## TIME HORIZON AND COOPERATION IN CONTINUOUS TIME

MARIA BIGONI
*University of Bologna, Bologna, Italy*

MARCO CASARI
*University of Bologna, Bologna, Italy and IZA*

ANDRZEJ SKRZYPACZ
*Graduate School of Business, Stanford University, Stanford, CA 94305, U.S.A.*

GIANCARLO SPAGNOLO
*SITE–SSE, Stockholm, Sweden, University of Rome Tor Vergata, EIEF, and*
*CEPR*

# TIME HORIZON AND COOPERATION IN CONTINUOUS TIME

By Maria Bigoni, Marco Casari,
Andrzej Skrzypacz, and Giancarlo Spagnolo[1]

We study social dilemmas in (quasi-) continuous-time experiments, comparing games with different durations and termination rules. We discover a stark qualitative contrast in behavior in continuous time as compared to previously studied behavior in discrete-time games: cooperation is easier to achieve and sustain with deterministic horizons than with stochastic ones, and end-game effects emerge, but subjects postpone them with experience. Analysis of individual strategies provides a basis for a simple reinforcement learning model that proves to be consistent with this evidence. An additional treatment lends further support to this explanation.

Keywords: Prisoner's dilemma, experiment, folk theorem, backward induction.

## 1. INTRODUCTION

IN THIS PAPER, we present results from laboratory experiments on dynamic social dilemmas played in (almost) continuous time. We discover behavior that is qualitatively different from what has been documented in social dilemmas played using discrete-time protocols.

Our experiment combines elements of the design of Dal Bó (2005) and of Friedman and Oprea (2012). In particular, we study cooperation patterns in (almost) continuously repeated prisoners' dilemmas by comparing deterministic versus stochastic termination rules for different lengths of the horizon. Our first main finding is that in short games, players cooperate more when the horizon is deterministic than when it is stochastic with the same expected duration. This pattern is opposite to that observed by Dal Bó (2005) in discretely repeated games. The second main finding is that while in games with deterministic horizon there is a clear end-game effect of the form that cooperation drops significantly in the last few seconds of each supergame, with experience subjects cooperate longer. This is also in contrast to what has been documented

in experiments with discrete rounds, where with experience subjects learn to defect sooner (see Embrey, Fréchette, and Yüksel (2013) for a comprehensive review of this literature and the latest evidence).

We are interested in studying games in continuous time because in many field situations actors can react to opponents' actions quickly. Examples include firms posting prices on the Internet or via a centralized and transparent marketplace (as airlines), financial markets with high-frequency/computerized trading, electricity markets with high-frequency bidding, workers in a team choosing effort, nearby restaurants choosing menus, and spouses sharing everyday chores. In such situations, the trade-off between immediate gains from a deviation and the continuation punishment could be of second order. Since that trade-off is at the center of discrete-time repeated games models and the vast experimental literature that followed them, we should be cautious in applying the lessons learned from these models and experiments to situations where reactions can be fast and asynchronous.

While the inherent difference between continuous-time and discrete-time games has been recognized in theoretical work, we offer the first empirical evidence that, indeed, qualitatively different patterns of behavior emerge in these environments. Our study provides an empirical foundation for using different models for discrete- and continuous-time games. Extrapolating from theories and experiments in discrete time to draw conclusions on situations with high-frequency interactions appears misguided in light of our findings that comparative statics in continuous time result in the opposite patterns to what is observed in discrete time. And since dynamic social dilemmas with fast reactions are relatively understudied (despite being common in the field), our results imply that we need more theoretical and empirical research specifically on them.

We suggest the following interpretation for our results. Under a deterministic horizon, standard models of repeated games in discrete time predict low cooperation rates in the last periods and unraveling of cooperation due to backward induction. While in experiments subjects rarely follow the predictions of full backward induction, typically with experience, subjects learn to defect sooner. In contrast, continuous-time games do not have "the last period" and, hence, unraveling arguments do not apply. Because counterparts can react quickly, immediate gains from deviation are small compared to the shadow of the future punishment. That provides a possible key of interpretation for our second main finding: while we do see some end-game effects in our deterministic-horizon treatments, we also see the opposite of unraveling with experience, which is consistent with the logic of continuous-time theories.

In discrete-time games, the unraveling argument implies that cooperation rates should be higher in games with stochastic than with deterministic horizon, especially in games with short duration where the horizon plays an important role. This prediction has been confirmed by Dal Bó (2005). At first our discovery of the opposite behavior is quite surprising, since the lack of unraveling under a deterministic horizon would lead at most to the prediction that

cooperation levels in deterministic and stochastic horizons should be similar (a prediction that is confirmed in our longer treatments).

To shed light on the causes of this unexpected behavior, we analyze individual strategies and their evolution across supergames. In all treatments, subjects change their strategies from defection to conditional cooperation as they gain experience. However, we find that the speed of this transition differs across treatments and is particularly slow in short stochastic treatments, which justifies the overall difference in cooperation rates. We then show that such a pattern of behavior is consistent with a simple reinforcement learning model: due to the unpredictability and variability of game duration, subjects receive particularly noisy signals about the relative profits of these types of strategies in short games with stochastic ending and, hence, it takes longer for the low-payoff strategies to be abandoned.

Our experiments have not been designed to discriminate among different learning models. Yet we ran one additional treatment that tries to isolate the effects of unpredictability of supergame duration from the variability of duration alone. In that treatment (see Section 7), subjects played games with deterministic duration that varied across supergames and matched exactly the realized durations in our short stochastic treatment. While both factors matter, we observe that the unpredictability of supergame durations slows down the convergence to cooperation more than their variability.

The next section reviews the related literature; Section 3 describes the experimental design; Section 4 presents our empirical results; Section 5 empirically identifies individual strategies employed; Section 6 shows the impact of experience and provides a model for the evolution of behavior; Section 7 presents additional empirical evidence; Section 8 discusses the theoretical background in more detail; Section 9 concludes.

