = 1/4, 3/8, 3/4, 1, 1/4, 3/8, 3/4, 1$$

We planned to get around 60 subjects in each order. Overall, 3 sessions (74 subjects in total) were run in the decreasing order²⁵ and 2 sessions (58 subjects in total) in the increasing order. Below we replicate the main results from Section IV using the data from the Within Sessions.

Initial and terminal cooperation: As reported in Table 8, first period cooperation is higher in the $\delta = 3/4$ treatment than that in $\delta = 1/4$ treatment. In fact, first-period cooperation across four δ treatments is increasing in δ . Similarly, last period cooperation is significantly lower in the $\delta = 3/4$ treatment than that in $\delta = 1/4$ treatment, and approximately decreasing in δ . The tests reconfirm Hypotheses 1 and 2. In column 2, we drop the first block of data for all subjects, and redo our analysis as a robustness check. All of the pair-wise tests come out to be significant, except the comparison of Period 5 cooperation rates between $\delta = 1$ and $\delta = 3/8$ treatments.

²⁴ $\delta = 1$ was chosen as it is the standard FRPD game. $\delta = 3/8$ was chosen as it is intermediate between $1/4$ and $3/4$.

²⁵After two decreasing-order sessions, we only had 46 subjects, well short of the 60 mark. So we ran a third session, but 28 subjects showed up, driving the total up to 74.

TABLE 8. Comparison of cooperation rates in first and terminal periods of Within Sessions (Compare with Table 2 from the Between sessions)

	Blocks 1 and 2 (All Data)		Only Second Block	
	Period 1	Period 5	Period 1	Period 5
$\delta_1 = 1$	33.71 (3.55)	9.85 (1.97)	31.82 (4.08)	8.33 (2.42)
$\delta_2 = 3/4$	26.14 (3.12)	6.06 (1.70)	21.97 (3.63)	6.06 (2.09)
$\delta_3 = 3/8$	20.08 (2.84)	12.88 (2.32)	15.15 (3.14)	12.88 (2.93)
$\delta_4 = 1/4$	17.42 (2.64)	16.67 (2.45)	9.85 (2.61)	17.42 (3.32)
N	1056	1056	528	528
δ_2 vs δ_4	.0067	.0000	.0009	.0005
δ_1 vs δ_4	.0000	.01	.0000	.01
δ_1 vs δ_3	.0000	.26	.0000	.20

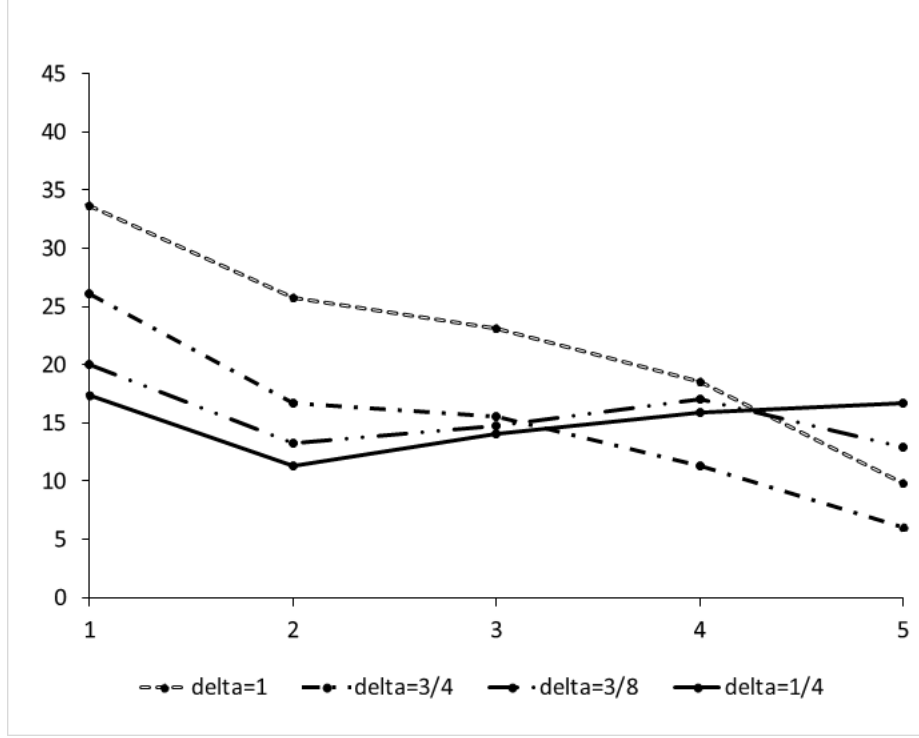
Standard errors, clustered at the level of 132 subjects are reported in parentheses below

Lower panel reports p-values from F test against $H_0 : \delta_i = \delta_j$

Cooperation is systematically reciprocal throughout even in Within sessions: In Table 9, we compare the relative frequencies of cooperation as a function of partner's last period behavior. Duty model is inconsistent with this data: first-period cooperators cooperate conditionally, and not unconditionally, even in low δ conditions. The second panel shows that reciprocal behavior continues even when imposed discounting has dissipated the pecuniary payoffs greatly in periods 4-5.

Effect of early behavior on late cooperation: In Table 10, we run a non-causal logistic regression of Period $t \in \{4, 5\}$ cooperation on own and partner's Period 1 cooperation ($Coop_{\tau=1}$, $PCoop_{\tau=1}$ respectively), and total instances of own and partner cooperation till Period $t - 1$ ($\sum_{\tau=1}^{\tau=t-1} Coop_{\tau}$, $\sum_{\tau=1}^{\tau=t-1} PCoop_{\tau}$

FIGURE 6. Percentage Cooperation across Periods 1-5 in Within sessions
(Compare with Figure 1 from the Between sessions)



respectively). Sequential reciprocity predicts that the coefficient on partner's Period 1 cooperation, $PCoop_{\tau=1}$, should be positive. The coefficients on partner's period 1 cooperation are instead negative (and potentially underpowered) for the Within sessions. The negative coefficients persist even if one removes the variables $Coop_{\tau=1}$ or $\sum_{\tau=1}^{t-1} Coop_{\tau}$ from the regression.

APPENDIX C. NESTED LOGIT MODELS

Different theories considered in this paper often make unrelated predictions. For e.g., Duty is the only theory that makes a prediction about own initial cooperation versus own later cooperation. Instead of comparing each model against the corresponding features of the data that result in its rejection, we provide a unified empirical test of fit across all the models of Altruism, Duty and Efficiency Seeking models. As a starting point, we specify a general model that nests the other-regarding utility of all three varieties. Under the assumptions used in the Finite Mixture Model estimation exercise from Section V this

model would imply:

$$\begin{aligned}\Pi^t(a_{i,t}^j = C) &= \beta_{o,t}((1 + \alpha_A)\Pi_{CC}^t + \alpha_D(6 - t) + \alpha_E(6 - t)) \\ &\quad + (1 - \beta_{o,t})(\delta^{t-1}200 + \alpha_A\delta^{t-1}2600 + \alpha_D + \Pi_{DD}^{t+1}) \\ \Pi^t(a_{i,t}^j = D) &= \beta_{o,t}(2600\delta^{t-1} + \alpha_A\delta^{t-1}200 + (1 + \alpha_A)\Pi_{DD}^{t+1}) + (1 - \beta_{o,t})(1 + \alpha_A)\Pi_{DD}^t\end{aligned}$$

We estimate this as a nested logit model in Column [1] of Table 11. Next, we also estimate models [2], [3] and [4] that are special cases of model [1], under $\alpha_E = 0$, $\alpha_A = 0$ and $\alpha_D = 0$ respectively. We compare the fit of each of the models [2], [3] and [4] individually to the unconstrained model [1], under the null hypothesis that the constrained or smaller model is the “best” model. The Likelihood Ratio (LR) test confirms that we can reject this null for all the comparisons²⁶, but for that between model [1] and [4]. Thus, the data rejects the null hypothesis that adding Duty preferences increases the model fit, but it cannot do the same for Efficiency-Seeking or Altruism. Finally, in column [5], we estimate the original model under the restriction $\alpha_E = \alpha_A = \alpha_D = 0$, imposing fully selfish preferences. This significantly reduces the fit that the model can offer, and the LR test indicates that adding other-regarding preferences significantly increases the model fit²⁷.

APPENDIX D. THEORETICAL RESULTS

We organize our theoretical results in five subsections. The first subsection discusses selfish reputational cooperation. Duty and Efficiency-seeking are pooled together in subsection D.2. We discuss a general model of reciprocity in subsection D.3. Sequential Reciprocity and Altruism follow in the subsections D.4 and D.5. All the theoretical results are summarized in Table 12 and proved in Section E. Selfish players, whose utility is identical to their payoffs, would be called *egoists*.

Whenever we mention “beliefs” in this section, we mean probability distributions over the partner’s strategy²⁸. For the theory results on initial cooperation, we use sizeBAD, which restricts each player to believe that their partner

²⁶The statistical software confirmed the p -value was lower than .0001.

²⁷The statistical software confirmed the p -value was lower than .0001.

²⁸A strategy, S_i , for a player i is a function that maps each of i ’s information sets into actions available at that information set.

is playing either Grim Trigger (GT) or Always Defect (AD). For all the other results in the paper, we do not need any restrictions on beliefs. Each player can hold arbitrary beliefs about their partner’s strategy, including beliefs like *partner either plays conditional cooperation or plays AD*, and the results still follow through. I use the solution concepts of belief-rational and dominant strategy which allows me to make predictions about actions without restricting beliefs about partner’s strategy.

D.1. Reputational Cooperation: In a standard FRPD, selfish players have to perform multiple steps of backward induction about own and others’ future actions to arrive at the subgame perfect Nash Equilibrium strategy of Always Defect. Thus, cooperation in the standard FRPD is often attributed to the failure of backwards induction, especially, given the rich evidence across games that players fail to perform backwards induction reasoning [Güth and Tietz, 1990, McKelvey and Palfrey, 1992, Binmore et al., 2002, Mantovani et al., 2014, Rampal, 2018].²⁹ In proposition 1, we show that in our $\delta = 1/4$ treatment, cooperation is ruled out not just under backwards induction, but also under the much weaker solution concept of dominant strategies.

Proposition 1. Reputational Cooperation for Egoists II: *In treatments with $\delta \leq 1/4$, the egoist has a unique weakly dominant strategy of playing Always Defect (AD).*

Here is a simple intuition for the result: In a $\delta = 0$ FRPD, AD is the dominant strategy as the FRPD essentially reduces to a one-shot PD game. By continuity, AD must stay a dominant strategy in a FRPD at δ values close enough to 0, and $1/4$ is such a value (by design).

In a $\delta = 1$ FRPD, if i believes that her partner plays a threshold m strategy³⁰, then i ’s best response would be to play a threshold $m - 1$ strategy. Learning backwards induction in FRPD is often the iterative process of understanding that both players would want to undercut the threshold duration from m to $m - 1$ to $m - 2$ and so on. Our $\delta = 1/4$ treatment was specifically

²⁹Chakraborty and Kendall [2022], Kendall and Chakraborty [2022] show that subjects also fail at backward induction when they are playing against their own selves.

