

# A Continuous Dilemma <sup>\*</sup>

Daniel Friedman<sup>†</sup>      Ryan Oprea<sup>‡</sup>

September 10, 2009

## Abstract

We study prisoner's dilemmas played in continuous time with flow payoffs over 60 seconds. In most cases, the median rate of mutual cooperation rises to 90% or more. Control sessions with 8-time repeated matchings achieve less than half as much cooperation, and cooperation rates approach zero in one-shot control sessions. In follow-up sessions with a variable number of subperiods, cooperation rates increase nearly linearly as the grid size decreases and, with one-second subperiods, they approach the level seen in continuous sessions. Our data support a strand of theory that explains how the capacity to respond rapidly stabilizes cooperation and destabilizes defection in the prisoner's dilemma.

**Keywords:** Prisoner's dilemma, game theory, laboratory experiment, continuous time game.

**JEL codes:** C73, C92, D74

---

<sup>\*</sup>We are grateful to the National Science Foundation for support under grant SES-0925039, and to seminar audiences at Harvard, NYU, UCSC and Santa Clara University. We thank Leo Simon, Guillaume Frechette, David Laibson, Ennio Stacchetti and Drew Fudenberg for helpful comments, James Pettit for crucial programming support, and Keith Henwood for research assistance. We retain sole responsibility for remaining idiosyncrasies and errors.

<sup>†</sup>Economics Department, University of California, Santa Cruz, CA, 95064. *dan@ucsc.edu*

<sup>‡</sup>Economics Department, University of California, Santa Cruz, CA, 95064. *roprea@ucsc.edu*

When does cooperation prevail in a prisoner’s dilemma? This question has inspired a grail quest among game theorists ever since Flood (1952) first used the game to show that equilibrium can be inconsistent with efficiency. In recent years, researchers have explored the efficiency-enhancing roles of spatial matchings, reward and punishment technologies and social preferences.

In this paper we examine a much simpler pro-efficiency device: continuous time. Existing theoretical and empirical work, discussed below, has explored the impact of finite and infinite repetition of the game under the assumption that decisions are made simultaneously and in discrete time. However, many real world environments (ranging from old fashioned team production problems such as construction projects to real time price competition in e-commerce markets) simply do not fit this mold—decisions are made asynchronously in real time rather than lockstep on a predefined grid. As explained below, standard game theory can mislead in such continuous games. There are, however, several different theoretical models of games played in continuous time. The predictions for the continuous time prisoner’s dilemma range from no cooperation to full cooperation to anything goes.

In this paper we use an experiment to help sort things out. Our laboratory subjects are matched anonymously in each 60-second period, within which they can freely switch between cooperation and defection. They accrue flow payouts from one of four parametrizations of the prisoner’s dilemma, and then are randomly rematched for the next period. Each session runs 32-36 periods. For comparison, we also run control sessions with one shot periods and repeated discrete time periods, using identical payoff matrices, period lengths and matching procedures.

The results are striking. Continuous time enables median mutual cooperation rates of 90% or more in most of the parametrizations, more than double the rates seen in our discrete (8 stages per period) repeated games, while mutual cooperation becomes quite rare in the one shot sessions. A separate series of sessions varies the number of discrete stages from 2 to 60 in each 60 second period. The data here show a smooth negative, almost linear, relationship between the the cooperation rate and the length of the stage game. Indeed, the mutual cooperation rates in the 30 second stages (2 per period) repeated game are not far from zero, and those in the one second stages (60 per period) are not far from the continuous cooperation rates. Both the aggregate data and the more detailed data seem broadly consistent with one of the theoretical models, Simon and Stinchcombe (1989).

The rest of the paper is organized as follows. Section I reviews some theoretical and experimental work related to our investigation, and section II lays out our experimental design. In section III we report results from our first set of sessions while section IV reports a second wave of sessions conducted for diagnostic purposes. Section V offers a broader discussion of the findings and remaining questions.

	<b>A</b>	<b>B</b>
<b>A</b>	10,10	0,x
<b>B</b>	x,0	y,y

Table 1: Generic form of prisoner’s dilemma. Temptation parameter  $x \in (10, 20)$  and punishment parameter  $y \in (0, 10)$ .

## 1 Theoretical Background

Table 1 parametrizes the standard prisoner’s dilemma payoff bimatrix (see Rapoport and Chammah, 1965, for a related parametrization). With no loss of theoretical generality, the table normalizes the “cooperation” payoff at 10 and the “sucker” payoff at 0. Strategy B is strictly dominant, and so (B,B) is the unique Nash equilibrium, as long as the “temptation” payoff satisfies  $x > 10$ . The restriction  $y < 10$  on the “punishment” payoff ensures that the Nash equilibrium is inefficient, and  $x < 20$  ensures that the sucker-temptation profiles (A,B) and (B,A) also yield a lower payoff sum than the cooperation profile (A, A). Thus the dilemma: the unique equilibrium is inefficient.

Since Flood (1952), legions of theorists have sought ways to evade the dilemma and to support cooperation. They soon discovered that finite repetition doesn’t help: as explained in any modern game theory text, backward induction eliminates every equilibrium profile sequence except all-defect—(B,B) every period. Patient pairs of players rematched over an infinite sequence of stages can support cooperation, but by the Folk Theorem (e.g., Fudenberg and Maskin, 1986), they can just as easily support (as a Nash equilibrium of the repeated game) all-defect and a wide variety of other inefficient profile sequences.

What happens if the game is played over a continuous finite time interval, say  $t \in [0, 1]$ ? Perhaps the most obvious approach is to specify a minimum reaction time  $\epsilon$  and to formalize the game as a finitely repeated game with  $1/\epsilon$  stages (rounding up to the nearest integer). The theoretical prediction again is that the dilemma persists, and only all-defect survives in Nash equilibrium.

Huberman and Glance (1993) offer a useful caution. They show that earlier positive results on spatial versions of the repeated prisoner’s dilemma evaporate when players move asynchronously, in real time. Their point is that coordination and cooperation get an artificial boost when all players must move simultaneously at discrete time intervals. Their results do not suggest more cooperation in the continuous time prisoner’s dilemma, but they do point up the need for more fully articulated models.

Bergin and MacLeod (1993) develop a general model of games played in continuous time. They

assume that actions cannot be reversed within time  $\epsilon$ , look for  $\epsilon$ -equilibria, and pass to the limit as  $\epsilon$  goes to 0. For the prisoner's dilemma, they obtain a Folk Theorem result: virtually any profile sequence that gives each player an average payoff of at least  $y < 10$  can be supported as a Nash equilibrium (indeed, one that is renegotiation-proof).

Simon and Stinchcombe (1989) propose a subtly different model of games played in continuous time. They consider discrete grids in the time interval  $[0, 1)$  for games with finite numbers of players and actions. In the limit as the grid interval approaches zero while the number of strategy switches remains uniformly bounded, they obtain upper and lower hemi-continuity conditions that guarantee well-defined games in continuous time. Subgame perfection is automatic in these games, but backward induction does not work because the real numbers are not well ordered (Anderson, 1984). For example, time  $t = 1$  has no immediate predecessor: for any previous time, say  $t = 1 - h$ , there are an infinite number of later times that fall before  $t = 1$ , e.g.,  $t = 1 - h/17$ . Consequently some repeated game equilibria disappear in the continuous limit, while new equilibria can appear. Simon and Stinchcombe focus on Nash equilibria in which weakly dominated strategies are not played, and find a unique such equilibrium for the Prisoner's dilemma: full cooperation at all times.

To summarize, existing theoretical literature offers three competing predictions for our continuous time experiment. The extended theory of finitely repeated games predicts that the all-defect profile (B,B) will predominate; Simon and Stinchcombe's model predicts that full cooperation profile (A,A) will predominate; and the extended Folk Theorem predicts virtually any profile sequence that does not too often give either player the sucker payoff of 0.

The noiseless theories just mentioned predict no impact from changing the parameters  $x, y$  within their admissible ranges. All noisy versions of which we are aware predict that players will choose A less often when either  $x$  or  $y$  increases.

The continuous time predictions have not yet been tested in the lab, but existing literature offers some tantalizing clues. Selten and Stoecker (1986), Andreoni and Miller (1993), Hauk and Nagel (2001) and Bereby-Myer and Roth (2006) all study behavior in discrete 10-stage repeated prisoner's dilemmas. These papers report that, after several repetitions of the repeated game, most subjects cooperate in early stages but cooperation begins to unravel around the fifth stage and is rare after the 8th stage. Thus the unravelling process of backwards induction seems at best incomplete in the laboratory data. Cooper et al. (1996) argue that the data are also inconsistent with more elaborate explanations such as altruism and the Kreps et al. (1982) model.