## 2. RELATED LITERATURE

The repeated (or "iterated") prisoner's dilemma with perfect monitoring has probably been the most important setup in which the question of what leads people to cooperate has been explored experimentally since the early work of Rapoport and Chammah (1965). A central and highly debated issue has been the role played by the time horizon, sometimes called the *termination rule*. The experimental literature has shown that the theoretical prediction that backward induction should apply to finitely repeated games with the features of a prisoner's dilemma often does not hold in the laboratory.[2] In field situations, the moment at which a relationship will come to an end is often uncertain. To capture this feature, several researchers, starting with Roth and Murnighan (1978) and Murnighan and Roth (1983), have tried to reproduce an indefinite,

---

[2]See, for example, Selten and Stoecker (1986), Andreoni and Miller (1993), Cooper, DeJong, Forsythe, and Ross (1996), Hauk and Nagel (2001), and Bereby-Meyer and Roth (2006).

uncertain horizon in the lab under a stochastic continuation/termination rule for the repeated game. Selten, Mitzkewitz, and Uhlich (1997) argued against the attempt to replicate a potentially infinite horizon in the lab, since no real experiment can have infinite duration, so subjects will be aware that the experiment will end in a reasonable amount of time and their beliefs may vary about when exactly. Based on previous experimental evidence (e.g., Selten and Stoecker (1986)), they proposed using finitely repeated games, given that the outcomes of repeated laboratory games with deterministic and stochastic horizons are similar, apart from the end-game effect that only takes place in the last rounds. Dal Bó (2005) offered experimental evidence against this last conclusion. He ran repeated prisoner's dilemma games with two different parameterizations of the stage-game payoffs and with deterministic and stochastic horizons with identical but short expected durations. Among other things, he found that cooperation rates in both the first and last rounds of the supergames are significantly lower in treatments with a deterministic horizon. Normann and Wallace (2012) also compared these termination rules (as well as a third, "unknown termination"), but in a different setup where the prisoner's dilemma is repeated 22 times before the different termination rules are introduced, finding instead no significant differences in cooperation rates.[3]

Friedman and Oprea (2012) is the first experiment on a prisoner's dilemma in continuous time. In their experiment, subjects can switch actions frequently (with latency times on the order of 0.02 seconds) and the game has a deterministic horizon of 60 seconds. As in our experiments, they observe high cooperation rates (a median between 81% and 93% after subjects gained experience) and, within a supergame, cooperation is typically sustained until the very last seconds of the game, when a short but drastic end-game effect takes place. They also compare cooperation rates for different frequencies of moves varying from a one-shot game to eight moves per supergame to one move per second. They find that games with higher frequency of moves have higher cooperation rates and that this effect is quantitatively large: the eight-moves-per-supergame treatments achieve less than half of the cooperation levels observed in the continuous-time treatment, providing support for theories that predict that a smaller per-period gain from defection (relative to continuation payoffs) stabilizes cooperation.[4]

---

[3]Notice, however, that Normann and Wallace do not allow subjects to gain experience, as they only play one supergame. See also Palfrey and Rosenthal (1994), who compared contributions to a public good in one shot versus indefinitely repeated games. Engle-Warnick and Slonim (2004) report little difference when comparing a trust game repeated exactly five times versus repeated with a continuation probability of 0.8. Aoyagi and Fréchette (2009) instead report results for a treatment with continuation probability of 0.9 and one with one-shot games, and report large differences in cooperation.

[4]Charness, Friedman, and Oprea (2011) ran a four-person public good experiment in continuous time and reported a somewhat lower impact of continuous-time interaction on cooperation.

Our paper differs from Friedman and Oprea (2012) because we show qualitative differences between continuous-time and discrete-time treatments. Their results are consistent with continuous-time games being a smooth limit of discrete-time games with higher and higher frequency of interaction (see their Result 5). By contrast, our experiment shows that in some crucial dimensions, the opposite patterns of behavior arise in discrete-time and continuous-time environments, especially with short durations. In terms of design, we differ from Friedman and Oprea (2012) in that we use Dal Bó (2005) as the discrete-time benchmark and run five different continuous-time designs (prisoner's dilemma games under a deterministic and a stochastic time horizon for two different durations and the new treatment with deterministic horizon with duration that varies across supergames). Less important differences include stage-game payoffs, the protocol to match subjects across supergames, and the starting action in each supergame (which was random in Friedman and Oprea (2012) and chosen by the subject in the present study).

Our work is also related to experimental studies of finitely repeated games played in discrete time at low frequency that, among other things, investigate whether experience induces subjects to cooperate more and to properly apply backward induction. A consistent finding in this literature, including Selten and Stoecker (1986), Andreoni and Miller (1993), Hauk and Nagel (2001), and Bereby-Meyer and Roth (2006), among others, is that subjects learn to cooperate more at early stages of the finitely repeated game and less at the final stages. Subjects also seemed to learn to anticipate the end-game effect, consistent with a partial learning backward induction story. The recent meta-analysis of previous work and additional experiments by Embrey, Fréchette, and Yüksel (2013) confirm that this is a consistent result in discretely repeated games experiments with a deterministic horizon: subjects tend to learn to anticipate the beginning of the end-game phase. In our continuous-time environment, we find precisely the opposite result.

Finally, the experimental literature on games in continuous time has blossomed during the last few years, so there are several less related studies focusing on strategic situations that are quite different from a prisoner's dilemma, such as games of network formation (Berninghaus, Ehrhart, and Ott (2006), Berninghaus, Ehrhart, Ott, and Vogt (2007)), minimum effort games (Deck and Nikiforakis (2012)), and hawk–dove games (Oprea, Henwood, and Friedman (2011)).

## 3. EXPERIMENTAL DESIGN

The experiment has five treatments. The two treatment variables are the expected duration of each supergame and the termination rule.[5]

---

[5]Table B.I in the Supplemental Material (Bigoni, Casari, Skrzypacz, and Spagnolo (2015)) summarizes the characteristics of each treatment.

TABLE I

STAGE GAME PAYOFFS[a]

|  | Cooperate | Defect |
|---|---|---|
| Cooperate | 1, 1 | −2, 2 |
| Defect | 2, −2 | 0, 0 |

[a]The numbers in each cell represent the payoff per second.

In all treatments, subjects played a series of (quasi-) continuous-time prisoner's dilemma games with stage-game payoffs as in Table I.[6]

Each session comprised a non-overlapping group of 24 subjects, who interacted in pairs for 23 supergames. Pairs were formed so that each subject met all the others once and only once in a session (round-robin matching).[7]

In all treatments, the stage game was as follows. Each subject had to select an initial action for the supergame, either cooperate (green) or defect (orange). When all subjects were done, the supergame began. Within a supergame, subjects could switch action up to six or seven times per second. More precisely, there was a tick every 0.16 of a second, which gave the participants the feeling of continuous time. The PCs had touch screens; hence, a switch of action could not be heard by others, as subjects simply touched the screen with a finger.[8]

Earnings for all possible combinations of actions were visible on the screen at all times. The payoff matrix showed earnings in tokens per second. The subject's current action was always highlighted in yellow in the payoff matrix. Moreover, every subject could observe her cumulative earnings on a continuously updated graph. Subjects' earnings in every supergame included an initial endowment (see Table B.I in the Supplemental Material), and could stay constant, increase, or decrease over time, depending on the choices of the pair. The graph showed these patterns of earnings as a flat, increasing, or decreasing line, respectively. A steeper line indicated a faster accumulation or depletion. The line color was green or orange depending on the subject's own action. Hence, from the graph, subjects could unambiguously infer the action taken in any moment by their opponent. The progression of the earnings line marked the timing of the supergame for the subjects. They could observe at every in-

---

[6]As the instructions explained, the experiment was in quasi-continuous time: "Within a period, both you and the other will be able to change action as many times as you wish. The time flows in very rapid ticks (of 0.16 of a second); in practice, there are between six and seven ticks every second, so that if you wish you can change action six or seven times per second." For brevity, from now on we will refer to it as a continuous-time experiment. Notice that in the instructions, supergames were referred to as "periods."