³⁰Grim Trigger (GT) till period m and AD thereafter.

designed such that *it does not require such iterative reasoning* for subjects to arrive at Always Defect (AD). Rather AD is the *dominant strategy* for selfish players: Subjects should play AD irrespective of what they believe their partner plays, even if she believes that her partner is playing a threshold strategy. The same intuition extends to intermediate values of δ too: the potential for early cooperation drops uniformly with δ , as we show in Remark 2 in the Appendix.

D.2. Duty and Efficiency Seeking:

Next we derive the implications of **Duty** and **Efficiency Seeking**. Under the Duty model, subject preferences are:

$$U_i = p_i + \alpha_i^D \text{ where } \alpha_i^D \geq 0 \text{ if Cooperate is played, } \alpha_i^D = 0 \text{ otherwise}$$

And, under the Efficiency-Seeking model, subject preferences are:

$$U_i = p_i + \alpha_i^E \text{ where } \alpha_i^E \geq 0 \text{ when both players have cooperated, } \alpha_i^E = 0 \text{ otherwise}$$

Initial cooperation: When a subject is contemplating whether to play a defecting strategy (like Always Defect) or a cooperative strategy in Period 1, she needs to determine which of these two strategies is the most profitable in expectation. This depends on the parameters of the game (for example, δ) and her belief that her partner wants to play a conditionally cooperative strategy (like GT). When a high belief about partner's conditional cooperation is required to justify own cooperation, it implies that the game or situation (δ) is less conducive to initial cooperation. Based on this idea, SizeBAD assumes that:

Assumption [sizeBAD]

(A1) All players are selfish ($u_i = \pi_i$).

(A2) Any player believes that her partner plays grim trigger (GT) with probability p and plays Always defect (AD) with probability $(1 - p)$.

SizeBAD, introduced by Dal Bó and Fréchette [2011], Blonski and Spagnolo [2015] for infinitely repeated games and adapted by Embrey et al. [2018] to FRPD, is calculated as the probability p that i must assign to partner j playing Grim Trigger, so that i herself is indifferent between playing Grim Trigger (cooperative strategy) and AD (defecting strategy).

To extend sizeBAD to other-regarding references in FRPD with discounted payoffs, we modify (A1) by replacing $u_i = \pi_i$ by Duty or Efficiency-Seeking (in Proposition 2) or Altruistic utility (in Proposition 7), maintain (A2), and use the discounted payoffs in calculating the total utility from each strategy. For example, to construct sizeBAD for the Efficiency-Seeking model, we assume

Assumption 2 [Modified sizeBAD for Efficiency-Seeking]

(A1) $U_i = p_i + \alpha_i^E$ where $\alpha_i^E \geq 0$ when both players cooperate, $\alpha_i^E = 0$ otherwise.

(A2) Any player believes that her partner plays grim trigger with probability p and plays Always defect (AD) with probability $(1 - p)$.

It is easy to show that the modified sizeBAD measure for player who has Efficiency-Seeking parameter α_i^E becomes

$$(3) \quad \rho_0 = \frac{\pi_{dd} - \pi_{cd}}{(x_5(\delta)(\pi_{cc} - \pi_{dd}) - \pi_{cd} - \pi_{dc} + 2\pi_{dd} + 5\alpha_i^E)}$$

instead of

$$(4) \quad \rho_0 = \frac{\pi_{dd} - \pi_{cd}}{(5(\pi_{cc} - \pi_{dd}) - \pi_{cd} - \pi_{dc} + 2\pi_{dd})}$$

under selfish preferences and $\delta = 1$, where,

$$(5) \quad x_n(\delta) = \begin{cases} \frac{1 - \delta^n}{1 - \delta} & \text{if } \delta < 1 \\ n & \text{if } \delta = 1 \end{cases}$$

Proposition 2. Initial Cooperation: *Under the Efficiency-Seeking model (or the Duty model), for any α_i , as δ decreases, the modified sizeBAD measure increases.*

For selfish subjects, the decision to cooperate depends on how the temptation to become the first defector compares to the potential loss from defecting too early. Subjects with Efficiency-Seeking (or Duty) preferences face an identical tradeoff, but with more (α_i^E or α_i^D) to lose from defecting too early, and this is what the modified sizeBAD above captures. When δ decreases, defecting too early becomes less costly for all types of subjects. Thus, the modified sizeBAD measure predicts the same comparative statics with respect to game

parameters (for example, the sucker's payoff π_{cd} or temptation payoff π_{cd}) for Efficiency-Seeking or Duty subjects, as the standard sizeBAD measure does for selfish subjects.

Mathematically speaking, as δ decreases, x_5 in equation (5) decreases, hence, the denominator of ρ_0 in (3) decreases and hence, ρ_0 itself increases. Thus, as δ decreases, initial cooperation should decrease *for all values of the Efficiency-Seeking parameter α_i^E* .

To generate predictions about initial cooperation from Proposition 2, we consider two alternative assumptions. First, initial beliefs are equally optimistic at all treatments. Second, as initial cooperation is more demanding at lower δ , subjects are (weakly) less optimistic at lower δ . Under Proposition 2, both assumptions imply lower cooperation in lower δ treatments. We plot average initial beliefs at every supergame in Figure 5. Our belief data empirically justifies less optimism at lower δ .

Terminal cooperation: To compare terminal cooperation, we do not assume any restriction on beliefs. Rather in Propositions 3, 4 and 6, we allow the players to hold any beliefs about their partner's strategy. We use the solution concepts of *belief-rational* or *dominant-strategy*, which allow for arbitrary beliefs.³¹

In particular, for a fixed player, we compare across δ treatments the set of beliefs (through set-inclusion) under which terminal cooperation becomes a best response. First, we define *belief-rational*, a condition that simply states that such a set is non-empty at period t for some player-type.

Definition 1. Cooperation is *belief-rational* in period t for Duty-type α_i^D (or Efficiency-type α_i^E) if there exists some period t belief about partner's strategy in the ensuing supergame, against which the best-response for player i involves cooperating in period t .

Cooperation could be belief-rational in any period of any treatment, as long as, a Duty or Efficiency-Seeking player i has a large enough α_i parameter. It is easiest to see this for the terminal periods because there are no future

³¹Among other things, this also implies that to predict terminal cooperation for i we do not need to assume that i believes her partner j has some particular utility type (Duty or Efficiency-Seeking for example).

reputational implications. As long as α_i is larger than $\min\{\pi_{dd} - \pi_{cd}, \pi_{dc} - \pi_{cc}\} = \min\{1000\delta^4, 600\delta^4\} = 600\delta^4$, terminal cooperation is belief-rational: for sufficiently large belief on the partner cooperating, cooperation is a best-response. Further, if α_i is larger than $\max\{\pi_{dd} - \pi_{cd}, \pi_{dc} - \pi_{cc}\} = 1000\delta^4$, terminal cooperation is strictly dominant under the Duty Model

Proposition 3. *Terminal cooperation:* *Let $\delta_h, \delta_l \in \{1, 3/4, 3/8, 1/4\}$ be two treatments with $\delta_h > \delta_l$.*

- i) Duty model: For any fixed α_i^D , if terminal-period cooperation is belief-rational at treatment δ_h , then it is a strictly dominant strategy at treatment δ_l . Further, there exists a range of duty parameters α_i^D for which final period cooperation is only belief-rational in δ_l but not in δ_h .*
- ii) Efficiency-Seeking Model: For any fixed α_i^E , if terminal-period cooperation is belief-rational at treatment δ_h , then so it is at treatment δ_l . Further, there exists a range of parameters α_i^E for which final period cooperation is only belief-rational in δ_l but not in δ_h .*

Our terminal-period predictions from Proposition 3, (i) and (ii), follow from how we have spaced out terminal-period payoffs through $\delta_h, \delta_l \in \{1, 3/4, 3/8, 1/4\}$ with $\delta_h > \delta_l$. The proof of (i) follows in three simple steps. The belief-rationality at δ_h implies that α_i must compensate for at least $600\delta_h^4$. By design, $600\delta_h^4$ is greater than $1000\delta_l^4$ for any pair of $\delta_h > \delta_l$. Finally, $\alpha_i^D > 1000\delta_l^4$ implies that cooperation is the weakly dominant action at δ_l . (ii) follows similarly. Results (i) and (ii) imply higher terminal cooperation under lower δ when terminal beliefs are equally optimistic at all treatments or (weakly) more optimistic at lower δ . Our belief data on terminal round optimism justifies this assumption (see Figure 5).

Remark 1. If the α_i -parameter was proportional to the stakes, the Duty or Efficiency-Seeking models would no longer predict the increase in terminal cooperation for small δ . The stake-independent α_i is crucial to cooperative forces dominating behavior in terminal rounds of low δ treatments.

Example 1 in Appendix E presents a numerical equilibrium-analysis for a population of Efficiency-Seeking players. Next, Proposition 4 states how first-period cooperation is linked to terminal period cooperation under the Duty Model.

Proposition 4. *Duty Model:* *Take a player i who has Duty model preferences with α_i^D . If Cooperation is belief-rational for i in the first period of the $\delta = 1/4$ Treatment, then Cooperation must be a dominant strategy for i in the last two periods of the $\delta = 1/4$ Treatment.*

Let S_c be the set of all strategies that cooperate in the first period. It is easy to see that the strategy of always cooperating (AC) would belong to this set S_c . Suppose, we fix i 's first-period belief b about her partner's strategy. For this fixed belief b , there must exist a Duty parameter $\alpha_i(b)$ such that some strategy $s \in S_c$ becomes the best response to b . For any b , such a value always exists as at $\alpha_i^D > 1000$, AC $\in S_c$ becomes a dominant strategy. We calculate a loose lower bound $\alpha_i^{\min}(b)$ on this $\alpha_i(b)$ by requiring that at least one strategy from S_c gives a higher expected payoff than playing AD. In the proof, we show that even $\inf_b \alpha_i^{\min}(b)$ is so large, that it would make Cooperation a dominant strategy in every subgame consisting of the last two periods.

Here is the intuition for why this result holds: Reputation alone cannot justify initial cooperation under $\delta = 1/4$, as there is little to be gained from the future. Cooperating in period 1 can only be rationalized under a large α_i^D , even under arbitrarily optimistic beliefs about partner's reciprocating behavior. Even the lowest α_i^D that makes initial cooperation belief-rational, would dominate the discounted pecuniary losses $1000\delta^3$ from cooperation by the fourth period.

D.3. Simple Reciprocity. In the Altruism, Duty and Efficiency-seeking models, subjects have the same other-regarding utility parameter α_i throughout the supergame. Here we consider an alternative model, called *Simple Reciprocity*, where partner j 's past defections determine and reduce how kindly i feels about j in period t . Suppose η_t is the percentage of times i 's partner has defected in periods $1, 2, \dots, t-1$. At every history in period t , player i evaluates

the supergame outcomes under the utility function

$$(6) \quad u_{i,t}(p_i, p_j, \eta_t) = p_i + \beta(\eta_t)p_j, \quad \beta \in (-1, 1), \quad \frac{d\beta}{d\eta_t} \leq 0$$

This nests the case of Pure Altruism under $\frac{d\beta}{d\eta_t} = 0$, but under $\frac{d\beta}{d\eta_t} < 0$, it allows i 's altruism to dwindle and then turn to spite. This model is not easily tractable: At period s , i has to foresee how her current decision could change j 's behavior at $s + 1$, which in turn would also change i 's own preferences at $s + 2$ and beyond. But it simplifies at period $t = 5$, as no future rounds remain.