Another branch of empirical literature studies behavior in “infinitely repeated” discrete time games, in which there is a known probability  $q$  that the matching ends after the current stage. Roth and Murningham (1978) produced mixed evidence for the theoretical prediction that cooperation is possible in such games. More recently, Dal Bo (2005) finds that individual cooperation rates respond sensitively to  $q$ , exceeding 50 percent in the most favorable case. Dal Bo and Frechette (2008) show that experience in these repeated games does not necessarily lead to greater cooperation but that variation in parameters similar to our  $x$  and  $y$  also have a significant impact on cooperation. Individual cooperation rates average roughly 35 percent and rise to 76 percent with particularly high continuation probabilities and stage game parameters considerably more conducive to cooperation than any studied in this paper. In both of these papers, the “shadow of the future” seems pivotal to cooperation.

There are several ways to extrapolate these empirical results to continuous time. In 60 second periods, the shadow of the future shrinks steadily to zero. Will cooperation also decline steadily to zero? Our sessions all contain more stationary repetitions (e.g., of the repeated game) than in previous studies, and continuous time provides unprecedented learning opportunities. Will subjects learn to unravel cooperation more completely? Or will asynchronous decisions or other aspects of continuous time twist the strategic behavior in a different direction? Answering such questions clearly requires new experiments.

## 2 Treatments and Experimental Design

We ran experiments using a new software package called ConG, for Continuous Games. Figure 1 shows the user interface. Each subject can freely switch between row strategies A and B by clicking a radio button (or pressing an arrow key), causing the chosen row to be shaded. In our main treatment (Continuous time) the other player’s current choice is shown as a shaded column, and the intersection is doubly shaded. The computer response time to strategy switches is less than 50 milliseconds, giving players the experience of continuous action. The screen also shows the time series of actions (coded here as 1 for A and 0 for B) for the player and her counterpart in the upper right graph, while flow payoffs for each player are shown in the lower right graph. The top of the screen also shows the time remaining and the accumulated flow payoff.

We study three treatments of time: Continuous, One-Shot, and Grid. In all treatments, each period lasts 60 seconds, during which subjects are allowed to change their strategies at will. In Continuous time, subjects observe the unfolding history of actions and payoffs, and at the end of the period

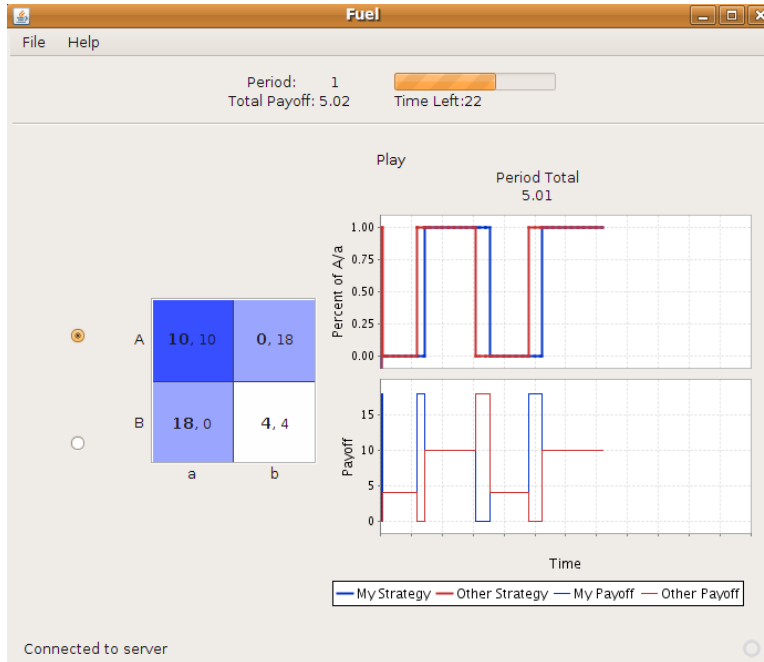


Figure 1: Screenshot of continuous display.

they earn the integral of the flow payoffs shown in the lower right hand graph of Figure 1.

In One-Shot time, subjects do not observe their counterpart's action until the period's end. They earn the lump sum payoffs for the strategy profile chosen at that point.

Grid time divides each sixty second period into  $n$  equal subperiods, during each of which subjects cannot see the choice of their counterparts. Payoffs for the subperiod are determined by the last strategy profile chosen in each subperiod and are plotted on the right hand side of the screen. That last profile becomes the initial profile of the next subperiod. End of period payoffs are the average of the lump sum subperiod payoffs or, equivalently, the integral across subperiods of the piecewise constant flow payoffs. Thus One-Shot time is the same as Grid time with  $n = 1$ , and Continuous time is closely approximated by Grid time with  $n > 300$ .

Our other treatment variable, payoff parameters, examines four different configurations of  $(x, y)$ , one from each quadrant of the admissible domain  $(10, 20) \times (0, 10)$ . They are Easy =  $(14, 4)$ , Mix-a =  $(18, 4)$ , Mix-b =  $(14, 8)$ , and Hard =  $(18, 8)$ .<sup>1</sup> The names reflect the conjecture that cooperation will be more difficult given either a larger temptation  $x$  or a larger punishment payoff  $y$ .

In all treatments, subjects are randomly rematched with a new counterpart each period. At period's beginning, each of the four possible initial profiles is chosen independently with probability 0.25.

<sup>1</sup>Subjects were paid 5 cents per point earned in the experiment.

We first ran 4 sessions for Continuous and 3 parallel sessions for One-Shot <sup>2</sup>. Ten subjects participated in each session (except for one Continuous session with only eight subjects), which consisted of 32 periods divided into 8 blocks. Each of the four parameter sets appears once, in random order, in each block, and the sequences are matched across the two time treatments. Then we ran another 4 matched sessions, again using the same sequences, under the Grid treatment with  $n = 8$  subperiods (hereafter called Grid-8 sessions). These sessions are comparable to 10-stage repeated game experiments noted earlier. We also ran three additional Grid sessions, to be described later, that varied  $n$  within session.

A key aspect of our design is that period lengths and payoff potentials are kept constant across Continuous, One-Shot and Grid-8 treatments. The only difference between these treatments is the frequency with which subjects can adjust their action choices.

Subjects in all sessions were inexperienced University of California, Santa Cruz undergraduates from the LEEPS lab subject pool. On arrival, subjects received written instructions (attached as Appendix A) which also were read out loud. Sessions lasted on average 75 minutes, and average earnings were roughly US\$17.50.

### 3 Main Results

Observed behavior can be summarized in the fraction  $\rho_{ipk}$  of time spent in each profile  $\rho$  by player  $i$  and her counterpart  $j(i, p, k)$  in period  $p$  of session  $k$ . Due to the symmetry of the game, the four strategy profiles reduce to three player-pair profiles:

- Cooperation ( $\rho = c$ ): Profile (A,A).
- Defection ( $\rho = d$ ): Profile (B,B).
- Sucker-Temptation ( $\rho = s$ ): Profile (B,A) or (A,B).

For example, if player 2 spent equal time in each of the four strategy profiles in period 3 of session 4, the data would be  $c_{234} = d_{234} = 0.25$  and  $s_{234} = 0.50$ . By definition,  $\rho_{ipk} \in [0, 1]$  and  $\sum_{\rho} \rho_{ipk} = c_{ipk} + d_{ipk} + s_{ipk} = 1$ . We have the further restriction that  $\rho_{ipk} = 0$  or 1 in the One-Shot treatment, and that  $\rho_{ipk} \in \{m/n : m = 0, 1, 2, \dots, n\}$  in the Grid- $n$  treatment.

---

<sup>2</sup>Due to a coding error, we were forced to drop several periods from a single One-Shot session. All results pertaining to the One-Shot treatment are robust to, instead, using the entire dataset or dropping the entire offending session.