[7]In the Short-Deterministic session run on February 2, 2011, due to a technical problem, subjects met their opponents of supergame 1 again in supergame 23. All reported results hold even if supergame 23 in that session is dropped.

[8]Figure B.1 in the Supplemental Material presents a screen shot of the computer interface.

stant the speed of the game, which ran at the same pace for all subjects in the session. For the Deterministic treatments, subjects could always check the time remaining before the end of a supergame by looking at the graph on the screen.

In the Long-Deterministic treatment, a supergame always lasted 60 seconds. In the Long-Stochastic treatment, a supergame lasted in expectation 60 seconds. Similarly for the Short treatments, where the expected duration was 20 seconds. In the stochastic treatments, the exact duration was selected at random supergame by supergame. As explained in the instructions for the Long (Short)-Stochastic treatment, the supergame duration depended on a random draw. "Imagine a box with 10,000 (1000) balls, of which 9973 (992) are black and 27 (8) are white. It is as if a ball is drawn after every tick. If the ball is white, the period ends. If the ball is black, the period continues and the ball is put back into the box. At the next tick, another ball is drawn at random. You have to imagine very fast draws, i.e., one every tick of 0.16 of a seconds. As a consequence of this procedure, we have estimated that periods will last, on average, 60 (20) seconds. There may be periods that are short and periods that are long." In case a supergame (which is called a period in the instructions) lasted beyond 60 seconds, the scale of the horizontal axis of the graph automatically shifted forward. The experimenters had no direct control and no ex ante information on the realized supergame durations.

In the Variable-Deterministic treatment, supergame ending was deterministic but supergames duration was variable. The sequence of supergame durations was calibrated to match exactly the realized durations in the Short-Stochastic treatment to allow for a tight comparison. At the beginning of each supergame, the current supergame duration was disclosed to the subjects both numerically—in terms of seconds—and graphically through a vertical line drawn in the payoff graph on the subjects' computer screen. Stochastic treatments are different from Deterministic treatments both because the ending of the supergame is random and because supergame durations are variable. The specific goal of this additional treatment is to understand which one of these factors has more impact on behavior.

Stage-game payoffs are such that cooperation should be easily achieved (at least in the stochastic ending treatments). In continuous time, cooperation is always sustainable because the instantaneous discount factor is 1: then a Grim-trigger strategy should, in theory, support cooperative play as an equilibrium, regardless of the arrival rate of the end of the game. But even if agents perceived the game to be played discretely, for example, because of minimal human reaction time, cooperation should be easily sustained with our parameterization. For example, if subjects react with a 1 second delay and treat it as a time interval length of 1 second, then given our stage-game payoffs (see Table I), cooperation can be sustained with an infinite horizon for discount factors higher than 1/2, which implies an expected duration of 2 seconds. If the time interval length is 0.25 of a second, then it would be enough to have an expected duration of 0.5 of a second and so on. Hence, the 20 seconds is quite far from the theoretical bound.

Instructions were distributed and then read aloud.[9] Subjects had the opportunity to ask questions, which were answered in private, and then went through three practice supergames with a robot opponent that was programmed to switch action in the middle of the supergame. After each practice supergame, subjects had to guess the actions taken by the robot and then complete a computerized quiz to verify their full understanding of the rules of the game. The experiment started as soon as all subjects answered correctly to all control questions.[10] The session ended with a questionnaire.

The experiment involved 240 subjects, mostly students at the University of Bologna, who took part in only one of the ten sessions and were assigned through an online recruitment software (Greiner (2004)). The experiment was run in the Bologna Laboratory for Experiments in Social Sciences using z-Tree (Fischbacher (2007)). Subjects were seated at visually isolated computer terminals and could not communicate. A session lasted, on average, 2 hours for the Long treatments and 1 hour and 20 minutes for the Short and Variable ones. At the end of the sessions, points were converted into euros at a rate of 1€ for every 150 points in the Long treatments, and at a rate of 1€ for every 50 points in the Short and Variable ones. Subjects earned, on average, 16.72€ and 15.47€, respectively, which include a show-up fee of 3€.

## 4. RESULTS

With our Long-Deterministic treatment, we replicate the results reported in Friedman and Oprea (2012). The median cooperation rate from supergame 13 on in Friedman and Oprea (2012) ranges between 81% and 93%, depending on the treatment, and in our Long-Deterministic treatment it is 91%. This provides a robustness check of their findings for different procedures, subject pools, and payoff levels. The novelty of this study, however, stems from the comparison across our treatments.

Cooperation rates are higher in supergames of longer (expected) duration, with the difference being large for stochastic ending and small for deterministic ending (Tables II and III). The impact of duration on cooperation rates is highly significant in the stochastic treatments, while being only marginally significant in the deterministic treatments. The unit of observation is the cooperation rate, which is defined as the fraction of time $R_{i,s}$ a subject $i$ spends cooperating within supergame $s$. Given that these observations are not independent, to assess the significance of the observed differences we take the

---

[9]An English translation of the experimental instructions is available in Appendix C in the Supplemental Material.

[10]In the three practice supergames, 71% of the subjects always made correct guesses about the sequence of actions taken by the robots. In answering the four control questions about the instructions, 53.8% of the subjects made at most one mistake.

TABLE II

COOPERATION RATES BY TREATMENT[a]

| Duration | | Termination Rule | | |
| | | Deterministic | | Stochastic |
| --- | --- | --- | --- | --- |
| Long | Mean | 65.5 | $\sim$ | 66.9 |
| | *Median* | *84.0* | | *84.8* |
| | | $\vee^{**}$ | | $\vee^{***}$ |
| Short | Mean | 63.3 | $>^{***}$ | 52.3 |
| | *Median* | *79.2* | | *47.0* |

[a]The mean cooperation rate of a session is the average across all 23 supergames and all 24 subjects. The unit of observation is a subject per supergame. For every treatment there are two sessions and 1104 observations. In this and in the following tables, the symbols *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Significance levels are derived from the regression presented in Table III.

average cooperation rate by subject across all supergames and run a linear regression with bootstrapped standard errors. Results are reported in Table III.[11] The outcome of this regression indicates that these differences are significant when controlling for individual characteristics.[12]

Let us now turn to our main result.