Proposition 5. *Simple Reciprocity:* *Take two players i_1 and i_2 who have identical preferences and identical $\frac{d\beta}{d\eta_t}$ as defined in equation (6). Suppose i_1 participates in a δ_h treatment and i_2 participates in a δ_l treatment where $\delta_h > \delta_l$. Further let i_1 and i_2 's experience from Periods 1 to 4 be summarized by η_5^h and η_5^l where $\eta_5^h \leq \eta_5^l$. Then, if cooperation is belief-rational for i_2 , then Cooperation must also be belief-rational for i_1 .*

If δ_l treatment has higher η_5 , then players should be relatively more kind to the partner in the terminal round of the δ_h treatment. Thus, the δ_h treatment should be more conducive to terminal cooperation.

D.4. Sequential Reciprocity: We first describe the Sequential Reciprocity model. For player i , we use S_i as the set of all her strategies and π_i to denote her supergame payoffs. Fix a player i at history h of the game. Suppose i is playing strategy $s_i \in S_i$ ³², believes that her partner j would play the strategy $s_{ij} \in S_j$, and, believes that j believes that i would play $s_{iji} \in S_i$. s_{ij} is i 's first order belief, and s_{iji} is i 's second-order belief. At history h , assume that the choices that define (or lead to) history h have been made with probability 1 under these strategies.

Let $E(S_x)$ be the set of efficient strategies for $x \in i, j$. A strategy belongs in this efficient set for $x \in i, j$, if there does not exist an alternative strategy by x that improves *everyone's payoffs, for every strategy played by other players, at all histories of play* with strict inequality for some (player-other's strategy-history).

³² S_i and S_j are the set of all possible strategies for i and j .

For a fixed s_{iji} , i considers the maximum payoff

$$\Pi_i^{max}(s_{iji}) = \max_{s_j \in S_j} \pi_i(s_{iji}, s_j)$$

and the minimum payoff

$$\Pi_i^{min}(s_{iji}) = \min_{s_j \in E(S_j)} \pi_i(s_{iji}, s_j)$$

that j could give her through her actions $s_j \in S_j$. Then, for her second-order belief s_{iji} , i calculates the equitable outcome $\Pi_i^e(s_{iji})$ as the average of $\Pi_i^{max}(s_{iji})$ and $\Pi_i^{min}(s_{iji})$. Given $\Pi_i^e(s_{iji})$, i judges j 's *kindness to i* as

$$\lambda_{iji}(s_{ij}, s_{iji}) = \pi_i(s_{iji}, s_{ij}) - \Pi_i^e(s_{iji})$$

Thus, kindness $\lambda_{iji}(\cdot)$ could be positive or negative, depending on if the perceived outcome $\pi_i(s_{iji}, s_{ij})$ is greater or smaller than the equitable outcome $\Pi_i^e(s_{iji})$. Similarly, given her first order belief s_{ij} , i judges her own kindness towards j at any action s_i as

$$\begin{aligned} k_{ij}(s_i, s_{ij}) &= \pi_j(s_i, s_{ij}) - \Pi_j^e(s_{ij}) \\ &= \pi_j(s_i, s_{ij}) - \frac{\max_{s_i \in S_i} \pi_j(s_i, s_{ij}) + \min_{s_i \in S_i} \{\pi_j(s_i, s_{ij}) : s_i \in E(S_i)\}}{2} \end{aligned}$$

Just like λ_{iji} , the kindness term k_{ij} can also be both positive and negative.

Finally, i 's utility is

$$U_i(s_i, s_{ij}, s_{iji}) = \pi_i(s_i, s_{ij}) + \alpha_i k_{ij}(s_i, s_{ij}) \lambda_{iji}(s_{ij}, s_{iji})$$

which is the sum of the monetary payoff and a reciprocity payoff that is the product of the two kindness factors. Positive and negative kindness has to be reciprocated with like to maintain a positive reciprocity payoff. Beliefs are updated using Bayes rule along the game.

Sequential Reciprocity, like other models of psychological games, is difficult to test as it involves subjects' first and higher-order beliefs which are unobservable. We do, however derive a testable implication of Sequential Reciprocity model that holds irrespective of subjects' first and second-order beliefs.

Proposition 6. *Sequential Reciprocity:* *If partner j has Defected in Period 1 of the $\delta = 1/4$ supergame, Defecting at all subsequent histories through periods 2-5 becomes a strictly best response for player i . This holds irrespective of α_i*

and irrespective of the first and second-order beliefs player i has about her partner j at every subsequent history.

The payoff gains from partner's cooperation are largest in Period 1 of the $\delta = 1/4$ treatment. A first-period Defect by partner j reduces i 's potential payoffs drastically. Payoffs decline exponentially thereafter and thus no future actions can push i 's total income above the equitable outcome. Thus, i regards j 's first-period Defection as an irreversibly unkind act, and maximizing her own reciprocity payoff now requires doing everything to reduce partner's payoff, which requires i to Defect throughout. Maximizing pecuniary payoffs also requires i to Defect throughout (see Proposition 1). Thus, Player 1 should Defect unconditionally hereon, irrespective of the relative weight she assigns to pecuniary and reciprocity payoffs. This logic goes through irrespective of first and second-order beliefs the player has for the rest of the supergame.

D.5. Pure Altruism: The preferences of any player under this model is given by

$$U_i = p_i + \alpha_i^A p_j, \quad 0 \leq \alpha_i^A < 1$$

Playing Defect is a weak best response in any stage game of any δ treatment as long as

$$1200 + 1200\alpha_i^A \geq 200 + 2600\alpha_i^A \quad \text{and} \quad 2600 + 200\alpha_i^A \geq 2000 + 2000\alpha_i^A$$

which gives $\alpha_i^A \leq \frac{1}{3}$. Similarly, for $\alpha_i^A \geq \frac{5}{7}$ players, playing Cooperation is a weakly best response in any stage game of any δ treatment. When two $\alpha_i^A \in (\frac{1}{3}, \frac{5}{7})$ types play each other, both (D,D) and (C,C) are equilibria of the stage-game, the latter being the payoff dominant outcome.

For $\alpha_i^A \geq \frac{5}{7}$ players, cooperating is a strictly dominant strategy of the stage game, and hence there is no value-risk tradeoff of cooperation! This also means that sizeBAD is an irrelevant statistic for them.

Proposition 7. *For the payoff parameters in our paper, sizeBAD is not defined for $\alpha_i \geq \frac{5}{7}$ players of the Pure Altruism model. For $\alpha_i < \frac{5}{7}$, modified sizeBAD decreases as δ increases.*

For $\alpha_i < \frac{5}{7}$ players, modified sizeBAD is given by

$$\rho_0 = \frac{(1 + \alpha_i^A)\pi_{dd} - (\pi_{cd} + \alpha_i^A\pi_{dc})}{((1 + \alpha_i^A)(\pi_{cc} - \pi_{dd})x_5 - (1 + \alpha_i^A)(\pi_{cd} + \pi_{dc}) + 2\pi_{dd}(1 + \alpha_i^A))}$$

which decreases as δ increases, at all values of α_i^A . We are going to assume the following condition on coordination behavior of any two $\alpha_i^A \in (\frac{1}{3}, \frac{5}{7})$ players:

Assumption 3: *Consider any two $\alpha_i^A \in (\frac{1}{3}, \frac{5}{7})$ players matched together in an FRPD supergame. They are no less likely to coordinate on the (C,C) outcome in a high- δ FRPD supergame than in a low- δ FRPD supergame.*

Comparisons of Payoff and Risk Dominance support this assumption. Let U^t be the Pure Altruism utility from period t stage-game payoffs.

- i) Payoff Dominance: $U^t(C, C) - U^t(D, D)$, the utility-gain from successful coordination on (C,C) vs that from (D,D) at period $t > 1$, is strictly increasing in δ . Thus, payoff-dominance is more salient at high δ treatments.
- ii) Risk Dominance: The value of α_i that solves the indifference condition of mixed strategies

$$U^t(C, \alpha_i^A C + (1 - \alpha_i)D) - U^t(D, \alpha_i^A C + (1 - \alpha_i)D) = 0$$

is independent of t and δ . The “risk factor” (from the risk-dominance literature) for playing (C,C) vs (D,D) for any player is the same at all periods, of all treatments. Thus risk dominance does not select stage-game equilibria differently across δ treatments.

Proposition 8. Pure Altruism: *Under Assumption 3 about coordination by $\alpha_i^A \in (\frac{1}{3}, \frac{5}{7})$ players, cooperation should either remain unchanged or decrease in all periods, as δ decreases.*

APPENDIX E. PROOFS

Prediction	Kreps et al. [1982] (Selfish)	Pure Altruism	Duty	Efficiency Seeking	DK (2004)
As δ increases, initial cooperation increases	✓	✓	✓	✓	
As δ decreases, terminal cooperation increases	X	X	✓	✓	
Initial cooperation in $\delta = 1/4$ guarantees later cooperation			✓		
Unconditional defection if partner defected initially in $\delta = 1/4$					X

TABLE 12. ✓ = Theory makes this prediction, X= Theory makes the opposite prediction. A blank box means the theory does not make a related prediction.

Proof. [Proposition 1] We need to show that for any strategy S_{-i} the partner plays, AD provides a weakly higher profit than any other strategy S_i . Thus, it must be that

$$\begin{aligned} \Pi_i(AD, S_{-i}) &\geq \Pi_i(S_i, S_{-i}) \quad \forall S_i, \forall S_{-i} \\ \iff \min_{S_{-i}} \min_{S_i \in \mathcal{S}_i} (\Pi_i(AD, S_{-i}) - \Pi_i(S_i, S_{-i})) &\geq 0 \end{aligned}$$

Fix some opponent strategy S_{-i} and let $\mathcal{H}$ be the histories reached by S_{-i} and AD. Now consider any deviation strategy S_i for i . Let A be the set of all deviation strategies that play defect at all histories in $\mathcal{H}$. As $S_i^a \in A$ and AD play identically at all histories reached,

$$(\Pi_i(AD, S_{-i}) - \Pi_i(S_i^a, S_{-i})) = 0 \quad \forall S_i^a \in A$$

For the more interesting case, consider some $S_i^b \in \mathcal{S}_i \setminus A$. By definition, S_i^b must cooperate at some history $h \in \mathcal{H}$ in period $t + 1$. Without loss of generality, let h also be the earliest history in $\mathcal{H}$ where S_i^b cooperates. Thus, S_i and AD

must have accumulated the same payoffs at all earlier histories, but accumulate difference payoffs period $t + 1$ (and history h) onwards.