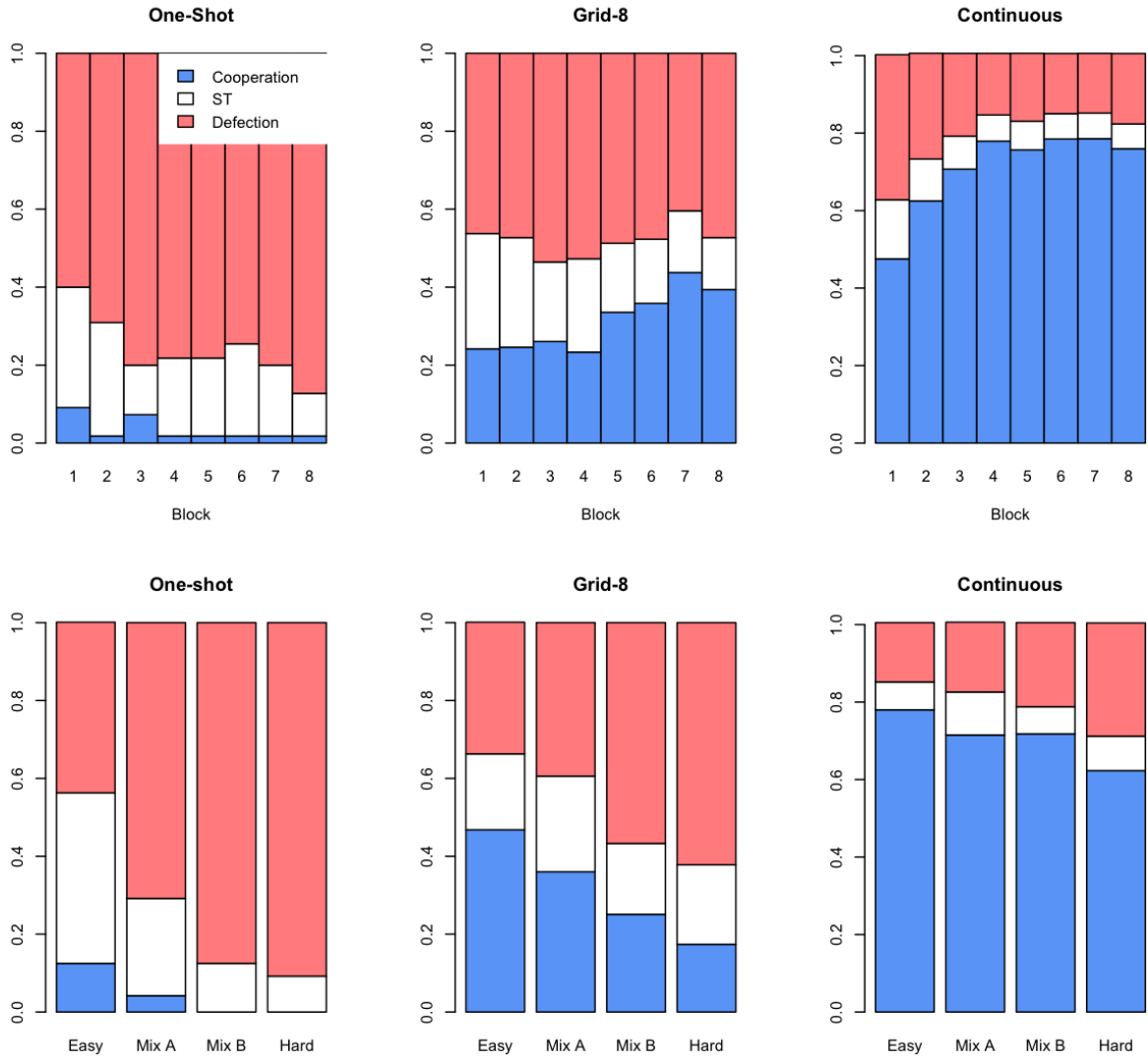


Figure 2: Main effects by treatment. The top row shows average profiles over blocks. The bottom row aggregates data by parameter set.



Figure 2 shows average behavior in the three time treatments,  $T \in \{\text{One-Shot}, \text{Grid-8}, \text{Continuous}\}$ .<sup>3</sup> The top row shows averages over successive 4-period blocks,

$$\rho_{bT} = \frac{\sum_{k \in T} \sum_{p \in b} \sum_i \rho_{ipk}}{\sum_{k \in T} \sum_{p \in b} \sum_i \sum_\rho \rho_{ipk}}. \quad (1)$$

The randomized block design ensures that each block  $b = 1, \dots, 8$  includes each of the four parameter sets.

Results are striking. Mean cooperation rates drop nearly to zero after the third block in One-Shot, but rise above 75 percent in Continuous.<sup>4</sup> Mean cooperation rates in Grid-8 are intermediate, mostly in the 20–40 percent range. Mean ST rates are low in One-Shot and Grid-8, and are even lower in Continuous.

To test these impressions, we take the averages  $c_{pk} = \sum_i c_{ipk} / \sum_i \sum_\rho \rho_{ipk}$  across subject pairs each period and run the regression

$$c_{pk} = \alpha + \beta \cdot D1S_k + \delta \cdot DG8_k + \psi_k + \epsilon_{pk}, \quad (2)$$

where  $DT$  are the indicator variables for treatments  $T = 1S$  and  $G8$  (One-Shot and Grid-8 respectively), and  $\epsilon$  is a normally distributed disturbance term while  $\psi$  is a normally distributed session random effect.

The intercept, measuring the overall cooperation rate in Continuous, is estimated at 0.708 (standard error 0.032) while the coefficient on D1S is  $-0.675$  (0.049) and on DG8 is  $-0.394$  (0.049), all significant at the one percent level. The coefficient estimates imply much lower cooperation rates in the non-continuous treatments—about 31 percent ( $=0.708-0.394$ ) in Grid-8 and under 5 percent in One-Shot—and the differences are highly significant.

To summarize,

**Result 1** *Cooperation prevails in the Continuous treatment, is less than half as common in Grid-8 and is quite rare in One-Shot.*

---

<sup>3</sup>Most previous authors report individual cooperation rates  $K_T$ . For reasons that will become increasingly clear, we focus on *mutual* cooperation and alternative profiles of pairs of players. Since  $K_T = c_T + 0.5s_T$ , the individual rates are easily recovered and compared to those in previous studies.

<sup>4</sup>Median cooperation rates in Continuous are even more impressive. Over the last four blocks, they are 81 percent in the Hard treatment, 90 percent in both Mix-a and Mix-b, and 93 percent in Easy.

Variable	One-Shot	Grid-8	Continuous
Intercept	0.123*** (0.019)	0.468*** (0.044)	0.781*** (0.046)
$X$	-0.081*** (0.022)	-0.107*** (0.040)	-0.067* (0.037)
$Y$	-0.123*** (0.022)	-0.217*** (0.040)	-0.063* (0.037)
$XY$	0.082*** (0.029)	0.030 (0.057)	-0.033 (0.052)

Table 2: Random effects coefficient estimates (and p-values) for prisoner’s dilemma parameters  $x$  (Temptation) and  $y$  (Punishment). One, two and three stars denote significance at the ten percent, five percent and one percent levels.

The bottom row of Figure 2 shows observed behavior disaggregated by parameter set,

$$\rho_{zT} = \frac{\sum_{k \in T} \sum_{p \in z} \sum_i \rho_{ipk}}{\sum_{k \in T} \sum_{p \in z} \sum_i \sum_\rho \rho_{ipk}}, \quad (3)$$

where  $z \in \{\text{Easy, Mix A, Mix B, Hard}\}$  denotes the periods using the given parameter configuration. As expected, cooperation is most prevalent under Easy parameters, least under Hard parameters and intermediate in the Mix cases. The effects seem stronger in Grid-8 than in either Continuous or One-Shot.

To test these impressions, we estimate the following model separately for each time treatment:

$$c_{pk} = \alpha + \beta \cdot X_p + \delta \cdot Y_p + \kappa \cdot X_p \cdot Y_p + \psi_k + \epsilon_{pk}, \quad (4)$$

where  $X_p$  is an indicator that the payoff parameter  $x = 18$ , its high level, in period  $p$ , and similarly  $Y_p$  indicates  $y = 8$ . Table 2 collects the results. The last column shows that moving to the high level of either parameter is associated with a (marginally significant) six percentage point decrease in cooperation rates. The impacts of both parameters are much larger and are highly significant in the Grid-8 sessions though here the punishment parameter  $Y$  has a much greater impact than the temptation parameter  $X$ . In One-Shot, parameter impacts are moderate but again highly significant.