RESULT 1: With deterministic duration, cooperation rates are equal or higher than with stochastic duration.

Support for Result 1 comes from Tables II and III. In the long duration treatments, cooperation rates are statistically indistinguishable between stochastic

[11]We obtain similar results with a panel regression with random effects at the subject level, where the unit of observation is the cooperation rate of a subject in a supergame, and standard errors are robust for heteroscedasticity (see Table B.II in the Supplemental Material). As a further robustness check, we also ran linear regressions with standard errors robust for clustering at the subject and pair level. The same treatment effects emerge if we compare the rates of mutual cooperation (Tables B.III and B.IV in the Supplemental Material) or the average profits per second (Tables B.V and B.VI in the Supplemental Material). Notice, however, that the difference between the Long- and Short-Deterministic treatments ceases to be statistically significant if we drop controls for individual characteristics, if we focus only on the last five supergames, or if we consider only the first 2/3 of each supergame.

[12]In this and in the following regressions, we control for a number of factors: (i) demographics—age, gender, field and degree of study, occupation, and Italian nationality (93.3% subjects); (ii) task comprehension—number of wrong answers to control questions and the total answering time; (iii) academic background—three dummies taking value 1 for subjects who have previously followed courses of economics, statistics, and game theory, respectively; (iv) nonincentivized questionnaire measures—risk attitudes, level of generalized trust, and two IQ-type questions. A full version of this regression, including the estimated coefficients for all the controls, is reported in Table B.II in the Supplemental Material.

TABLE III

LINEAR REGRESSION ON COOPERATION RATES[a]

| *Dependent Variable: Cooperation Rate* | | |
|---|---|---|
| | Coefficient | (s.e.) |
| Short-Deterministic | −5.537** | (2.760) |
| Long-Stochastic | 1.682 | (2.577) |
| Short-Stochastic | −17.799*** | (3.830) |
| Constant | 63.798*** | (14.128) |
| Controls for individual characteristics | Yes | |
| $N$ | 192 | |
| $R$-squared | 0.220 | |

[a]Linear regression with bootstrapped standard errors. To take care of the potential correlation between observations coming from the same session, we take sessions as resampling clusters. The unit of observation is the fraction of time a subject spends cooperating within a supergame, averaged by subject across all supergames. Default treatment: Long-Deterministic. See footnote 12 for a list of controls of individual characteristics.

and deterministic horizons ($p$-value $> 0.1$; see Table III). By contrast, in the short duration treatments, cooperation rates are significantly higher with a deterministic horizon than with a stochastic horizon ($p$-value $< 0.001$; see Table III). The absolute difference in cooperation between the two treatments is 11.0 points in terms of means and 32.2 points in terms of median. This result is in stark contrast with experiments on repeated games in discrete time, where cooperation is typically higher with stochastic than with deterministic duration.[13]

The next result shifts the focus on the dynamics within each supergame, as the same cooperation rate $R_{ip}$ can result from different sequences of actions, especially in continuous-time games. The evidence suggests that subjects do not learn to apply backward induction.

RESULT 2: With deterministic duration, end-game effects exist, do not unravel cooperation, and appear later with experience.

Support for Result 2 comes from Figure 1 and Table IV. Figure 1 presents the time profile of the mean share of cooperators, taken across supergames and sessions. A subject can change action every 0.16 seconds. Our unit of observation is the share of cooperators $S_{tp}$ over time $t$ within a supergame $p$.

---

[13]For example, in a repeated game with short expected duration, Dal Bó (2005) finds that "for every round, [...] the percentage of cooperation in infinitely repeated games [...] is greater than in finitely repeated games of the same expected length [...], with $p$-values of less than 0.01." (The expected number of action choices is 125 in our Short treatments and 375 in our Long treatments, while it ranges between 2 and 4 in his treatments.) More specifically, when the expected duration is two (four) periods, the average cooperation rate is 28.3% (35.2%) with stochastic ending and 12.5% (24.8%) with deterministic ending.
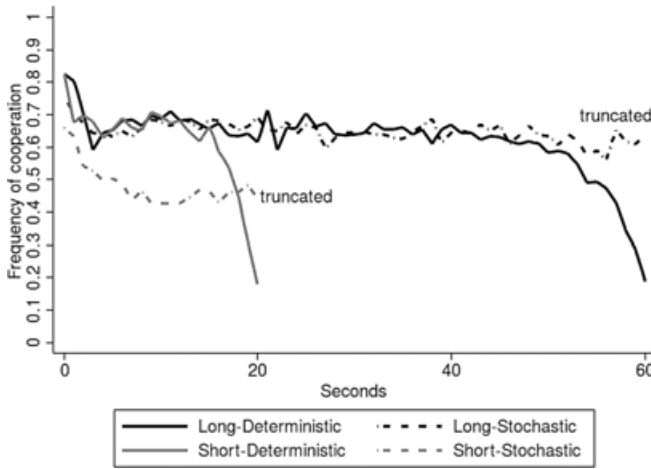
FIGURE 1.—Time profile of the share of cooperative actions. *Notes*: The unit of observation is the share of cooperative actions in every second of a supergame. A subject could switch action every 0.16 of a second. All subjects and all supergames are included for the first 20 or 60 seconds. In the Long-Stochastic treatments, 45.7% of supergames lasted more than 60 seconds. In the Short-Stochastic treatments, 30.4% of supergames lasted more than 20 seconds.

In both the Short- and Long-Deterministic treatments, there is a clear end-game effect: the share of cooperators suddenly drops a few seconds before the end of the supergame (Figure 1).[14] There are, of course, many ways to quantitatively measure the timing of this switch from cooperation to permanent defection. We measured it by focusing on all pairs that at some point during

TABLE IV

TIMING OF THE END-GAME EFFECT[a]

|  | Supergames | | | | |
| Treatment | 1–6 | 7–12 | 13–18 | 19–23 | Overall |
| --- | --- | --- | --- | --- | --- |
| Long-Deterministic | 17.7 | 11.5 | 11.4 | 7.1 | 11.9 |
| Short-Deterministic | 7.4 | 5.3 | 4.1 | 3.8 | 4.9 |

[a]Average distance (in seconds) between the end of the supergame and the time of a permanent switch of a pair from mutual cooperation $(C, C)$ to defection, that is, to $(C, D)$, $(D, C)$, or $(D, D)$.