If S_{-i} cooperates at h , then, AD and S_i^b obtain $2600\delta^t$ and $2000\delta^t$ from period $t + 1$ respectively. Thus,

$$\begin{aligned}
\min_{S_i^b} (\Pi_i(AD, S_{-i}) - \Pi_i(S_i^b, S_{-i})) &\geq \min_{S_i^b} \Pi_i(AD, S_{-i}) - \max_{S_i^b} \Pi_i(S_i^b, S_{-i}) \\
&\geq \left(2600\delta^t + \sum_{s=t+1}^4 1200\delta^s \right) - \left(2000\delta^t + \sum_{s=t+1}^4 2600\delta^s \right) \quad \forall t \\
&= \left(600\delta^t - \sum_{s=t+1}^4 1400\delta^s \right) \quad \forall t \\
&\geq \left(600\delta^t - \sum_{s=t+1}^{\infty} 1400\delta^s \right) \quad \forall t \\
&= \delta^t \left(600 - \delta \frac{1400}{1-\delta} \right) \quad \forall t \\
&= \frac{1}{4^t} \left(600 - \frac{1400}{3} \right) \quad \forall t \\
&\geq 0 \quad \forall t
\end{aligned}$$

To bound the expressions of the minimum and maximum, we have simply picked the minimum and maximum stage game payoffs possible in each following period, and thus *our bounds are robust to how S_{-i} and S_i^b play in the future periods.*

Similarly, if S_{-i} defects at h , then, AD and S_i^b obtain $1200\delta^t$ and $200\delta^t$ from

the from period $t + 1$ respectively. Thus,

$$\begin{aligned}
\min_{S_i^b \in \mathcal{S}_i \setminus A} (\Pi_i(AD, S_{-i}) - \Pi_i(S_i, S_{-i})) &\geq \min_{S_i^b \in \mathcal{S}_i \setminus A} \Pi_i(AD, S_{-i}) - \max_{S_i^b \in \mathcal{S}_i \setminus A} \Pi_i(S_i, S_{-i}) \\
&= \left(1200\delta^t + \sum_{s=t+1}^4 1200\delta^s \right) - \left(200\delta^t + \sum_{s=t+1}^4 2600\delta^s \right) \\
&= \left(1000\delta^t - \sum_{s=t+1}^4 1400\delta^s \right) \\
&\geq \left(600\delta^t - \sum_{s=t+1}^{\infty} 1400\delta^s \right) \quad \forall t \\
&\geq 0 \quad \forall t
\end{aligned}$$

Thus, irrespective of what S_{-i} plays at h ,

$$\min_{S_i \in \mathcal{S}_i} (\Pi_i(AD, S_{-i}) - \Pi_i(S_i, S_{-i})) \geq 0$$

Given we showed this for an arbitrary S_{-i} , this statement must hold for all S_{-i} . Thus,

$$\min_{S_{-i}} \min_{S_i \in \mathcal{S}_i} (\Pi_i(AD, S_{-i}) - \Pi_i(S_i, S_{-i})) \geq 0$$

and AD must be a weakly dominant strategy. The proof of the uniqueness of AD as a weakly dominant strategy follows similarly, and is hence omitted. $\square$

Remark 2. SizeBAD: For egoists, the sizeBAD measure is not defined for $\delta \leq .51$. In the range $\delta \in [.52, 1]$, sizeBAD increases as δ decreases.

Proof. Before we get to the proof, here is the main intuition for this result: For selfish subjects, the decision to cooperate depends on how the temptation to become the first defector compares to the potential loss from defecting too early. When δ decreases, the potential loss from defecting too early falls, as payoffs from future rounds are heavily discounted. This makes the environment less conducive to initial cooperation.

First we modify *sizeBAD* to fit our FRPD games. Following [Embrey et al. \[2018\]](#), we assume that an egoistic player (P1) playing a $T = 5$ period FRPD in δ treatment is deciding between playing Grim Trigger and Always Defect. She believes that her partner (P2) too is limited to playing these two strategies.

Let ρ_0 be the belief on P2 playing Grim Trigger that makes P1 indifferent between playing Grim Trigger and Always Defect.

Let

$$x_n(\delta) = \begin{cases} \frac{1 - \delta^n}{1 - \delta} & \text{if } \delta < 1 \\ n & \text{if } \delta = 1 \end{cases}$$

Thus, $x_T = 1 + \delta x_{T-1}$ and $\delta x_{T-1} = x_T - 1$.

P1's indifference implies

$$\begin{aligned} \rho_0(\pi_{cc}x_T) + (1 - \rho_0)(\pi_{cd} + \pi_{dd}\delta x_{T-1}) &= \rho_0(\pi_{dc} + \pi_{dd}\delta x_{T-1}) + (1 - \rho_0)(\pi_{dd}x_T) \\ \iff \rho_0(\pi_{cc}x_T) + (1 - \rho_0)(\pi_{cd} + \pi_{dd}(x_T - 1)) &= \rho_0(\pi_{dc} + \pi_{dd}\delta x_{T-1}) + (1 - \rho_0)(\pi_{dd}x_T) \\ \iff \rho_0((\pi_{cc} - \pi_{dd})x_T + \pi_{dd} - \pi_{cd}) + \pi_{cd} + \pi_{dd}\delta x_{T-1} &= \rho_0(\pi_{dc} + \pi_{dd}\delta x_{T-1} - \pi_{dd}x_T) + \pi_{dd}x_T \\ \iff \rho_0((\pi_{cc} - \pi_{dd})x_T + \pi_{dd} - \pi_{cd}) + \pi_{cd} + \pi_{dd}\delta x_{T-1} &= \rho_0(\pi_{dc} - \pi_{dd}) + \pi_{dd}x_T \\ \iff \rho_0((\pi_{cc} - \pi_{dd})x_T - \pi_{cd} - \pi_{dc} + 2\pi_{dd}) &= \pi_{dd} - \pi_{cd} \\ (7) \iff \rho_0 &= \frac{\pi_{dd} - \pi_{cd}}{((\pi_{cc} - \pi_{dd})x_T - \pi_{cd} - \pi_{dc} + 2\pi_{dd})} \end{aligned}$$

Given the stage game parameters, the denominator turns zero just below $\delta = .52$, and hence sizeBAD is undefined thence. As δ increases, the denominator increases and the sizeBAD measure decreases. $\square$

Proof. [Proposition 2] We use the same notation as in the proof of Remark 2. Game length, $T = 5$. Let

$$x_n(\delta) = \begin{cases} \frac{1 - \delta^n}{1 - \delta} & \text{if } \delta < 1 \\ n & \text{if } \delta = 1 \end{cases}$$

Under the Efficiency-Seeking model, P1's indifference in the first-period implies

$$\begin{aligned} \rho_0(\pi_{cc}x_T + 5\alpha_i^E) + (1 - \rho_0)(\pi_{cd} + \pi_{dd}\delta x_{T-1}) &= \rho_0(\pi_{dc} + \pi_{dd}\delta x_{T-1}) + (1 - \rho_0)(\pi_{dd}x_T) \\ \iff \rho_0 &= \frac{\pi_{dd} - \pi_{cd}}{((\pi_{cc} - \pi_{dd})x_T - \pi_{cd} - \pi_{dc} + 2\pi_{dd} + 5\alpha_i^E)} \end{aligned}$$

For any α_i^E -player, as δ decreases, x_T and the denominator decreases, thus predicting a high ρ_0 and declining cooperation. The result for the Duty model

follows similarly, with the caveat that for very large Duty parameters, cooperation would become a dominant strategy of every stage-game, and sizeBAD would become irrelevant. $\square$

Proof. [Proposition 3] (i) Last period cooperation is never reputational and always altruistic or warm-glow in nature. In the final period if a subject believes that the other player is going to cooperate with probability p , then she would find it worth cooperating as long as

$$\begin{aligned}\delta^4 [p(2000) + (1-p)(200)] + \alpha_i^A &\geq \delta^4 [p(2600) + (1-p)(1200)] \\ \alpha_i^A &\geq \delta^4 [p(2600 - 2000) + (1-p)(1200 - 200)] \\ &= 600\delta^4 p + 1000\delta^4 (1-p)\end{aligned}$$

The RHS is bounded above by $1000\delta^4$ and bounded below by $600\delta^4$. The next set of inequalities show that the lower bound for any δ_h is larger than the upper bound for δ_l .

$$\begin{aligned}600 \times (1)^4 &> 1000 \times (3/4)^4 \\ 600 \times (3/4)^4 &> 1000 \times (3/8)^4 \\ 600 \times (3/8)^4 &> 1000 \times (1/4)^4\end{aligned}$$

Thus, if altruistic cooperation in the final period is belief-rational for any player in a high δ treatment, say $\delta = 1$, then $\alpha_i^A \geq 600 \times 1^4$ and hence, $\alpha_i^A \geq 1000 \times (3/4)^4$, and hence it must be a dominant strategy for her at $\delta = 3/4$, and by extension at all lower δ treatments. For values of α_i^A that lie in the intervals defined by any of the strict inequalities, final-period cooperation could only be a rational response in the corresponding lower δ treatment. For example, with $\alpha_i^A \in (1000 \times (3/4)^4, 600)$, “final period cooperation” could only be a rational response in the $\delta = 3/4$ treatment, but not in the $\delta = 1$ treatment. As long as there are subjects in the experimental population with α_i^A in the relevant range, one would expect to see strictly higher last period cooperation as δ decreases.

(ii) The proof is very similar to (i) and hence omitted. $\square$

Example 1. Efficiency-Seeking Equilibrium: Suppose nature assigns Reciprocation Cooperation parameter α_i of 450, 70 and 0 with probability .2, .2, .6 respectively. Let us call the players with these α s as Types 1, 2, and 3 respectively. Cooperation is strictly dominated in the final period for all Types in the $\delta = 1$ treatment. But terminal period cooperation is reasonable in the $\delta \in \{3/8, 1/4\}$ treatments.

For example, in the $\delta = 3/8$ treatment, there is a Perfect Bayesian Equilibrium with the following strategies.

- Type 1 cooperates in period 1 and conditionally cooperates (plays Tit for Tat) thereafter. Every other Type defects in both the first and second periods.
- In the third period, Type 2 reveals itself by cooperating only if partner was Type 1, i.e, had cooperated in period 1. Type 2 conditionally cooperates thereafter. Type 3 never cooperates.
- Cooperation rates in the 5 periods in this equilibrium would be 20%, 04%, 08%, 12%, 12% respectively.

Proof. [Proposition 4] Let the preferences of the player under consideration be $U_i = p_i + \alpha_i$.

We start by deriving an $\alpha_i^{\min}$ below which first-period cooperation cannot be belief-rational: any player cooperating in the first period must have an α_i at least as large as $\alpha_i^{\min}$. We then show that for $\alpha_i^A \geq \alpha_i^{\min}$, cooperation in the last two periods at all histories is a strictly best response.

For any strategy S_c where i cooperates in the first period, and for any strategy S_{-i} played by the partner, i defects between 1 to 4 times between periods 2-5. The maximum expected payoff from any extensive-form strategy S_c is no more than the maximum of the four following quantities, where in Π_{max}^i she defects

i times in periods 2-5.

$$\begin{aligned}\Pi_{\max}^1 &= \rho_0 200 + (1 - \rho_0) 2000 + 863.28 + \alpha_i \\ \Pi_{\max}^2 &= \rho_0 200 + (1 - \rho_0) 2000 + 860.94 + 2\alpha_i \\ \Pi_{\max}^3 &= \rho_0 200 + (1 - \rho_0) 2000 + 851.56 + 3\alpha_i \\ \Pi_{\max}^4 &= \rho_0 200 + (1 - \rho_0) 2000 + 814.06 + 4\alpha_i\end{aligned}$$

These maximal bounds on $\Pi_{\max}^i$ are derived by allowing the player to enjoy the (Defect, Cooperate) payoff whenever she Defects, and, the (Cooperate, Cooperate) payoff with α_i^A whenever she Cooperates. For first-period cooperation to be belief rational, it must outperform all possible strategies, including Always Defect. Always Defect provides a payoff no less than

$$\Pi = \rho_0 1200 + (1 - \rho_0) 2600 + 398.43$$

with this lower bound achieved by assuming the (Defect, Defect) outcome throughout. Hence, for initial Cooperation to be a better response than Always Defect, one must have $\max_i \Pi_{\max}^i \geq \Pi$. Each inequality $\Pi_{\max}^i \geq \Pi$ implies a lower bound on α_i . For example, the first inequality is

$$\begin{aligned}\Pi_{\max}^1 &\geq \Pi \\ \iff \rho_0 200 + (1 - \rho_0) 2000 + 863.28 + \alpha_i^A &\geq \rho_0 1200 + (1 - \rho_0) 2600 + 398.43 \\ \iff \alpha_i^A &\geq 135.16\end{aligned}$$

We have used the most optimistic and the most pessimistic pecuniary payoffs in the LHS ($\Pi_{\max}^1$) and RHS (Π) respectively. This makes sure that $\Pi_{\max}^1 \geq \Pi$ is impossible at any value of α_i^A that is lower than 135.16.