**Result 2** *An upward shift in either payoff parameter  $x$  or  $y$  slightly reduces cooperation in Continuous, moderately reduces cooperation in One-Shot sessions and substantially reduces cooperation in Grid-8 sessions.*

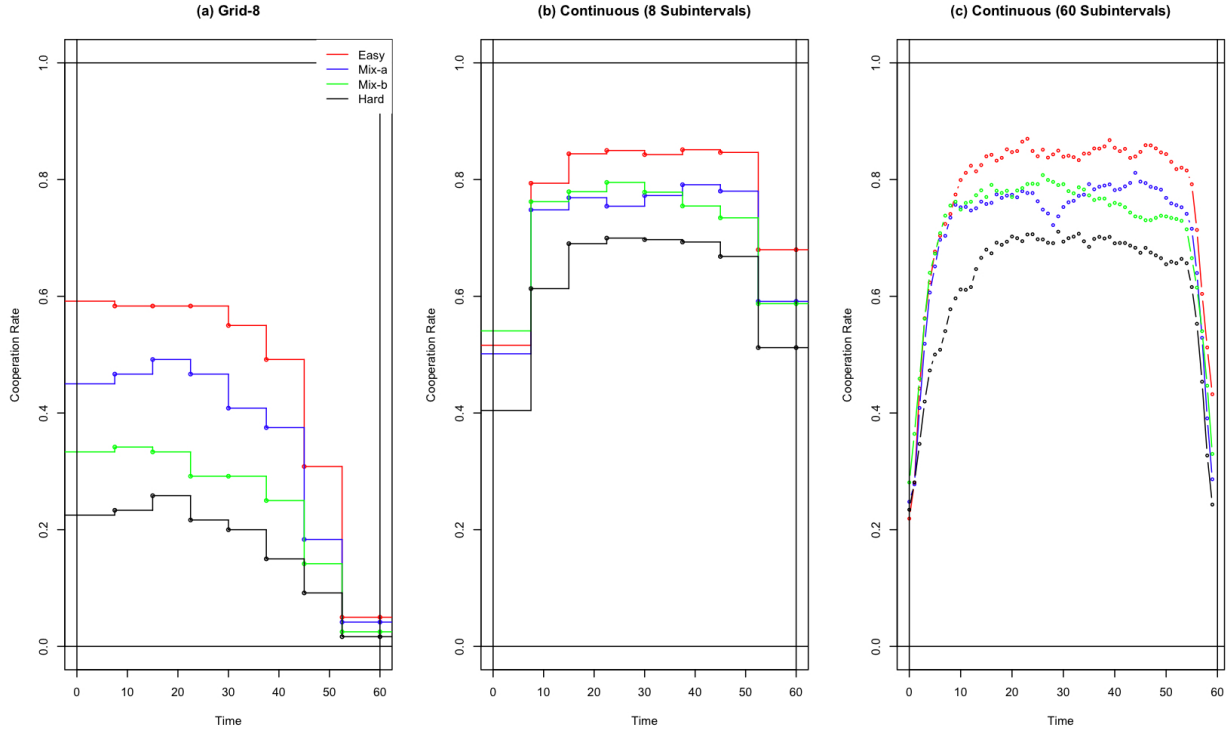


Figure 3: Mutual cooperation rates by subperiod and subinterval.

### 3.1 Behavior Within Period

The patterns of behavior within each Continuous and Grid-8 period provide some insight into our first two results. Panel (a) of Figure 3 presents the average level of cooperation in each subperiod of Grid-8. As in the finitely repeated pairings of Selten and Stoecker (1986) and later authors, some level of cooperation is maintained for several subperiods, but begins to decline roughly halfway through and collapses to nearly zero by the end. The initial level of cooperation is considerably higher with Easy parameters than with Hard ones, and the Mixes lie in between.

Panel (b) of the Figure displays our Continuous data in a comparable manner. It breaks the  $c_{ipk}$  data into 8 equal and consecutive subintervals, and takes the average over each subinterval. Three departures from the Grid-8 data are apparent. First, initial levels of cooperation are higher for Hard and Mix-b parameters. Second, and most striking, Continuous cooperation rates all rise substantially over the first several subintervals, and plateau at much higher levels than their Grid-8 counterparts. Third, cooperation tends not to decay until the very last subinterval, and remains considerably higher than in Grid-8.

Panel (c) provides a finer resolution by reporting average behavior over 60 one-second subintervals.

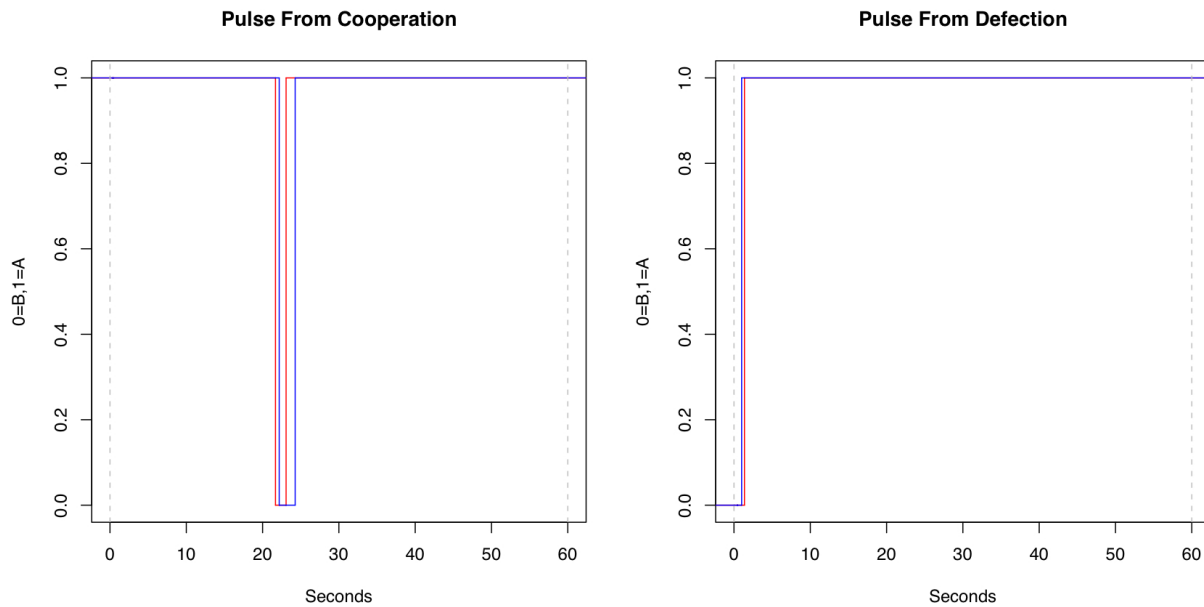


Figure 4: Time series from two sample pairings. Choice A is coded as 1, and choice B as 0; in red for one player and blue for the other.

Due to the random initial strategy assignments, cooperation rates all begin at roughly 25 percent. Cooperation rates rise swiftly and approach their maximum within 10-20 seconds. Then they plateau and only in the final 5 seconds do they drop off sharply. Even in the very last second, however, they remain considerably higher than in the final subperiod of Grid-8.

To summarize,

**Result 3** *Only in the Continuous sessions do cooperation rates increase sharply early in the period. As compared to Grid-8 sessions, cooperation reaches higher levels by midperiod in Continuous sessions, and late in the period it decays more slowly and to a lesser degree.*

### 3.2 Phases, Transitions and Pulses

How do players manage to increase initial cooperation rates in the Continuous time treatment, and sustain cooperation for so long? Figure 4 provides some clues. In panel (a) the player pair was randomly initialized at the Cooperation profile, and they stay there for about 20 seconds. Red then defects to strategy B. However, within a second, Blue follows her, landing them in the Defection profile. Red is now worse off than before and her one second in the ST profile did not earn her much. She soon returns to strategy A, Blue follows after a few seconds, and they earn the Cooperation

Parameters	Cooperation	Defection	ST
Easy	19.93	3.10	0.88
Hard	15.72	4.45	0.92
Mixa	15.10	3.16	1.13
Mixb	18.67	4.62	0.90

Table 3: Mean duration (in seconds) of phases by parameter set and profile type.

payoff the rest of the period. The accumulated flow payoff that period is close to the maximum, and is almost equally split.

In panel (b) the pair of subjects is randomly assigned initially to the Defection profile. Blue soon switches to A, Red quickly follows and they remain in the Cooperation profile for the rest of the period. Even had Blue refused to cooperate, Red could have quickly returned to the defection profile, losing very little for her attempt.

Episodes like these are common in our data. To study them more systematically and to gauge their impact, we divide each period/pair combination into a sequence of “phases” in which the strategy profile is constant. Phase 1 begins when subjects are assigned their initial strategy profile, and subsequent phases begin with a strategy change by either player. Each phase ends with the next strategy change, or with the end of the period, whichever comes first. Each phase is characterized by its duration (in seconds) and its pair profile  $\rho \in \{c, d, s\}$ . For example, in Panel A of Figure 4, the phase sequence is  $c, s, d, s, c$  and duration is quite short for the middle three phases.