[14]Friedman and Oprea (2012) also report an end-game effect. They find that "cooperation level falls below 75 percent only when 5 seconds remain and below 50 percent only when 1 second remains."

a supergame reached simultaneous cooperation (C, C) and then switched to defection before the end of the supergame, that is, (C, D), (D, C), or (D, D).[15]

It is evident from Table IV that in the Short-Deterministic treatment, the end-game effect kicks in significantly later than in the Long-Deterministic treatment: about 5 seconds from the end in contrast to about 12 seconds (Table IV; $p$-value $< 0.001$).[16] It is also evident that the end-game effect kicks in later and later, as subjects gain experience (3.6–10.6 seconds later; Table IV). The impact of experience is significant both in the Long-Deterministic ($p$-value $< 0.001$) and in the Short-Deterministic treatment ($p$-value $< 0.001$).

One reason behind the postponing of the end-game effect may be that subjects become faster in reacting to defections as they gain experience. Indeed, across supergames, we observe a decrease in reaction time, which is measured as the time interval between a deviation from mutual cooperation and the beginning of the punishment phase, that is, the time between the moment we see a switch from (C, C) to (C, D) until the moment we see a switch from (C, D) to (D, D).

The correlation between reaction times and timing of the end-game effect, however, is not significant (Table B.VII in the Supplemental Material). When controlling for the average reaction time in the regression, the decrease in the duration of the end-game effect across supergames is still significant (see Table A.I).

## 5. IDENTIFICATION OF INDIVIDUAL STRATEGIES

Studying strategy adoption may shed light on the differences in aggregate levels of cooperation across treatments, in particular the low cooperation rates in the Short-Stochastic treatment.

The adopted approach to strategy estimation is based on the assumption that subjects (i) adopt one of the 20 strategies presented in Table V, (ii) may change strategy across supergames, and (iii) may make errors on actions, that is, with some probability may choose an action that is not recommended by the strategy.[17] We considered all strategies analyzed in Fudenberg, Rand, and Dreber (2012), which include, in particular, three sets: (A) fully cooperative strategies, which prescribe cooperation at all points in time if matched with each other;

[15]This calculation includes the lion's share of the observations. Out of a total of 552 subject–supergame observations per treatment, we have 468 and 460 in the Long and Short treatment, respectively, such that both subjects cooperated simultaneously at least once in that supergame. Of these, some (54/468 and 51/460, respectively) kept on cooperating until the end of the supergame, while in the other cases (414/468 and 409/460, respectively) at least one of the subjects in the pair switched to permanent defection.

[16]The $p$-values reported in this paragraph are obtained from panel regressions with random effects at the session level. The unit of observation is a session in a supergame. Regression's results are reported in Table A.I in Appendix A.

[17]Table A.III in Appendix A presents a description of the 20 strategies.

(B) fully defective strategies, which prescribe permanent defection if matched with each other; and (C) other strategies that would prescribe a combination of cooperation and defection. Notice that Table V does not include a Cut-Off strategy, that is, a strategy that prescribes to play Grim-trigger until close to the end of the supergame, then switch to permanent defection. The Cut-Off strategy is particularly relevant in deterministic treatments, as it could account for the aforementioned end-game effect (Table IV). The reason we exclude it is that including it would have required us to arbitrarily set the cutoff time. As a consequence, subjects following these strategies will be most likely classified as Grim-trigger, because the two strategies differ by at most one action (the first defection). For strategy estimation, we adopted a methodology introduced in Dal Bó and Fréchette (2011) for repeated games in discrete time.[18] To apply this methodology to our continuous-time setup, we considered subjects' actions in a sequence of specific instants of play (every second) and disregarded the rest of the data.[19]

For each of the strategies, we define the sequence of actions prescribed by the strategy to the subject, given the sequence of the subject's and the opponent's previous actions. The maximum likelihood procedure is based on the comparison between the actual sequence of actions taken by the subject and the 20 sequences corresponding to the 20 strategies. It returns an estimate of the frequency with which each strategy is played in the population.[20]

To study how strategy adoption evolved with experience, we carried out the maximum likelihood estimation over four blocks of supergames: 1–6, 7–12, 13–18, 19–23 (Figure 2).

RESULT 3: With experience, the frequency of Always Defect declines while the frequency of conditionally cooperative strategies increases. For all levels of experience and in all treatments, the combined frequency of these strategies is above 80%.

Support for Result 3 comes from Figure 2 and Table V: the share of subjects who adopt fully defective strategies (B) is initially high (between 25% and 41%) and sharply decreases with experience. The decrease, however, is

---

[18]The same estimation approach was also used, among others, by Fudenberg, Rand, and Dreber (2012), Camera, Casari, and Bigoni (2012), Dal Bó and Fréchette (2012), Fréchette and Yuksel (2013), Rand, Fudenberg, and Dreber (2013), Vespa (2013).

[19]Using the whole sequence of data, that is, one observation every 0.16 seconds, would have implied a computational burden that goes beyond the limits of the ICT resources available to us. We chose not to use the modal action in a second, because the mode does not necessarily preserve the temporal sequence of actions, which is crucial for strategy estimation. As a robustness check, we repeated the estimation by taking one action every 2 and every 3 seconds. Our main results are robust to these changes. We thank Guillaume Fréchette for kindly providing the original code for strategy estimation.

[20]We refer the reader to the on-line appendix of Dal Bó and Fréchette (2011) for a more detailed explanation of the estimation approach.

TABLE V

ESTIMATED STRATEGIES IN SUPERGAMES 19–23[a]