The next three inequalities result in $\alpha_i^A \geq 68.75$, $\alpha_i^A \geq 48.96$, and $\alpha_i^A \geq 46.09$. For first-period cooperation to be belief-rational, one requires $\max_i \Pi_{\max}^i \geq \Pi$, and hence atleast $\alpha_i^A > \alpha_i^{\min} = 46$.

The payoff parameters from the last three periods of the $\delta = 1/4$ supergame are given in Table 13. Cooperation and Defect are written as C and D respectively. As $46 > \max\{\pi_{dd} - \pi_{cd}, \pi_{dc} - \pi_{cc}\} = \max\{4.69 - .78, 10.16 - 7.81\}$,

a player would with $\alpha_i^A \geq 46$ would find it a dominant strategy to Cooperate in the last period. Cooperating in fourth-period risks losing at most $\max\{40.63 - 31.25, 18.75 - 3.13\} = 15.62$ on stage-game payoff and at most $(\pi_{cc} - \pi_{cd}) = (7.81 - .78) = 7.03$ in the last period payoff (she already knows that she will play Cooperate in the terminal period). The non-pecuniary utility $\alpha_i^A > 46$ gained through cooperation in period 4 more than makes up for the maximum total pecuniary loss $(15.62 + 7.03)$. Thus, she would find Cooperate to be a dominant strategy also in the fourth period.

Such a player would cooperate in the last two periods, at all histories, and for all beliefs that she might have how her actions might affect her opponent's play. $\square$

Proof. [Proposition 5] Let $\beta_1 = \beta(\eta_5^h)$ and $\beta_2 = \beta(\eta_5^l)$. As $\eta_5^h \leq \eta_5$, we have $\beta_1 \geq \beta_2$. Belief rationality at δ_l requires

$$\begin{aligned} & \min\{((\delta_l^4 200 + \beta_2 \delta_l^4 2600) - (\delta_l^4 1200 + \beta_2 \delta_l^4 1200)), \\ & ((\delta_l^4 2000 + \beta_2 \delta_l^4 2000) - (\delta_l^4 2600 + \beta_2 \delta_l^4 200))\} \geq 0 \end{aligned}$$

$$\iff \delta_l^4 \min\{(\beta_2 1400 - 1000), (\beta_2 1800 - 600)\} \geq 0$$

$$\iff \delta_h^4 \min\{(\beta_2 1400 - 1000), (\beta_2 1800 - 600)\} \geq 0$$

$$\iff \delta_h^4 \min\{(\beta_1 1400 - 1000), (\beta_1 1800 - 600)\} \geq 0$$

$$\begin{aligned} & \min\{((\delta_h^4 200 + \beta_1 \delta_h^4 2600) - (\delta_h^4 1200 + \beta_1 \delta_h^4 1200)), \\ & ((\delta_h^4 2000 + \beta_1 \delta_h^4 2000) - (\delta_h^4 2600 + \beta_1 \delta_h^4 200))\} \geq 0 \end{aligned}$$

Thus, belief rationality of terminal cooperation at δ_l implies the belief rationality of terminal cooperation at δ_h . $\square$

Sequential Reciprocity:

Proof. [Proposition 6] We will show that subjects consider partner's first-period defection as irreversibly unkind under [Dufwenberg and Kirchsteiger \[2004\]](#), irrespective of her first and second-order beliefs.

Take any history where j has defected on i in period 1. Suppose i cooperated in period 1. By consistency of beliefs and actions, i thinks that j believes that i is playing an extensive form strategy s_{iji} that involves playing C in the first period. The case where s_{121} involves Player 1 playing D initially follows

similarly.

Let i 's belief about j 's strategy be s_j . Let S_{AC} (and S_{AD}) be the strategies of j where j always cooperates (and always defects, respectively) at every history. For any s_{iji} , the maximum payoff that j can allow i , $\max_{s_j} \pi_i(s_{iji}, s_j)$, is no less than $\pi_i(s_{iji}, S_{AC})$. Thus,

$$\max_{s_j} \pi_i(s_{iji}, s_j) \geq \pi_i(s_{iji}, S_{AC})$$

Thus, for any fixed s_{iji} , $\max_{s_j} \pi_i(s_{iji}, s_j)$ is bounded below by

$$\begin{aligned} \Pi_1 &= \inf_{s_{iji}} \max_{s_2} \pi_i(s_{iji}, s_j) \geq \inf_{s_{iji}} \pi_i(s_{iji}, S_{AC}) \\ &= 2664.03 \end{aligned}$$

(The infimum is evaluated when s_{iji} is a strategy of cooperation at every history. 2664.03 is equal to the sum of (C,C) outcomes across five periods of $\delta = 1/4$. See Table 13.)

Similarly, for any s_{121} , $\min_{s_j \in E(S_j)} \pi_i(s_{iji}, s_j)$, is no less than

$$\begin{aligned} \Pi_2 &= \inf_{s_{iji}} \min_{s_j \in E(S_j)} \pi_i(s_{iji}, s_j) \geq \inf_{s_{121}} \pi_i(s_{iji}, S_{AD}) \\ &= 266.403 \end{aligned}$$

(The infimum is evaluated by summing the (C,D) payoff across five periods of $\delta = 1/4$. We get an extreme lower bound, as Player 1's payoffs can't get any lower. See Table 13. We have used the fact that $AD \in E(S_j)$.)

The outcome considered as kind by Player 1 under [Dufwenberg and Kirchsteiger \[2004\]](#) is,

$$\Pi_i^e(s_{iji}) = \frac{\max_{s_2} \pi_i(s_{iji}, s_j) + \min_{s_j \in E(S_j)} \pi_i(s_{iji}, s_j)}{2}$$

This is no less than

$$\Pi^* = \frac{(\Pi_1 + \Pi_2)}{2} = \frac{(2664.03 + 266.403)}{2} = 1465.234$$

The maximum Player 1 could get if Player 2 has already Defected in the first period is $\Pi_m = 1063.281$. This generous upper bound is calculated as

$$200 + 2600 \sum_{s=1}^4 (1/4)^s$$

by assuming she enjoys the (C,D) outcome for rounds 2-5, and thus upper bound is still below Π^* . Thus, P2's first-period defect guarantees P1 an unfair outcome, and hence, must be considered irredeemably unkind by Player 1.

Player 1's final utility is the sum of pecuniary payoff and α_i -times reciprocity payoff. Maximizing the reciprocity payoff requires her to act as unkindly as possible to Player 1. Playing unkindly is equivalent to playing Defect throughout. Maximizing pecuniary payoffs also requires her to Defect throughout (see Proposition 1). Thus, Player 1 should Defect unconditionally hereon, irrespective of α_i .

The proof for the case where s_{121} involves a second order believe about Player 1 playing D in the first period of the Quarter treatment is very similar, with Π_1 , Π_2 , Π^* and Π_m equal to 3264, 1266.41, 2265.23 and 2063.28 respectively. $\square$

Proof. [Proposition 7] Following the steps in Proposition 2, one could show that under Pure Altruism, sizeBAD is given by

$$\rho_0 = \frac{(1 + \alpha_i)\pi_{dd} - (\pi_{cd} + \alpha_i\pi_{dc})}{((1 + \alpha_i)(\pi_{cc} - \pi_{dd})x_5 - (1 + \alpha_i)(\pi_{cd} + \pi_{dc}) + 2\pi_{dd}(1 + \alpha_i))}$$

where

$$x_n(\delta) = \begin{cases} \frac{1 - \delta^n}{1 - \delta} & \text{if } \delta < 1 \\ n & \text{if } \delta = 1 \end{cases}$$

For the range of parameters where $\rho_0 \in (0, 1)$, increasing δ increases ρ_0 . In our paper, $\pi_{dd} = 1200, \pi_{cd} = 200, \pi_{dc} = 2600$.

For α_i close to 1, ρ_0 is negative for all δ . Intuitively, for subjects with α_i close to 1, cooperating is a strictly dominant strategy of the stage game, and hence there is no value-risk tradeoff of cooperation!

Same holds for the payoffs in Embrey et al. [2018]: $\pi_{dd} = 39, \pi_{cd} = 5, \pi_{dc} = 87$ for the Difficult PD. $\pi_{dd} = 39, \pi_{cd} = 22, \pi_{dc} = 63$ for the Difficult PD.

[Proposition 8] We shorten Pure Altruism to PA for the following text. In an FRPD supergame with a lower δ , the PA-players with $\alpha_i \leq \frac{1}{3}$ would play reputational cooperation less often as gains from future-cooperation are lower. The $\alpha_i \in (\frac{1}{3}, \frac{5}{7})$ players, who try to coordinate on (C,C) or (D,D) would also play cooperation less often, as matched reputational players Defect more often under a low δ and due to Assumption 3. The $\alpha_i \geq \frac{5}{7}$ type plays Cooperation as a dominant strategy. Their behavior does not change with δ .

Thus, we can conclude that as δ decreases, cooperation across periods should either remain the same (in absence of $\alpha_i < \frac{1}{3}$ PA-players), or strictly decrease. These conclusions hold for all periods of the supergame. $\square$

APPENDIX F. EXTENDING THE EXISTING MODELS OF BEHAVIOR

Equipped with the hindsight of the experimental findings (Table 7), we discuss how the other theories could be altered to make them fit better. The Pure Altruism model is already consistent with the reciprocity behavior in the data: $\alpha_i \in (\frac{1}{3}, \frac{5}{7})$ players try to coordinate on (C,C) or (D,D). It is also consistent with higher initial cooperation under higher δ . But we could not think of any version of it which would also be consistent with increasing cooperation in the terminal periods.

The kindness calculations under Dufwenberg and Kirchsteiger [2004] can be extended to allow for only short recall periods³³. This would better fit the “fast to forgive” reciprocal behavior observed in the current paper (Table 4) and some other papers (Fudenberg et al. [2012]).

The *sizeBAD* notion could also be extended to better predict our cooperation-data. First, we allow it to be defined for any m -th period of a n -period game. Second, we retrofit the Efficiency-Seeking utility model into it. One would get

$$\rho_m^{RecCoop} = \frac{\pi_{dd} - \pi_{cd}}{((\pi_{cc} - \pi_{dd})(1 + \delta.. + \delta^{n-m}) + \delta^{-(m-1)}(n - m)\alpha_{coop} - \pi_{cd} - \pi_{dc} + 2\pi_{dd})}$$

³³For example, if every period kindness is re-calculated from the payoff consequences that took place in the last two period’s actions only.