Table 3 shows that, contrary to standard theory, cooperation phases are quite stable behaviorally relative to defection phases. On the other hand, the  $s$  phase is indeed quite transitory. Hence one can usefully think of phase transitions as in Figure 5. From the quasi-stable phases  $c$  and  $d$ , a player can “pulse” briefly to the transitory phase  $s$ . Then, usually very soon thereafter, the next transition returns the player pair either to  $c$  or  $d$ .

What are subjects’ incentives to pulse? As noted in Figure 5, the immediate opportunity cost is roughly measured as the difference  $\pi_P$  between the subject’s actual earnings during the pulse phase and her counterfactual earnings in the departure profile ( $c$  or  $d$ ) over the same time interval. We plot the empirical CDFs of these “pulse returns” in Figure 6. Of course, by definition pulse returns from defection are always negative (representing the sucker’s shortfall over some duration, usually short) while those from cooperation are positive (representing the excess temptation payoff), so the CDFs do not overlap.

The message of the Figure is that pulse returns  $\pi_P$  are typically very close to zero. At the median,

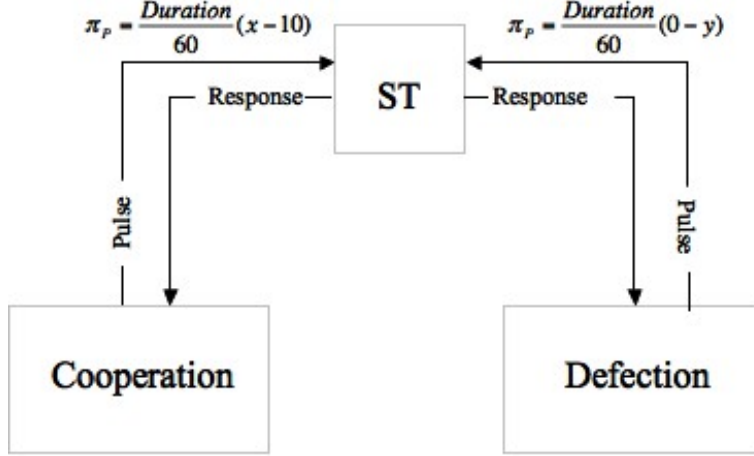


Figure 5: Diagram of transitions between phases.

subjects sacrifice less than one cent by pulsing from Defection. Likewise, the median subject earns less than one cent by pulsing from the Cooperation profile  $c$ .

Of course, the immediate return is not the end of the story—there is a response phase, typically with a much longer duration than the pulse. Therefore we define the net return  $\pi_N$  as the pulse return plus the incremental earnings from the response phase. That is,  $\pi_N$  is the subject’s actual earnings in the pulse phase and the subsequent phase minus the counterfactual earnings over that time interval had she remained in the departure profile ( $c$  or  $d$ ).

Figure 7 plots the empirical CDFs for net returns. Forty percent of pulses from Defection yield positive returns, often quite large, because the other player matches the switch from B to A. The remaining 60 percent of net returns are mostly very close to zero, as the pulsing player typically returns back to B.

For a more formal examination of the question, we run the regression

$$\pi_{Nipk} = \alpha + \beta \cdot X_p + \delta \cdot Y_p + \kappa \cdot X_p \cdot Y_p + \psi_k + \epsilon_{ipk}. \quad (5)$$

Table 4 uses fitted coefficients from this equation to construct the estimated values of  $\pi_N$  for each parameter set. Pulsing from Cooperation leads to returns either negative or insignificantly different from zero. Pulsing from Defection generally leads to significantly positive returns.

**Result 4** *Net returns from pulsing from a Defection profile are significantly positive, while net returns from pulsing from a Cooperation profile are insignificant or negative.*

These findings provide a potential explanation of the observed patterns of behavior. The direct

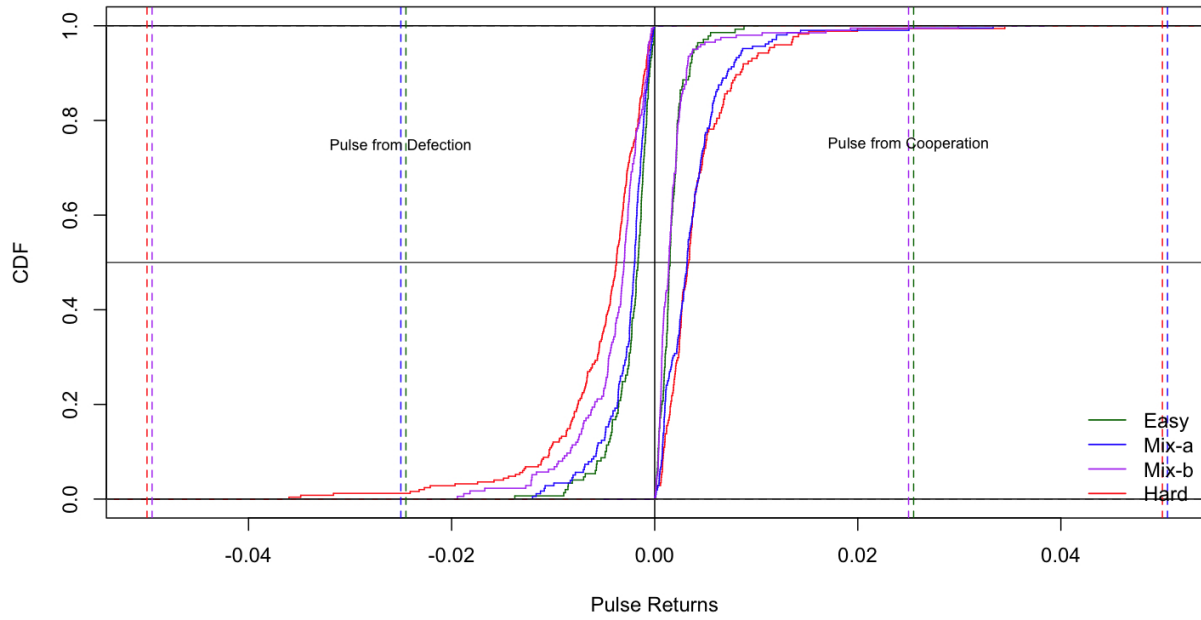


Figure 6: Empirical cumulative distributions of pulse returns. Vertical dashed lines indicate minimum pulse returns in Grid-8.

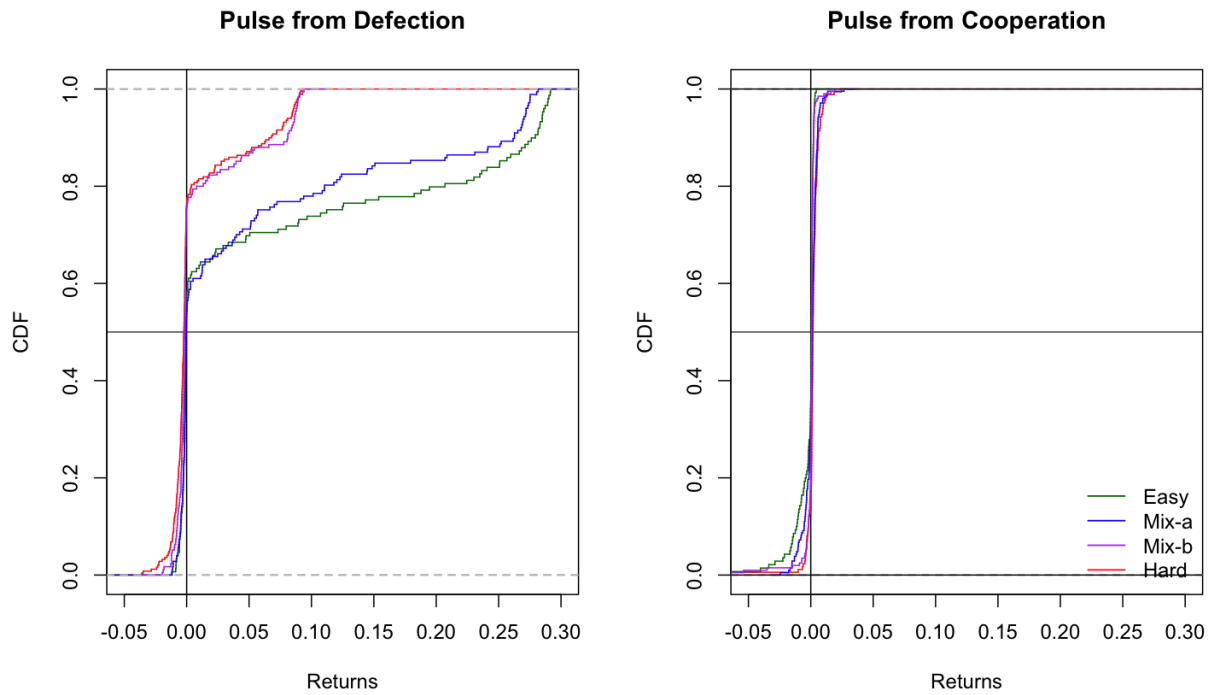


Figure 7: Empirical cumulative distributions of net returns from pulsing.