| Strategy | Long Deterministic Coefficient | s.d. | Long Stochastic Coefficient | s.d. | Short Deterministic Coefficient | s.d. | Short Stochastic Coefficient | s.d. |
|---|---|---|---|---|---|---|---|---|
| A. Fully cooperative strategies | **0.963** | | **0.888** | | **0.987** | | **0.814** | |
| Always Cooperate | 0.043** | 0.025 | 0.079* | 0.061 | 0.000 | 0.000 | 0.026* | 0.020 |
| *Conditionally cooperative strategies (A1 + A2)* | *0.920* | | *0.809* | | *0.987* | | *0.788* | |
| *A1. Forgiving cooperative strategies* | *0.338* | | *0.659* | | *0.247* | | *0.282* | |
| Tit-for-tat | 0.119*** | 0.038 | 0.203*** | 0.041 | 0.126* | 0.084 | 0.000 | 0.024 |
| Win-stay–lose-shift | 0.007 | 0.017 | 0.034 | 0.039 | 0.000 | 0.011 | 0.000 | 0.006 |
| T2 | 0.068*** | 0.029 | 0.049* | 0.038 | 0.063** | 0.036 | 0.039** | 0.023 |
| Tit-for-2-tats | 0.007 | 0.014 | 0.105** | 0.056 | 0.015 | 0.019 | 0.041*** | 0.016 |
| Tit-for-3-tats | 0.042* | 0.028 | 0.064** | 0.029 | 0.000 | 0.006 | 0.027* | 0.019 |
| 2-tits-for-tat | 0.037* | 0.026 | 0.047 | 0.048 | 0.042 | 0.040 | 0.087** | 0.046 |
| 2-tits-for-2-tats | 0.060** | 0.032 | 0.126** | 0.068 | 0.000 | 0.024 | 0.070 | 0.068 |
| Tit-for-tat-2 | 0.000 | 0.018 | 0.031** | 0.016 | 0.000 | 0.000 | 0.018 | 0.017 |
| *A2. Unforgiving cooperative strategies* | *0.582* | | *0.150* | | *0.740* | | *0.505* | |
| Grim-trigger | 0.495*** | 0.065 | 0.069** | 0.030 | 0.609*** | 0.158 | 0.350** | 0.211 |
| Lenient grim 2 | 0.036** | 0.020 | 0.041 | 0.047 | 0.105*** | 0.034 | 0.105*** | 0.042 |
| Lenient grim 3 | 0.051*** | 0.018 | 0.040*** | 0.016 | 0.026** | 0.016 | 0.050** | 0.027 |
| B. Fully defective strategies | **0.034** | | **0.071** | | **0.005** | | **0.152** | |
| Always Defect | 0.027 | 0.025 | 0.043*** | 0.016 | 0.000 | 0.006 | 0.075*** | 0.021 |
| Suspicious tit-for-tat | 0.008 | 0.009 | 0.028** | 0.014 | 0.005 | 0.012 | 0.077 | 0.075 |
| C. Other strategies | **0.002** | | **0.041** | | **0.008** | | **0.034** | |
| False cooperator | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.002 |
| Suspicious tit-for-2-tats | 0.000 | 0.000 | 0.021 | 0.033 | 0.000 | 0.002 | 0.000 | 0.000 |
| Suspicious tit-for-3-tats | 0.000 | 0.000 | 0.020 | 0.020 | 0.008 | 0.008 | 0.034 | 0.029 |
| Suspicious lenient grim 2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Suspicious lenient grim 3 | 0.002 | 0.006 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| D alternator | 0.000 | | 0.000 | | 0.000 | | 0.000 | |
| Gamma | 0.308*** | 0.022 | 0.348*** | 0.012 | 0.358*** | 0.028 | 0.335*** | 0.035 |
| Beta | 0.962 | | 0.946 | | 0.942 | | 0.952 | |

[a]The table reports the estimated frequency of each strategy in the population. In bold, we present the aggregate frequency for relevant families of strategies. The estimation methodology is described in the main text, while Table A.III in Appendix A presents a description of the strategies.
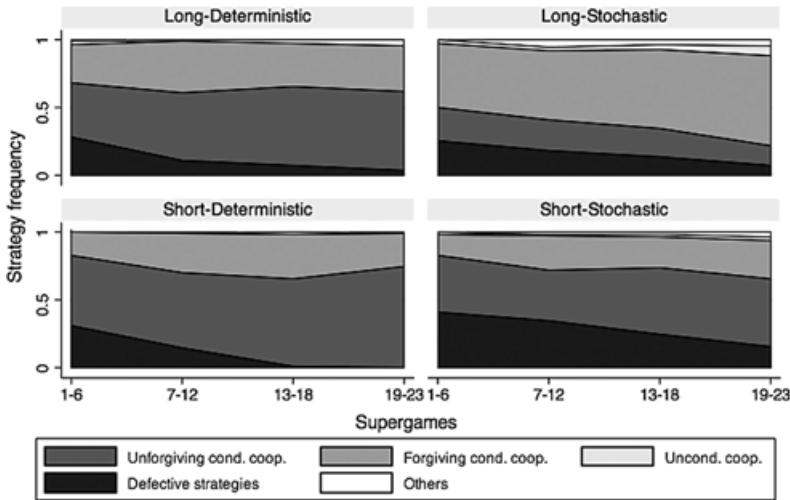
FIGURE 2.—Evolution of strategies across supergames.

slower for the Stochastic treatments and, in particular for the Short-Stochastic one. When considering only experienced play (supergames 19–23; Table V) the estimated frequency of fully defective strategies is substantially higher in the Short-Stochastic than in the other treatments (15% vs. 1–7%).

Table V also reveals that in the last five supergames, Grim-trigger is by far the most frequent strategy in the Deterministic treatments (50% in the Long-Deterministic and 61% in the Short-Deterministic; Table V). As previously mentioned, this probably indicates that in these treatments most subjects play Cut-Off strategies, which is in line with the analysis in Friedman and Oprea (2012).

Taken together, these results suggest that the difference in cooperation rates across treatments is to be found in a different dynamics of behavior, in particular, in the relatively slow reduction in the share of defective strategies in the Short-Stochastic treatment. In the next section, we put forward a reinforcement learning model that can explain why the take-off of cooperation is slower in the Short-Stochastic treatment than in the other three treatments.[21]

---

[21]Contrary to what Figure 2 might suggest, initial and average cooperation rates in the first supergame are higher in the Short-Stochastic treatment (60.4% and 40.7%, respectively) than in the other three treatments (50% and 25.4% in the Long-Deterministic; 54.2% and 35.4% in the Long-Stochastic; 52.1% and 24% in the Short-Deterministic). Hence, we believe that the higher prevalence of defective strategies estimated for the first block of six supergames in the Short-Stochastic treatment is determined by how play evolves with experience, rather than by the initial inclinations of the subjects.

## 6. DYNAMICS ACROSS SUPERGAMES

The analysis of strategies presented in the previous section shows the presence of a common trend in strategy adoption across all treatments, but this trend is characterized by different speeds. In particular, we observe that in all treatments, conditionally cooperative strategies tend to become more widespread with experience, even if this trend is slower in the Short-Stochastic treatment (Figure 2). In this section, we first present additional data about the diffusion of cooperative strategies with experience and then model its dynamics across supergames, showing that a simple reinforcement learning model fits our data very well. Finally, we provide an intuition for why learning to cooperate should be slower when supergame duration is stochastic and shorter in expectations: this intuition is based on the relatively lower profitability of cooperative strategies in very short supergames.