At $m = 1$, the second term in the denominator is independent of δ . Thus, for initial cooperation, an increase in δ increases the denominator, making cooperation easier. As m increases, the second term in the denominator would eventually dominate the effect of m in the denominator (due to its multiplicative exponential nature). Lower the δ , higher the denominator, and higher would be cooperation. This could indeed predict the two central traits of our data: lower δ favors higher terminal cooperation and lower initial cooperation.

APPENDIX G. APPENDIX VI: EXPERIMENTAL INSTRUCTIONS (FOR ONLINE APPENDIX)

The following are the experimental instructions, first for Between($\delta = 1/4$ and $\delta = 3/4$) and then for Within treatments.

Instructions

WELCOME

This experiment is a study of economic decision making. The instructions are simple. If you follow them carefully and make good decisions you may earn a considerable amount of money. The money you earn will be paid to you, in cash, at the end of the experiment. **Do not discuss your choices or payoffs with any other player!**

- **Description of payments:**

In all your decisions, you will make a choice between two possible actions Top Row (T) and Bottom Row (B). The other player will also make a similar choice- he/ she would choose between Left Column (L) and Right Column (R). You have to make the decision in each round without knowing what your partner's choice, and your partner makes her decision without knowing yours. Your (and your partner's) payments will depend on the combination of your action and your counterpart's simultaneous choices. These payment possibilities will be represented in a MATRIX like the one below.

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Notice that the cells contain a pair of numbers. In each cell, the first number(which is always in **bold**) shows *your payment* for this round, and the second number shows *the payment of the other player*. Your and the other player's choice determines which of the cell payments are activated.

Summarizing this MATRIX:

You select T and the other selects L, you each make 2000.

You select T and the other selects R, you make 200 while the other makes 2600.

You select B and the other selects L, you make 2600 while the other makes 200.

You select B and the other selects R, you each make 1200.

All payments are written in lab currency. During the payment stage, total earnings in lab currency will be converted to cash payments at the exchange rate **300 lab currency=\$1**, and approximated to the closest multiple of \$1.

Description of a game:

You will be playing 8 games with 8 different partners. At the end of the experiment, one of those 8 games would be randomly chosen and your payment would be earnings from that 5 rounds of that game+\$6 participation fee. Each game in turn will have 5 rounds of decisions, and start with the same payoff MATRIX. As the rounds progress within a game, the payments (i.e, payment MATRIX) from your actions would change. **In each subsequent round, the payments would shrink to a quarter of their previous value, for all pair of actions by you and your partner.** To see in more detail, the first round payoff MATRIX *for all games* would be

Round 1:

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Then, the second round payoff MATRIX would have all the possible payments reduced to a quarter ($\frac{1}{4}$) of their value, i.e, the matrix would look like

Round 2:

		Other player	
		L	R
You	T	500 , 500	50 , 650
	B	650 , 50	300 , 300

Take a moment to see how (500, 500) from the top-left cell in second round, is exactly a quarter of (2000, 2000) from the top-left cell in first round. So, the same pair of actions from the first round yield only a quarter of the value to both you and your partner now. You should be able to guess at this point, that the payment MATRIX in the next (third) round will look like

Round 3:

		Other player	
		L	R
You	T	125 , 125	12.5 , 162.5
	B	162.5 , 12.5	75 , 75

This would continue similarly till the 5th or last round (as described in the next page).

Round 4:

		Other player	
		L	R
You	T	31.25 , 31.25	3.13 , 40.63
	B	40.63 , 3.13	18.75 , 18.75

Round 5:

		Other player	
		L	R
You	T	7.81 , 7.81	.78 , 10.16
	B	10.16 , .78	4.69 , 4.69

At the start of each setting you are matched with a new player, and you play all the 5 rounds of that setting with that same player. **You can see all your plays and results from all previous rounds and games. For example, in Round 3 of Game 6, you can see your full history from Games 1-5, and Rounds 1-2 from Game 6 itself.**

Thus you can only see what your current partner chose in all the previous rounds of the current game she is playing with you, and the same holds for your partner. So, Round 2 onwards, both you and your partner can make your decisions in a game based how the other person has previously played in the previous rounds of that same game.

After the 5 rounds in a game end, you never face the same player again. You move on to the next game where you have a new partner. **Your new partner does not get any information about your play in previous games, and you do not get to know about her history from previous games she played with someone else.**

Prediction tasks:

Before each game, you would be asked to make 4 predictions about your new partner in the ensuing game. You have never played with this person before, hence, this is merely a prediction, and we would like you to give it your best shot. **These predictions are confidential, and will never be disclosed to any other participant.** You would be asked the following questions:

- How likely is your partner to play L on the first round of the this game?
- How likely is your partner to play L in the very next round if you played T in the previous round of the game?
- How likely is your partner to play L in the very next round if you played B in the previous round of the game?
- How likely is your partner to play L on the very last (5th) round of the this game?

At each question you are asked to guess how likely something is, and you are asked to respond on a scale of 0 to 10. You can answer with any number between 0 and 10 on each question, as you see fit. The more likely you think the event is, the higher the number you should enter.

For reference, we ask you to use the following rule:

- *A response of 0 (lowest point of the scale) would mean "never".*
- *5 (midway point of the scale) would mean "as likely as getting Heads on a fair coin toss/ 50-50 odds",*
- *10 (right extreme of the scale) would mean "surely".*
- *Events more likely than "never" and less likely than heads on a fair coin toss, should be rated between 0 and 5, and so on.*
- *Of course, you might want to use intermediate numbers as you see fit.*

Of course, you might get to see how close your predictions were to the reality when you play the game.

Read this paragraph only if you prefer to think of likelihoods in terms of probabilities. In that case, you can think of the prediction task in the following way: if you think that the likelihood has a probability x , you should the integer closest to $10x$. So, you can think of the scale 0 – 10 as the probability scale of 0 – 1 in steps of .1. For example, if you think the probability that your partner plays L on the first round with probability .3, your response should be 3 on the 0 – 10 scale for that question.

We will start by giving you a short tutorial about how the GUI works, and then we will go right away to the Predictions and Games 1-8. Best of luck.

Instructions

WELCOME

This experiment is a study of economic decision making. The instructions are simple. If you follow them carefully and make good decisions you may earn a considerable amount of money. The money you earn will be paid to you, in cash, at the end of the experiment. **Do not discuss your choices or payoffs with any other player!**

- **Description of payments:**

In all your decisions, you will make a choice between two possible actions Top Row (T) and Bottom Row (B). The other player will also make a similar choice- he/ she would choose between Left Column (L) and Right Column (R). You have to make the decision in each round without knowing what your partner's choice, and your partner makes her decision without knowing yours. Your (and your partner's) payments will depend on the combination of your action and your counterpart's simultaneous choices. These payment possibilities will be represented in a MATRIX like the one below.

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Notice that the cells contain a pair of numbers. In each cell, the first number(which is always in **bold**) shows *your payment* for this round, and the second number shows *the payment of the other player*. Your and the other player's choice determines which of the cell payments are activated.

Summarizing this MATRIX:

You select T and the other selects L, you each make 2000.

You select T and the other selects R, you make 200 while the other makes 2600.

You select B and the other selects L, you make 2600 while the other makes 200.

You select B and the other selects R, you each make 1200.

All payments are written in lab currency. During the payment stage, total earnings in lab currency will be converted to cash payments at the exchange rate **300 lab currency=\$1**, and approximated to the closest multiple of \$1.

Description of a game:

You will be playing 8 games with 8 different partners. At the end of the experiment, one of those 8 games would be randomly chosen and your payment would be earnings from that 5 rounds of that game+\$6 participation fee. Each game will have 5 rounds of decisions, and start with the same payoff MATRIX. As the rounds progress within a game, the payments (i.e, payment MATRIX) from your actions would change. **In each subsequent round, the payments would shrink to three quarters ($\frac{3}{4}$) of their previous value, for all pair of actions by you and your partner.** To see in more detail, the first round payoff MATRIX *for all games* would be

Round 1:

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Then, the second round payoff MATRIX would have all the possible payments reduced to a quarter of their value, i.e, the matrix would look like

Round 2:

		Other player	
		L	R
You	T	1500 , 1500	150 , 1950
	B	1950 , 150	900 , 900

Take a moment to see how (1500, 1500) from the top-left cell in second round, is exactly a three quarters of (2000, 2000) from the top-left cell in first round. So, the same pair of actions from the first round yield only three quarters of the value to both you and your partner now. You should be able to guess at this point, that the payment MATRIX in the next (third) round will look like

Round 3:

		Other player	
		L	R
You	T	1125, 1125	112.5 , 1462.5
	B	1462.5 , 112.5	675 , 675

This would continue similarly till the 5th or last round (as described in the next page).

Round 4:

		Other player	
		L	R
You	T	843.75 , 843.75	84.38 , 1096.88
	B	1096.88 , 84.38	506.25 , 506.25

Round 5:

		Other player	
		L	R
You	T	632.81 , 632.81	63.28 , 822.66
	B	822.66 , 63.28	379.69 , 379.69

At the start of each setting you are matched with a new player, and you play all the 5 rounds of that setting with that same player. **You can see all your plays and results from all previous rounds and games. For example, in Round 3 of Game 6, you can see your full history from Games 1-5, and Rounds 1-2 from Game 6 itself.**

Thus you can only see what your current partner chose in all the previous rounds of the current game she is playing with you, and the same holds for your partner. So, Round 2 onwards, both you and your partner can make your decisions in a game based how the other person has previously played in the previous rounds of that same game.

After the 5 rounds in a game end, you never face the same player again. You move on to the next game where you have a new partner. **Your new partner does not get any information about your play in previous games, and you do not get to know about her history from previous games she played with someone else.**

Prediction tasks:

Before each game, you would be asked to make 4 predictions about your new partner in the ensuing game. You have never played with this person before, hence, this is merely a prediction, and we would like you to give it your best shot. **These predictions are confidential, and will never be disclosed to any other participant.** You would be asked the following questions:

- How likely is your partner to play L on the first round of the this game?
- How likely is your partner to play L in the very next round if you played T in the previous round of the game?
- How likely is your partner to play L in the very next round if you played B in the previous round of the game?
- How likely is your partner to play L on the very last (5th) round of the this game?

At each question you are asked to guess how likely something is, and you are asked to respond on a scale of 0 to 10. You can answer with any number between 0 and 10 on each question, as you see fit. The more likely you think the event is, the higher the number you should enter.

For reference, we ask you to use the following rule:

- *A response of 0 (lowest point of the scale) would mean "never".*
- *5 (midway point of the scale) would mean "as likely as getting Heads on a fair coin toss/ 50-50 odds",*
- *10 (right extreme of the scale) would mean "surely".*
- *Events more likely than "never" and less likely than heads on a fair coin toss, should be rated between 0 and 5, and so on.*
- *Of course, you might want to use intermediate numbers as you see fit.*

Of course, you might get to see how close your predictions were to the reality when you play the game.

Read this paragraph only if you prefer to think of likelihoods in terms of probabilities. In that case, you can think of the prediction task in the following way: if you think that the likelihood has a probability x , you should the integer closest to $10x$. So, you can think of the scale 0 – 10 as the probability scale of 0 – 1 in steps of .1. For example, if you think the probability that your partner plays L on the first round with probability .3, your response should be 3 on the 0 – 10 scale for that question.