Variable	Cooperation	Defection
Easy	-0.0074 ** (0.0022)	0.0790 *** (0.0064)
Mix-A	0.0019 (0.0022)	0.0449 *** (0.0064)
Mix-B	-0.0006 (0.0022)	0.0142 ** (0.0064)
Hard	0.0017 (0.0022)	0.0103 * (0.0064)

Table 4: Estimated net returns (in cents) from pulsing from Cooperation or Defection, based on the coefficient estimates for equation (5).

cost of pulsing from the Defection profile and the direct benefit of pulsing from Cooperation both are much smaller in Continuous than in Grid-8, as indicated by the vertical dashed lines in Figure 6. This tends to stabilize  $c$  in the Continuous time treatment and to destabilize  $d$ . Moreover, the longer expected duration of a  $c$  phase in the Continuous treatment increases the net return to pulsing away from  $d$ , at least until late in the period. Thus the returns to pulsing seem to account qualitatively for all aspects of Result 3, as well as for Results 1 and 2.

## 4 Variable Grid Results

The true test of any explanation lies in its excess predictive power—in the verifiable implications beyond the facts it was constructed to explain. The key implication in the present instance is that if we raise either the direct cost of pulsing from  $d$  or the direct benefit of pulsing from  $c$  or both, then we should observe less cooperation and more defection.

Our Grid- $n$  treatment enables a sharp test of this implication. In Grid- $n$ , the number  $\tau = 1/n$  is a lower bound on the duration of the  $s$  phase. Thus, as illustrated in Figure 8, we can use  $n$  to exogenously control the pulse returns.

The Figure repeats the Mix-a Continuous data from Figure 6 and overlays vertical lines representing the minimal pulse return  $\pi_{P|c} = \tau(x - 10) = 8\tau$  from  $c$ , and  $\pi_{P|d} = -\tau y = -4\tau$  from  $d$ , for Mix-a parameters and various values of  $n$ . For  $n = 60$  (or one subperiod per second), the minimal pulse returns are quite close to the actual  $\pi_P$  observed in Continuous data. For larger  $\tau$  or smaller  $n$ , the pulse returns will be approximately the minimal returns, which are linear in  $\tau$ . The implication is that we will see less cooperation and more defection as  $\tau$  increases in Grid- $n$  for  $n < 60$ .



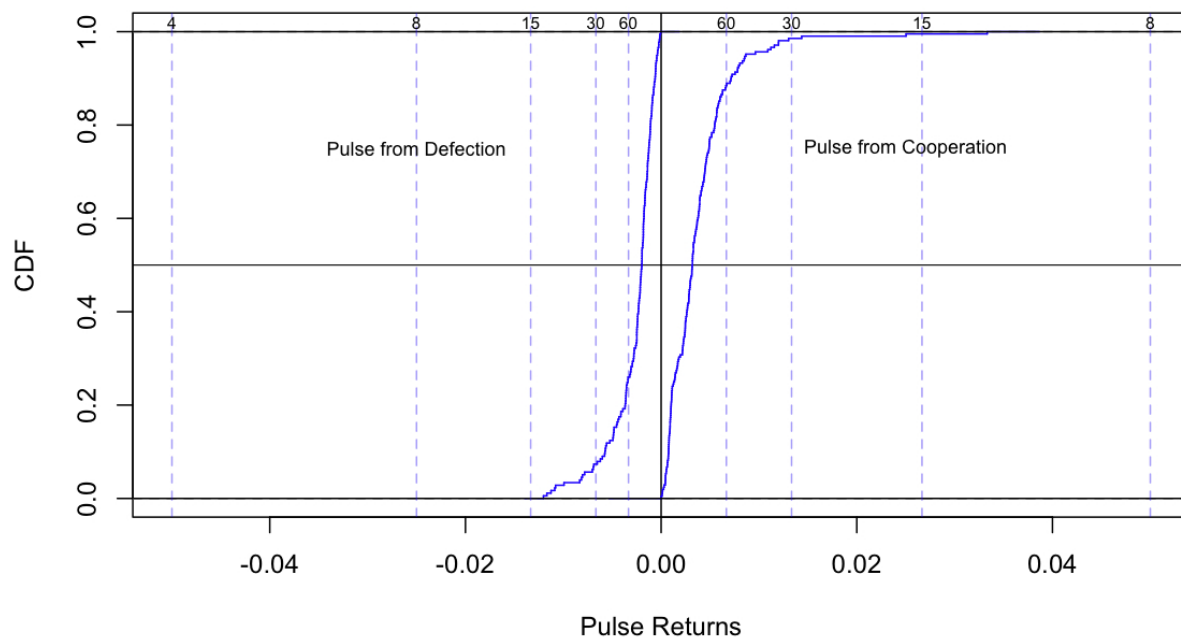


Figure 8: Empirical cumulative distributions of pulse returns for Mix-a parameters. Vertical dashed lines indicate minimum pulse returns in Grid- $n$  for indicated values of  $n$ .

This prediction motivates our three Variable Grid sessions. Those sessions each last 36 periods and, for clarity, use only the Mix-a parameters. In each of three 12 period Blocks, we run each  $n$  twice in consecutive periods and vary  $n$  in the sequence  $\text{Incr} = (2,4,8,15,30,60)$  or  $\text{Decr} = (60,30,15,8,4,2)$  or Random. In two sessions the blocks were sequenced  $\text{Incr-Decr-Random}$  and in the other session the blocks were sequenced  $\text{Decr-Incr-Random}$ . Note that the within-session variation of  $\tau$  is conservative, and is likely (if anything) to understate its impact. To focus on settled behavior, we drop the first two observations of each  $n$  in each session; the results are similar (though a bit noisier) if we include all data.

Figure 9 summarizes the results. At the lowest values of  $\tau$ , cooperation rates are nearly (yet not quite) at levels observed in Continuous. Cooperation rates drop almost linearly as  $\tau$  increases, falling to the One-Shot level at  $\tau = 0.5$  or  $n = 2$ . One sees roughly parallel patterns in the other profiles, except that, at the higher levels of  $\tau$ , the  $d$  rates are lower and the  $s$  rates are higher than in the One-shot data.

A simple regression of cooperation rates on  $\tau$  using data averaged by session and period lends further support to the prediction. The intercept term is 0.573 ( $p=0.000$ ) and the slope is -1.001