Results in Section 5 reveal that the vast majority of subjects played either Always Defect or one of the conditionally cooperative strategies (Result 3). The empirical approach adopted there estimates the prevalence of each strategy in the population; it does not associate a specific individual to a strategy. As a consequence, it does not allow us to assess the empirical profitability of each strategy. Such an assessment is crucial to estimate a reinforcement learning model that can explain the evolution of strategy adoption. To this purpose, here we employ the following simple approach. We classify all individuals into two non-overlapping categories: cooperators and defectors. The former category includes all subjects who start cooperating and never defect before their opponents or defect only during the last fifth of the (expected) duration of the supergame.[22] The latter category includes all other subjects. This classification is broadly consistent with the estimates in Table V: in supergames 19–23, cooperators account for 73–90% of the population, depending on the treatment.[23]

RESULT 4: The frequency of cooperators increases with experience in all treatments, but the increase is slowest in the Short-Stochastic treatment.

Support for Result 4 comes from Figure 3(a), which reports the frequency of cooperators. An upward trend emerges in all treatments, but this trend is weakest in the Short-Stochastic treatment.[24] Our results show that when play-

---

[22]The category of cooperators hence comprises also subjects adopting a Cut-Off strategy. A finer partition of the cooperators category is empirically troublesome because the same individual might be classified under multiple strategies. As Camera, Casari, and Bigoni (2012) show, cooperative strategies are frequently observationally equivalent.

[23]In Bigoni, Casari, Skrzypacz, and Spagnolo (2013), we present a similar reinforcement learning model based on initial actions, rather than on this strategy classification. Results are qualitatively the same.

[24]A similar trend characterizes initial and average cooperation rates (Figure B.2 in the Supplemental Material).
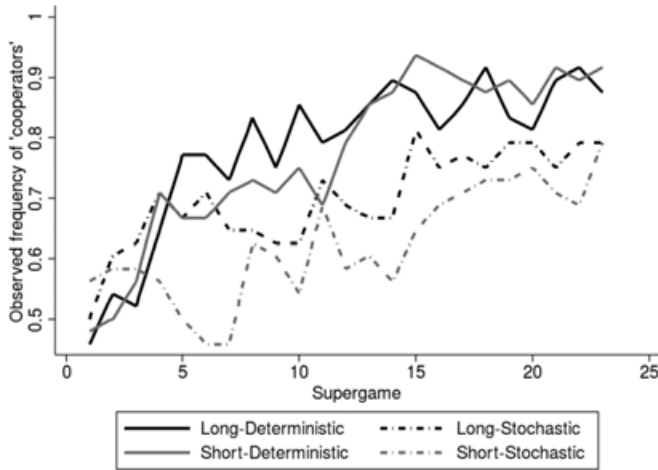
(a) Observed frequencies
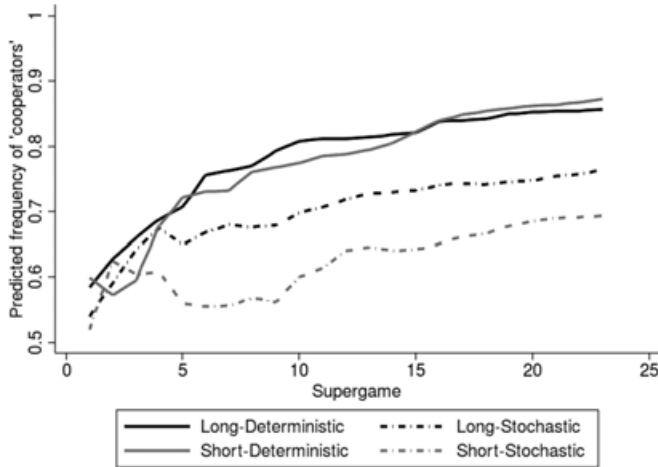


(b) Predicted probabilities

FIGURE 3.—Frequency of cooperators.

ing repeated games, the amount of experience is a critical determinant of out-comes. This is consistent with the findings of Dal Bó and Fréchette (2011) for discretely repeated games.

We now show that the observed dynamics of strategy adoption across su-pergames are consistent with a reinforcement learning model. Along the lines of Mookherjee and Sopher (1994, equation (2.1)), we estimate the probability $p_i^t$ that subject $i$ is a cooperator in supergame $t$ in the model

$$(1) \qquad p_i^t = f\big(\phi A_{i,C}^{t-1} + \chi A_{i,D}^{t-1} + \psi \mathbf{x}_i + \omega \mathbf{z}_i\big),$$

where $f(\cdot)$ denotes the logistic function, $A_{i,C}^{t-1}$ is the average payoff per second obtained by subject $i$ up to supergame $t-1$ when adopting a cooperative strategy, $A_{i,D}^{t-1}$ is the average payoff per second obtained by subject $i$ up to supergame $t-1$ when adopting a defective strategy,[25] $\mathbf{x}_i$ is a vector of three dummy variables corresponding to treatments Short-Deterministic, Long-Stochastic, and Short-Stochastic, and $\mathbf{z}_i$ is a vector of individual characteristics.[26]

Results estimating the model in equation (1) via a logit regression are presented in Table VI (Model 1). Figure 3(b) presents the predicted probability of adopting a cooperative strategy across treatments and supergames, estimated through Model 1 (equation (1)). This reinforcement learning model fits our data well; the estimated coefficients $\phi$ and $\chi$ have the expected sign, and are both highly significant.

Table VI (Model 2) also presents results from an alternative model, where the only independent variables are the individual characteristics and the treatment dummies ($\phi = \chi = 0$). The comparison between the two models in Table VI reveals that the observed differences in the rates of adoption of cooperative strategies across treatments are almost fully captured by the reinforce-

TABLE VI

PANEL REGRESSION ON SUBJECTS' STRATEGY ADOPTION[a]

*Dependent Variable: Subject's Strategy (1 = Cooperator, 0 = Defector)*

|  | Model 1 | | Model 2 | |
|---|---|---|---|---|
|  | Coefficient | (s.e.) | Coefficient | (s.e.) |
| $A_{i,C}^{t-1}$ | 3.043*** | (0.262) |  |  |
| $A_{i,D}^{t-1}$ | −0.753** | (0.311) |  |  |
| Short-Deterministic | −0.000 | (0.315) | −0.419 | (0.350) |
| Long-Stochastic | −0.072 | (0.275) | −0.645** | (0.286) |
| Short-Stochastic | −0.383 | (0.463) | −1.599*** | (0.430) |
| Constant | −1.670* | (1.008) | 0.332 | (0.667) |
| Controls for individual characteristics | Yes | | Yes | |
| $N$ | 4416 | | 4416 | |
| Log likelihood | −1859.435 | | −2081.673 | |

[a]Panel logit regression with random effects at the subject level and standard errors robust for heteroscedasticity. Default treatment: Long-Deterministic. Standard errors are reported in parentheses. The unit of observation is a subject's strategy (1 = cooperator, 0 = defector) in a supergame. See footnote 12 for a list of controls of individual characteristics. For a full version of this regression, see Table B.VIII in the Supplemental Material.