We will start by giving you a short tutorial about how the GUI works, and then we will go right away to the Predictions and Games 1-8. Best of luck.

Subjects' Instructions

WELCOME

This experiment is a study of economic decision making. The instructions are simple. If you follow them carefully and make good decisions you may earn a considerable amount of money. You should read the following instructions carefully, as they would help you perform successfully in the game.

In this experiment you will be playing under 4 different Settings: 1) One, 2) Three-quarters, 3) Three-over-eight, 4) Two-over-eight.

For each Setting you will be paired with one other player through a computer network to play a game. You will be playing 2 games under each setting, thus a total of 8 games. At the end of the experiment, one of the 8 games will be randomly chosen, and the total earnings in 5 rounds of that game will be converted from lab currency to cash payments at the exchange rate 300 lab currency=\$1, and approximated to the closest multiple of \$1. Your total earnings would be calculated as (show-up fee of \$5+your earnings from that particular game) and paid to you immediately after the experiment. Under each game, you and the other player will be playing 5 Rounds. Thus in 8 games you will playing a total of $(8 \times 5) = 40$ Rounds. Remember, after each game you are matched with a different opponent, hence, you play each of these 8 games with 8 different opponents.

Your decisions and payoffs are confidential. Do not discuss your choices or payoffs with any other player!

- **Description of Rounds:**

As said before, you will be playing 6 rounds under each Setting. Every round you will make a choice between two possible actions Top Row(T) and Bottom Row(B). The other player will also make a similar choice each round- he/ she chooses between Left Column(L) and Right Column(R). Your payments will depend on the combination of your action and your counterpart's action. These payment possibilities will be represented in a MATRIX like the one below.

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Notice that the cells contain a pair of numbers. In each cell, the first number(which is always in **bold**) shows your payment for this round, and the second number shows the payment of the other player. Your and the other player's choice determines which of the cell payments are activated. If you choose **T (top row)** and the other player chooses **R (right column)**, the **top right cell** (located at the intersection of top row and right column) is activated.

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Hence, **your payment** for this round is 200, while the **other player's payment** for this round is 2600.

Similarly, if you choose **B (bottom row)** and the other player chooses **R (right column)**, the **bottom right cell** is activated. Hence, **your payment** for this round is 1200, and the **other player's payment** for this round is also 1200.

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

All payments here are in lab currencies.

- **Description of Settings:**

In this experiment you will be playing under 4 different Settings: 1) **Same**, 2) **Three-quarters**, 3) **Three-over-eight**, 4) **Two-over-eight**. The name of the particular setting, suggests how the payments diminish after each round. For example, the **Three-quarters** Setting means that the payments get diminished to **Three-quarters** its value each subsequent round. Similarly the **Two-over-eight** setting means that the payments get diminished to **Two-over-eight (or a quarter)** times its value each subsequent round. The setting **Same** is the only exception, as it would have exactly identical payoffs in all the rounds.

To see in more detail, under the “**Three-quarters**” setting, the first round payoff MATRIX looks like

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Then, the second round payoff MATRIX would have all the possible payments halved, i.e, the matrix would look like

		Other player	
		L	R
You	T	1500 , 1500	150, 1950
	B	1950 , 150	900 , 900

Take a moment to see how (1500, 1500) from the top-left cell in second round, is exactly $3/4$ times of (2000, 2000) from the top-left cell in first round. You should be able to guess at this point, that the payment MATRIX in the next (third) round will look like

		Other player	
		L	R
You	T	1125 , 1125	112.5 , 1462.5
	B	1462.5 , 112.5	675 , 675

This would continue similarly till the 5th or last round.

At the start of each setting you are matched with a new player, and you play all the 5 rounds of that setting with that same player. After the 5 rounds in a setting end, you never face the same player again. You move on to the next setting where you are matched with a new player.

Setting Same:

All 5 rounds are played with the following payment scheme:

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Setting Three Quarters:

Round 1:

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Round 2

		Other player	
		L	R
You	T	1500 , 1500	150, 1950
	B	1950 , 150	900 , 900

Round 3:

		Other player	
		L	R
You	T	1125 , 1125	112.5 , 1462.5
	B	1462.5 , 112.5	675 , 675

Round 4:

		Other player	
		L	R
You	T	843.75 , 843.75	84.38 , 1096.88
	B	1096.88 , 84.38	506.25 , 506.25

Round 5:

		Other player	
		L	R
You	T	632.81 , 632.81	63.28 , 822.66
	B	822.66 , 63.28	379.69 , 379.69

Setting Three-over-eight:

Round 1:

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Round 2

		Other player	
		L	R
You	T	750 , 750	75 , 975
	B	975 , 75	450 , 450

Round 3:

		Other player	
		L	R
You	T	281.25 , 281.25	28.13 , 365.63
	B	365.63 , 28.13	168.75 , 168.75

Round 4:

		Other player	
		L	R
You	T	105.47 , 105.47	10.55 , 137.10
	B	137.10 , 10.55	63.28 , 63.28

Round 5:

		Other player	
		L	R
You	T	39.55, 39.55	3.96 , 51.42
	B	51.42 , 3.96	23.73 , 23.73

Setting Two-over-eight:

Round 1:

		Other player	
		L	R
You	T	2000 , 2000	200 , 2600
	B	2600 , 200	1200 , 1200

Round 2

		Other player	
		L	R
You	T	500 , 500	50 , 650
	B	650 , 50	300 , 300

Round 3:

		Other player	
		L	R
You	T	125 , 125	12.50 , 162.50
	B	162.50 , 12.50	75 , 75

Round 4:

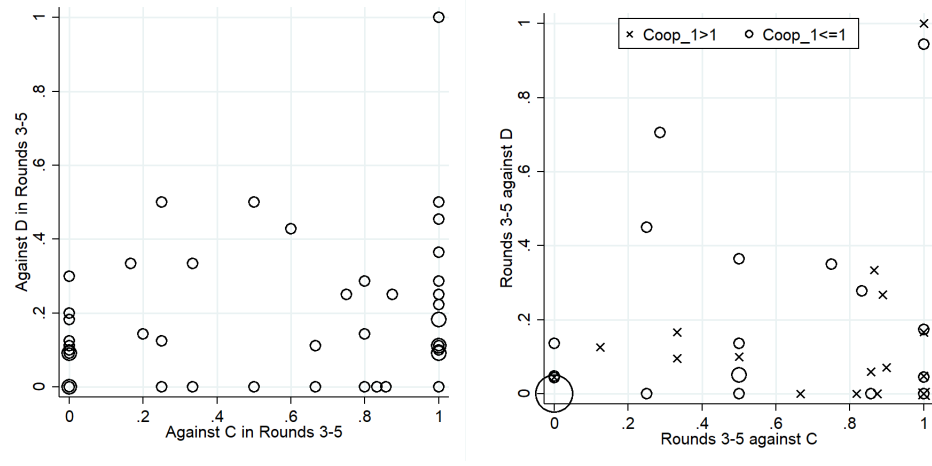
		Other player	
		L	R
You	T	31.25 , 31.25	3.13 , 40.63
	B	40.63 , 3.13	18.75 , 18.75

Round 5:

		Other player	
		L	R
You	T	7.81 , 7.81	.78 , 10.16
	B	10.16 , .78	4.69 , 4.69

APPENDIX H. ADDITIONAL GRAPHS AND TABLES

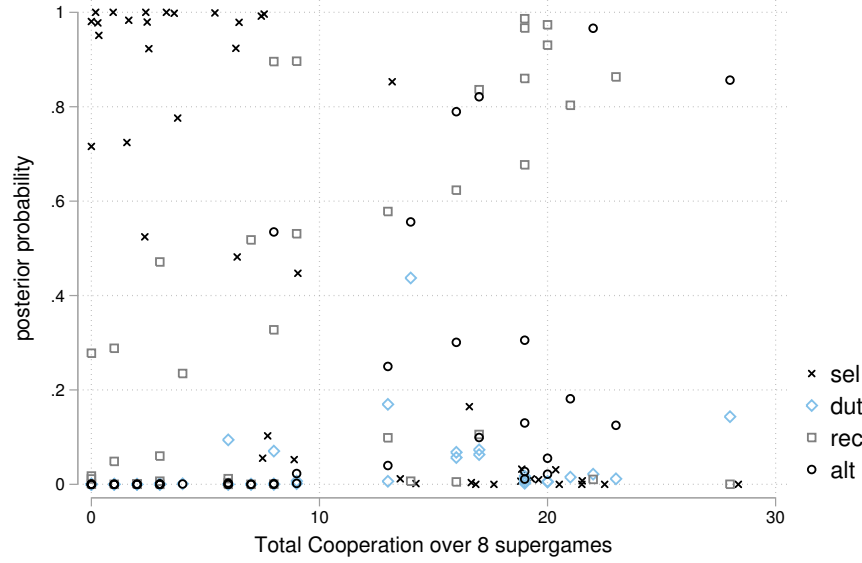
FIGURE 7. Heterogeneity in conditional cooperation



(i) $\delta \leq \frac{3}{8}$ Within, Against C vs Against D.

(ii) $\delta = \frac{3}{4}$ Between, separately for subjects who cooperate less or more than once in Period 1.

FIGURE 8. Jittered scatter plot of posterior type-probabilities against the total count of cooperation over 8 supergames. ($\delta = 3/4$ treatment)



REFERENCES

- James Andreoni and John H Miller. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *The economic journal*, 103(418):570–585, 1993.
- Yoella Bereby-Meyer and Alvin E Roth. The speed of learning in noisy games: partial reinforcement and the sustainability of cooperation. *The American economic review*, 96(4):1029–1042, 2006.
- Ken Binmore, John McCarthy, Giovanni Ponti, Larry Samuelson, and Avner Shaked. A backward induction experiment. *Journal of Economic theory*, 104(1):48–88, 2002.
- Matthias Blonski and Giancarlo Spagnolo. Prisoners' other dilemma. *International Journal of Game Theory*, 44:61–81, 2015.
- Pedro Dal Bó. Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *American Economic Review*, 95(5):1591–1604, December 2005. URL <http://ideas.repec.org/a/aea/aecrev/v95y2005i5p1591-1604.html>.

- Luis Cabral, Erkut Y Ozbay, and Andrew Schotter. Intrinsic and instrumental reciprocity: An experimental study. *Games and Economic Behavior*, 87: 100–121, 2014.
- Evan Calford and Ryan Oprea. Continuity, inertia, and strategic uncertainty: A test of the theory of continuous time games. *Econometrica*, 85(3):915–935, 2017.
- Anujit Chakraborty and Chad W Kendall. Noisy foresight. Technical report, National Bureau of Economic Research, 2022.
- Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *The quarterly journal of economics*, 117(3):817–869, 2002.
- Gary Charness, Luca Rigotti, and Aldo Rustichini. Social surplus determines cooperation rates in the one-shot prisoner’s dilemma. *Games and Economic Behavior*, 100:113–124, 2016.
- R. Cooper, D.W. DeJong, and T.W. Ross. Cooperation without reputation: Experimental evidence from prisoner’s dilemma games. Papers 36, Boston University - Industry Studies Programme, 1992. URL <http://ideas.repec.org/p/fth/bostin/36.html>.
- Caleb A Cox, Matthew T Jones, Kevin E Pflum, and Paul J Healy. Revealed reputations in the finitely repeated prisoners’ dilemma. *Economic Theory*, 58(3):441–484, 2015.
- Pedro Dal Bó and Guillaume R Fréchette. The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, 101(1):411–29, 2011.
- Robyn M Dawes. Social dilemmas. *Annual review of psychology*, 31(1):169–193, 1980.
- Martin Dufwenberg and Georg Kirchsteiger. A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268 – 298, 2004. ISSN 0899-8256. doi: <https://doi.org/10.1016/j.geb.2003.06.003>. URL <http://www.sciencedirect.com/science/article/pii/S0899825603001908>.
- Matthew Embrey, Guillaume R Fréchette, and Sevgi Yuksel. Cooperation in the finitely repeated prisoner’s dilemma. *The Quarterly Journal of Economics*, 133(1):509–551, 2018.