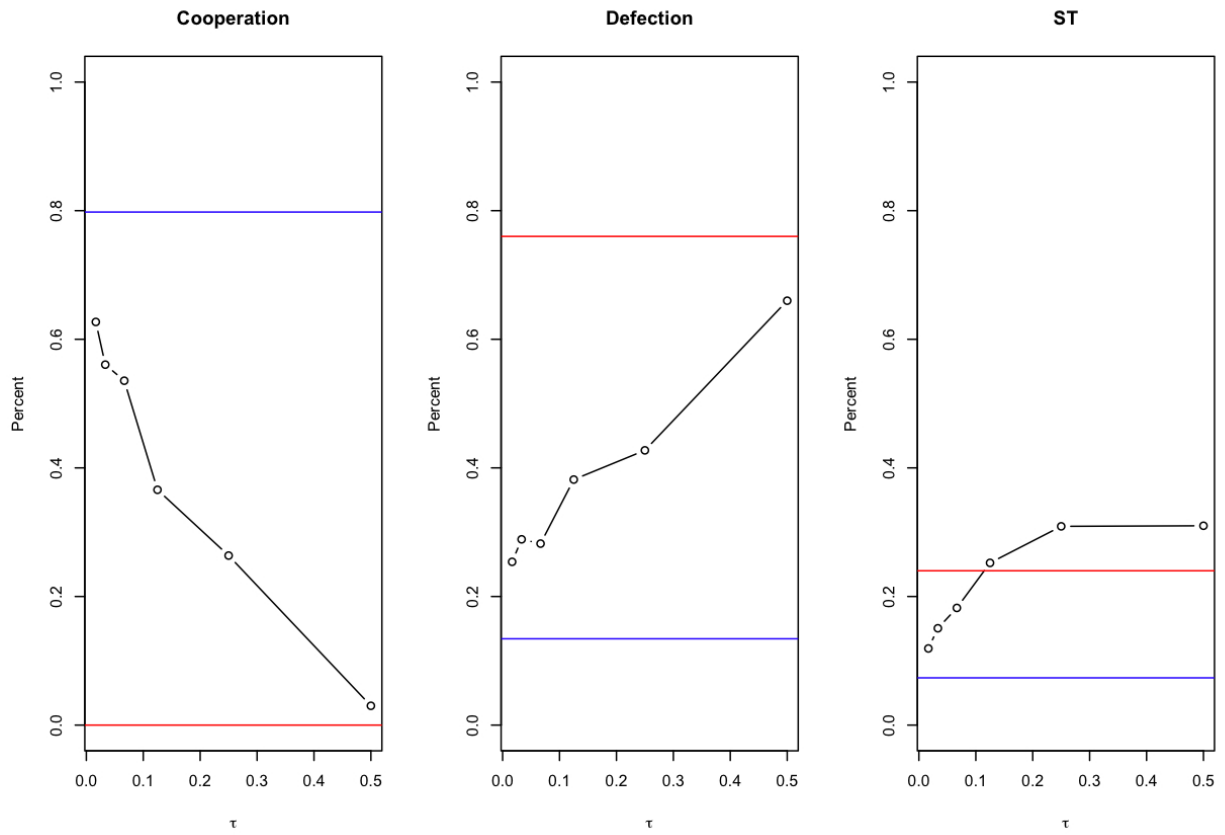


Figure 9: Variable Grid session data. Profile frequencies in periods 13-32 are shown as black lines. Blue horizontal lines represent levels in Continuous and red lines levels in One-Shot sessions with Mix A.

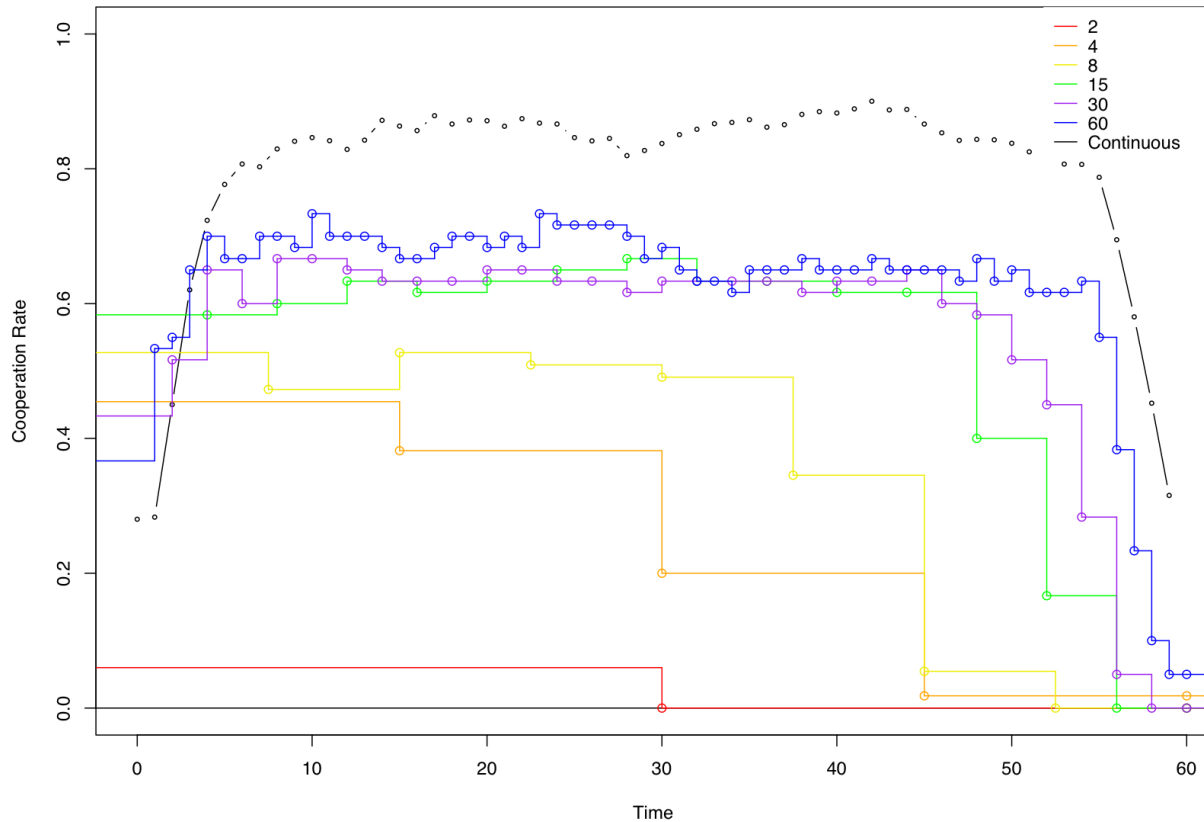


Figure 10: Mutual cooperation rates by  $n$  and subperiod in periods 13-36 of Variable Grid sessions.

( $p = 0.000$ ). Thus we have:

**Result 5** *As predicted, rates of cooperation are monotonically decreasing in the grid size.*

Figure 10 shows the Variable Grid data in more detail, along with the Mix-a data from Panel (c) of Figure 3. Using the same conventions as in Panel (a) of Figure 3, it plots the average rate of cooperation for each value of  $n$ . Figure 10 seems to confirm the explanation, and can be summarized as follows.

**Result 6** *Grid size  $\tau$  affects cooperation in three ways. First, smaller values of  $\tau$  are associated with higher early levels of cooperation. Second, when  $\tau$  is sufficiently small, we see increases in cooperation cooperation rates early in the period. Third, cooperation decays later in the period when  $\tau$  is smaller.*

## 5 Discussion

Our principal findings can be summarized briefly. First and foremost, in the Continuous time treatment, we found remarkably high levels of cooperation in all four parametrizations of the prisoner’s dilemma. Even in the “hard” parametrization with maximal temptation and minimal efficiency loss, the all-cooperate profile was played 81 percent of the time by the median pair of subjects in later periods. The other parametrizations led to median cooperation rates of 90 percent or more. By contrast, in the Grid-8 control treatment, with rapid repeat pairings over 8 subperiods, defection was more prevalent than cooperation, and cooperation became quite rare in the One shot control treatment.

Second, the parametrization had a considerably stronger impact (in the predicted direction) in discrete time than in Continuous time.

Third, early in the 60 second Continuous pairings, average levels of cooperation rose rapidly, stayed at a high level until the last few seconds of the period, and then dropped abruptly (but not to as low levels as in Grid-8 or One-Shot).

We proposed an explanation of these striking patterns in terms of the immediate and net returns to “pulsing” away from defection or cooperation. Data from the Variable Grid sessions, which varied the number of subperiods from 2 to 60, confirmed the explanation.

What are the theoretical implications? Routine application of well-known standard theory, in particular backward induction and iterated dominance, leads to the incorrect prediction of very little cooperation in any of our treatments. More flexible theories of continuous time play, such as that of Bergin and MacLeod (1993), are consistent with almost anything, including our data.

Simon and Stinchcomb’s (1989) model predicts complete cooperation in the continuum limit, and our Continuous data approach that mark. Equally important, the logic of the model seems consistent with our finer cuts of the data. The model specifies (epsilon-) equilibrium strategies in which players unilaterally pulse away from the defection profile because it costs very little to do so but there is much to be gained if the other player matches the move as soon as possible. Also, the strategies do not allow players to pulse away from the cooperation profile, because a rapid match by the other player sharply limits the potential gain and increases the subsequent loss.

Just short of the continuum limit, these strategies are almost, but not quite, best responses; in the limit they are best responses that weakly dominate alternative strategies. Despite the inevitable frictions of laboratory games, our human players often seem to employ strategies of just that sort,

and to do quite well with them. Thus our data suggest that Simon and Stinchcomb's (1989) model is right for the right reason.

Of course, our four sets of sessions are not definitive. Future laboratory studies could test robustness with different payoff parameters, longer or shorter periods, more periods, and variations on near continuous time, e.g., alternating moves, or perceptible lags in implementing strategy switches, or temporary strategy lock-ins.