- Dirk Engelmann and Martin Strobel. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American economic review*, 94(4):857–869, 2004.
- Ernst Fehr and Klaus M Schmidt. A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868, 1999.
- Guillaume R Fréchette and Sevgi Yuksel. Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination. *Experimental Economics*, 20(2):279–308, 2017.
- Daniel Friedman and Ryan Oprea. A continuous dilemma. *The American Economic Review*, 102(1):337–363, 2012.
- Drew Fudenberg, David G. Rand, and Anna Dreber. Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, 102(2):720–49, April 2012. doi: 10.1257/aer.102.2.720. URL <http://www.aeaweb.org/articles?id=10.1257/aer.102.2.720>.
- Werner Güth and Reinhard Tietz. Ultimatum bargaining behavior: A survey and comparison of experimental results. *Journal of Economic Psychology*, 11(3):417–449, 1990.
- Yuval Heller. Three steps ahead. *Theoretical Economics*, 10(1):203–241, 2015.
- Philippe Jehiel. Limited foresight may force cooperation. *The Review of Economic Studies*, 68(2):369–391, 2001.
- Chad Kendall and Anujit Chakraborty. Future self-proof elicitation mechanisms. *Available at SSRN 4032946*, 2022.
- David M Kreps, Paul Milgrom, John Roberts, and Robert Wilson. Rational cooperation in the finitely repeated prisoners’ dilemma. *Journal of Economic Theory*, 27(2):245 – 252, 1982. ISSN 0022-0531. doi: 10.1016/0022-0531(82)90029-1. URL <http://www.sciencedirect.com/science/article/pii/0022053182900291>.
- Marco Mantovani et al. Limited backward induction: Foresight and behavior in sequential games. *Unpublished paper, Working Papers*, 289, 2014.
- Richard D McKelvey and Thomas R Palfrey. An experimental study of the centipede game. *Econometrica*, 60(4):803–836, 1992.
- Friederike Mengel. Learning by (limited) forward looking players. *Journal of Economic Behavior & Organization*, 108:59–77, 2014.

- Friederike Mengel. Risk and temptation: A meta-study on prisoner's dilemma games. *The Economic Journal*, 128(616):3182–3209, 2018.
- Thomas R Palfrey and Howard Rosenthal. Private incentives in social dilemmas: The effects of incomplete information and altruism. *Journal of Public Economics*, 35(3):309–332, 1988.
- Matthew Rabin. Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 83(5):1281–1302, December 1993. URL <https://ideas.repec.org/a/aea/aecrev/v83y1993i5p1281-1302.html>.
- Roy Radner. Can bounded rationality resolve the prisoner's dilemma? 1986.
- Jeevant Rampal. Opponent's foresight and optimal choices. *working paper*, 2018.
- Alvin E. Roth and J.Keith Murnighan. Expecting continued play in prisoner's dilemma games. *Journal of Conflict Resolution*, 27(2):279–300, 1983.
- Gerardo Sabater-Grande and Nikolaos Georgantzis. Accounting for risk aversion in repeated prisoners' dilemma games: An experimental test. *Journal of economic behavior & organization*, 48(1):37–50, 2002.
- Emanuel Vespa. An experimental investigation of cooperation in the dynamic common pool game. *International Economic Review*, 61(1):417–440, 2020.

parameters	FMM			
	[1]	[2]	[3]	[4]
$\lambda_{1,\delta=3/4}$	491.27 (44.15)	237.32 (12.04)	253.48 (12.88)	520.69 (46.07)
$\lambda_{1,\delta=1/4}$	299.58 (26.90)	86.79 (7.34)	80.37 (6.02)	293.70 (35.52)
$r_{\delta=3/4}$	.64 (.02)			.65 (.02)
$r_{\delta=1/4}$	.26 (.01)			.26 (.01)
$\beta_{DD,3/4}$	0.00 (.00)	0.00 (.00)	0.00 (.00)	0.00 (.00)
$\beta_{CD,3/4}$	.05 (.05)	.12 (.04)	.05 (.03)	.04 (.05)
$\beta_{DC,3/4}$	.26 (.04)	.33 (.04)	.18 (.03)	.25 (.03)
$\beta_{CC,3/4}$	.81 (.06)	.96 (.09)	.99 (.08)	.86 (.07)
$\beta_{DD,1/4}$	0.00 (.00)	0.00 (.00)		0.00 (.00)
$\beta_{CD,1/4}$	.02 (.01)	.04 (.05)		.01 (.00)
$\beta_{DC,1/4}$	.02 (.005)	.01 (.06)		.14 (.02)
$\beta_{CC,1/4}$	1.00 (.00)	.48 (.06)		1.00 (.00)
α_D	214.65 (35.88)	184.69 (39.01)	168.10 (33.58)	
α_E	252.05 (24.99)	221.36 (13.13)	214.35 (10.82)	227.00 (30.28)
α_A	.37 (.03)	.56 (.02)	.50 (.02)	.50 (.04)
p_S	.40 (.06)	.47 (.06)	.49 (.06)	.43 (.06)
p_D	.04 (.02)	.02 (.02)	.01 (.01)	
p_E	.36 (.06)	.45 (.06)	.36 (.06)	.38 (.06)
p_A	.20 (.05)	.06 (.03)	.14 (.04)	.19 (.04)
LL	-1222.00	-1636.32	-1648.80	-1261.17
AIC	2480.01	3304.63	3321.60	2554.34

TABLE 6. Finite Mixture Model analysis of the four utility models. Column [2] restricts $r_{\delta=3/4} = r_{\delta=1/4} = 1$. Column [3] further restricts the same belief parameters across both treatments, that is, $\beta_{o,\delta=3/4} = \beta_{o,\delta=1/4}$. Column [4] is identical to Column [1], but restricts $p_D = 0$. We report the asymptotic standard errors below the estimates. The last two rows report the log-likelihood and AIC.

Data features	Kreps et al. [1982] (Selfish)	Pure Altruism	Duty	Efficiency Seeking	DK (2004)
As δ increases, initial cooperation increases	✓	✓	✓	✓	
As δ decreases, terminal cooperation increases	X	X	✓	✓	
Initial cooperation does not guarantee later cooperation			X		
Partner's initial defection does not guarantee later defection					X

TABLE 7. ✓ = Data matches prediction, X= Data rejects prediction. A blank box means the theory does not make a related prediction.

	Periods 2-5		Periods 4-5
	First period Cooperators		All subjects
	$\delta \leq 3/8$	$\delta > 3/8$	$\delta \leq 3/8$
Partner cooperated in previous period	$\frac{78}{123}$ (63%)	$\frac{160}{226}$ (71%)	$\frac{88}{163}$ (53.99%)
Partner defected in previous period	$\frac{43}{273}$ (16%)	$\frac{51}{406}$ (12%)	$\frac{77}{893}$ (8.62%)

TABLE 9. Conditional cooperation instead of duty: Frequency of Cooperation in Periods 2-5 by First Period Cooperators in Within treatments. (Compare with Table 4 from the Between sessions)

$\delta = 1/4$	Within	
Treatment	(1) Period 4	(2) Period 5
$Coop_{\tau=1}$	-1.74 (.79)	-1.10 (.66)
$PCoop_{\tau=1}$	-.98 (.87)	-1.08 (.85)
$\sum_{\tau=1}^{\tau=t-1} Coop_{\tau}$	1.76 (.34)	1.00 (.25)
$\sum_{\tau=1}^{\tau=t-1} PCoop_{\tau}$	1.10 (.33)	.49 (.21)
Constant	-2.88 (.30)	-2.39 (.24)
Observations	264	264
Standard errors, clustered at the level of subjects, are reported in parentheses below		

TABLE 10. Logistic regression of cooperation in Period 4 or Period 5 on behavior from period $\tau = 1$ and total previous cooperation. (Compare with Table 5 from the Between sessions)

parameters	Nested Logit Models				
	[1]	[2]	[3]	[4]	[5]
	unrestricted	$\alpha_E = 0$	$\alpha_A = 0$	$\alpha_D = 0$	$\alpha_E = \alpha_A = \alpha_D = 0$
$\lambda_{1,\delta=3/4}$	275.04 (26.58)	284.12 (28.17)	505.54 (35.36)	275.04 (26.58)	499.93 (33.93)
$\lambda_{1,\delta=1/4}$	722.55 (71.36)	800.2 (76.50)	467.80 (40.19)	722.55 (71.36)	455.86 (39.26)
$r_{\delta=3/4}$	.56 (.02)	.53 (.02)	.70 (.02)	.56 (.02)	.71 (.02)
$r_{\delta=1/4}$	.24 (.01)	.20 (.01)	.23 (.01)	.24 (.01)	.25 (.01)
$\beta_{CD,3/4}$	0.00 (.00)	0.00 (.00)	.15 (.07)	0.00 (.00)	.15 (.07)
$\beta_{DC,3/4}$	.14 (.05)	.11 (.05)	.33 (.05)	.14 (.05)	.35 (.05)
$\beta_{CC,3/4}$	1.00 (.00)	.99 (.07)	1.00 (.00)	1.00 (.00)	1.00 (.00)
$\beta_{CD,1/4}$	.53 (.05)	.64 (.05)	.42 (.06)	.53 (.05)	1.00 (.00)
$\beta_{DC,1/4}$	.33 (.05)	.43 (.05)	.24 (.05)	.33 (.05)	.33 (.05)
$\beta_{CC,1/4}$	1.00 (.00)	1.00 (.00)	1.00 (.00)	1.00 (.00)	1.00 (.00)
α_D	0.00 (.00)	1.38 (.12)	0.00 (.00)	—	—
α_E	4.37 (.72)	—	10.76 (1.68)	4.37 (.72)	—
α_A	.20 (.02)	.25 (.02)	—	.20 (.02)	—
LL	-1356.07	-1380.11	-1388.71	-1356.07	-1532.22

TABLE 11. Nested logit models: Column [1] is the unconstrained model that nests the features of Altruism, Efficiency-Seeking and Duty. Models [2], [3] and [4] are special cases of model [1], under $\alpha_E = 0$, $\alpha_A = 0$ and $\alpha_D = 0$ respectively. Model [5] has purely selfish preferences. We report the asymptotic standard errors below the estimated coefficients. We imposed $\beta_{DD,\delta} = 0$ given it was consistently 0 in all versions of the FMM.

	(D,D)	(D,C)	(C,D)	(C,C)
1	1200	2600	200	2000
2	300	650	50	500
3	75	162.5	12.5	125
4	18.75	40.63	3.13	31.25
5	4.69	10.16	.78	7.81

TABLE 13. Payoff parameters in the five periods of the $\delta = 1/4$ supergame approximated to 2 decimal places.