Our study also points up the potential for more sophisticated structural econometrics. The sequence of phases (or strategy profiles) is endogenous, so it is difficult to generate unbiased estimates of how pulse returns affect the incidence of pulsing. The net returns to pulsing from defection are clearly determined simultaneously with the net returns of pulsing from cooperation, since both depend on duration. And duration depends on the time remaining in the period, as well as the history observed so far. We leave it as a challenge to future researchers to create experimental designs that allow identification and reliable structural estimates of continuous time strategies.

## References

- Anderson, Robert M.** 1984. "Quick Response Equilibria." University of California, Berkeley, C.R.M. I.P. No. 323
- Andreoni, James., and John Miller.** 1993. "Rational Cooperation in the Repeated Prisoner's Dilemma: Experimental Evidence." *Economic Journal*, 103: 570–585.
- Berbey-Meyer, Yoella., and Alvin E. Roth.** 2006. "The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation." *American Economic Review*, 96(4): 1029–1042.
- Bergin, James., and Bentley MacLeod.** 1993. "Continuous Time Repeated Games." *International Economic Review*, 34(1): 21–37.
- Cooper, Russel, Douglas V. DeJong, Robert Forsythe, and Thomas W. Ross.** 1996 "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games." *Games and Economic Behavior*, 12(2): 187–218.
- Dal Bo, Pedro.** 2005 "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games." *American Economic Review*, 95(5): 1591–1604.
- Dal Bo, Pedro., and Guillaume R. Frechette** 2008 "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence" Working Paper.
- Flood, Merrill M.** 1952. "Some Experimental Games." RAND Research Memorandum, RM-789-1.
- Fudenberg, Drew., and Eric Maskin.** 1986. "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information." *Econometrica*, 54: 533–554.
- Hauk, Esther and Rosemarie Nagel.** 2001 "Choice of Partners in Multiple Two-Person Prisoner's Dilemma Games: An Experimental Study." *Journal of Conflict Resolution*, 45(6): 770–793.
- Huberman, Bernardo A. and Natalie S. Glance.** 1993. "Evolutionary Games and Computer Simulations." *Proceedings of the National Academy of Sciences*. 90: 7716–7718.
- Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson.** 1982. "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma." *Journal of Economic Theory*, 17: 245–252.

- Rapoport, Anatol and Albert M. Chammah.** 1965. "Prisoner's Dilemma: A Study in Conflict and Cooperation." Ann Arbor: University of Michigan Press.
- Roth, Alvin E. and Keith J. Murnighan.** 1978. "Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma." *Journal Mathematical Psychology*, 17(2): 189–198.
- Selton, Reinhard. and Rolf Stoecker.** 1986. "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach." *Journal of Economic Behavior and Organization*, 7: 47–70.
- Simon, Leo K. and Maxwell. B. Stinchcombe.** 1989. "Extensive Form Games in Continuous Time: Pure Strategies." *Econometrica*, 57: 1171–1214.

## Appendix

The following is the instructions from Continuous sessions. Grid and One-Shot instructions carefully match Continuous ones, deviating only where necessary, and are available upon request.

### Instructions (C)

Welcome. This is an experiment in the economics of decision-making. If you pay close attention to these instructions, you can earn a significant amount of money that will be paid to you in cash at the end of the experiment.

### The Basic Idea

A	<b>10, 10</b>	<b>0, 14</b>
B	<b>14, 0</b>	<b>6, 6</b>
	a	b

In each of several periods you will be randomly matched with a counterpart, a person in this room. Each period you will choose one of two actions: “A” or “B”. Your counterpart will choose “a” or “b”.

Your earnings will depend on the combination of your action and your counterpart’s action. These earnings possibilities will be represented in a matrix like the one above. Your action will determine the row of the matrix (A or B) and your counterpart’s will determine the column of the matrix (a or b). The cell corresponding to this combination of actions will determine earnings as follows. In each cell are two numbers. The first of the two numbers (shown in bold) is your earnings from this action combination. The second number is your counterpart’s earnings.

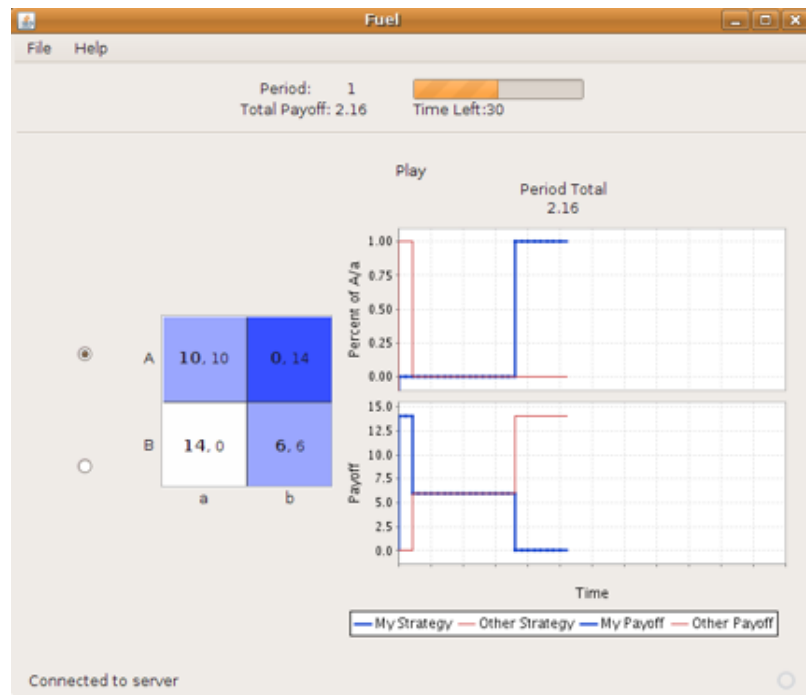
For example, in the sample matrix above, if you chose A and your counterpart a, you would earn 10 points and your counterpart 10. If instead your counterpart chose b, you would earn 0 points and your counterpart would earn 14.

When you start a new period, you will be randomly matched with a different counterpart. We do our best to ensure that you and your counterpart remain anonymous.



The payoff matrix in a new period is not always the same as in the previous period, so you should always look at it carefully at the beginning of the period.

### Details on How to Play



Periods will each last 60 seconds and the clock at the top of the screen will show how much time is left. Each period, the computer will start you and your counterpart off with randomly chosen actions. At any moment during the period, you will be able to change your action by clicking the two radio buttons (or by using the arrow keys). The row corresponding to your chosen action and the column corresponding to your counterpart's action will both be highlighted in blue. You and your counterpart may change your decision as many times as you want over the course of the period.

The numbers in the payoff matrix are the period payoffs you would earn if you maintained the same action throughout the period. For instance if you played B for the entire period and your counterpart b in the example above, you would earn 6 points and your counterpart 6 points.

In each period the payoffs depend only on how much time is spent in each cell of the payoff matrix. The more time you spend in any one cell, the closer final payoffs will be to the payoffs in that cell. For example, if you played A for half of the period and B for the other half while your counterpart played b for the entire period, then your earnings would be  $(1/2)0 + (1/2)6 = 3$  and your counterpart's earnings would be  $(1/2)14 + (1/2)6 = 10$ . This is because you spent half of the period

in the upper left cell and half the period in the lower right cell of the payoff matrix.

To the right of the payoff matrix are two graphs showing outcomes over the course of the period. The top graph shows your action (in blue) and your counterpart's action (in red) so far during the period. The graph is labeled "Percentage of A." If this is at 100% it means that at that moment you are using action A. If it is 0% it means at that moment you are using action B. The graph tracks any changes you and your counterpart make during the period.

The bottom graph shows your earnings over the course of the period in blue. On the horizontal axis is your payoff and on the vertical axis is the percentage of time elapsed. The more area below your earnings curve, the more you have earned. The red line on this graph shows the corresponding payoffs for your counterpart.

### **Earnings**

You will be paid at the end of the experiment based on the sum of point earnings throughout the experiment. These total earnings are displayed as the Total Payoff at the top of the screen. The conversion rate from points to US dollars is written on the whiteboard.

### **Frequently Asked Questions**

Q1. Is this some kind of psychology experiment with an agenda you haven't told us?

Answer. No. It is an economics experiment. If we do anything deceptive or don't pay you cash as described then you can complain to the campus Human Subjects Committee and we will be in serious trouble. These instructions are meant to clarify how you earn money, and our interest is in seeing how people make decisions.

Q2. If I choose the rows and my counterparty chooses the columns, does his (or her) screen look different than mine?

Answer. On his or her screen, the same choices are shown as rows. For example, if s/he chooses row B then it shows up on your screen as a choice of the b column. Of course, the payoff numbers are the same on both screens, just shown in a different format.