[25]Variables $A_{i,C}^{0}$ and $A_{i,D}^{0}$ are set equal to 0.

[26]To allow for correlation between observations belonging to the same subject, this vector also includes a dummy taking value 1 if, in the first supergame, the subject adopted a cooperative strategy, and 0 otherwise. We thank an anonymous referee for the suggestion.

ment learning terms ($A_{i,C}^{t-1}$ and $A_{i,D}^{t-1}$), whose inclusion into the model substantially improves the estimated log likelihood.[27]

Let us now provide some motivation for why it may be easier to experience the relative benefits of cooperation over defection with a deterministic rather than a stochastic horizon.

Consider for simplicity a situation where all subjects play—with no errors—either Always Defect or one of the conditionally cooperative strategies (described in Table V). No matter what the specific strategy is, when two conditional cooperators meet, they cooperate at all points in time, while if a conditional cooperator meets a defector, he will cooperate at the beginning and eventually switch to permanent defection.

There exist both empirical and theoretical reasons to consider a restriction to these two types of strategies meaningful. On the empirical side, the analysis of individual strategies in Section 5 shows that for all levels of experience and in all treatments, the combined frequency of these strategies is above 80% (Result 3). On the theoretical side, justifications for restricting the focus to this class of strategies may come from the model of Jehiel (2005) or from the model of Neyman (1985). Albeit different in underlying assumptions, the logic behind these two models can be applied to high-frequency/continuous-time games and can lead to very similar predictions. Jehiel (2005) shows that conditional cooperation can be an equilibrium strategy if strategies depend on a coarse set partition of all possible histories. For example, consider the following partition: set A contains all histories in which the opponent never defected (including the empty history) and set B contains all other histories. If both players use such two coarse analogies, the following strategies form equilibria for long games: (i) both players play Always Defect and (ii) both players follow a Grim-trigger, that is, cooperate if history is in set A and defect otherwise. Neyman (1985), instead, shows that a cooperative equilibrium exists in finitely repeated prisoner's dilemmas if subjects are bound to adopt strategies of limited complexity. More specifically, if players are restricted to use two-state automata (for instance Always Defect and Grim-trigger strategies), then they can sustain a cooperative equilibrium (unless the game is very short).

In practice, when player 1 follows Always Defect and player 2 uses a conditionally cooperative strategy, in continuous time player 2 does not respond with defection immediately, but there is some time period when one subject cooperates and the other defects. Let us denote this duration with $\kappa$. The length of $\kappa$ determines the profits of a subject playing Always Defect against an opponent playing one of the conditionally cooperative strategies. Let us denote by $\bar{\pi}(\kappa, H)$ the per-second expected profits from defection in a supergame, which

---

[27]In Model 1 there is no statistically significant difference between the coefficients of the Short-Stochastic and Short-Deterministic treatment dummies ($p$-value $> 0.1$). In Model 2, this difference is highly significant ($p$-value $< 0.001$).

depend on $\kappa$ and on whether the time horizon is deterministic or stochastic ($H = d, s$). More precisely,

$$\bar{\pi}(\kappa, d) = \frac{2\kappa}{T} \quad \text{(deterministic treatments)}, \tag{2}$$

$$\bar{\pi}(\kappa, s) = 2\left[1 - e^{-\kappa/T} - \frac{\kappa}{T}\,\text{Ei}\left(-\frac{\kappa}{T}\right)\right] \quad \text{(stochastic treatments)}, \tag{3}$$

where $T$ represents the (expected) total duration of the supergame and $\text{Ei}(-\frac{\kappa}{T})$ is the exponential integral function.[28]

The length of $\kappa$ depends on the leniency of the specific conditional strategy adopted by the cooperator (it can be chosen strategically or simply driven by reaction time). We computed an empirical measure of $\kappa$ by considering those pairs in the experiment that initiated a supergame with one subject cooperating and the other defecting, and then directly transitioned to mutual defection (i.e., the cooperator switched to defection before the end of the supergame). The average duration of this phase was 1.75 seconds in the Long-Deterministic, 1.51 seconds in the Long-Stochastic, 1.40 seconds in the Short-Deterministic, and 1.61 seconds in the Short-Stochastic treatments. Plugging the empirical measures of $\kappa$ into equations (2) and (3) yields the payoff matrix presented in Table VII, where $\bar{\pi}(\kappa, d) = 0.06$ in the Long-Deterministic treatment, $\bar{\pi}(\kappa, d) = 0.14$ in the Short-Deterministic, $\bar{\pi}(\kappa, s) = 0.21$ in the Long-Stochastic, and $\bar{\pi}(\kappa, s) = 0.48$ in the Short-Stochastic treatments.

For these values of $\bar{\pi}(\kappa, H)$, the set of beliefs that make a conditionally cooperative strategy optimal (i.e., its "basin of attraction"; see Myerson (1991, Chapter 7.11)) is much smaller and the riskiness of cooperating (Blonski and Spagnolo (2015)) is much larger in the Short-Stochastic than in the other three

TABLE VII

MATRIX OF THE EXPECTED PROFIT PER SECOND

| | Strategy of the Column Player | |
| Strategy of the Row Player | Always Defect | Conditional Cooperation |
| --- | --- | --- |
| Always Defect | 0, 0 | $\bar{\pi}(\kappa, H), -\bar{\pi}(\kappa, H)$ |
| Conditional Cooperation | $-\bar{\pi}(\kappa, H), \bar{\pi}(\kappa, H)$ | 1, 1 |

[28]For the stochastic treatments, this profit is equal to 2 in all supergames that last less than $\kappa$, which have probability $\int_0^\kappa \lambda e^{-x\lambda}\,dx$, where $\lambda = \frac{1}{T}$ is the parameter of the geometric distribution from which supergame duration is drawn. For supergames whose realized duration $\tau$ is longer than $\kappa$, instead, the per-second profit from playing Always Defect against a conditionally cooperative strategy is $\frac{2\kappa}{\tau}$. Hence, the expected per-second profit from defection is $2\int_0^\kappa \lambda e^{-x\lambda}\,dx + \int_\kappa^\infty \frac{2\kappa}{x}\lambda e^{-x\lambda}\,dx = 2(1 - e^{-\frac{\kappa}{T}}) + \frac{2\kappa}{T}(-\text{Ei}(-\frac{\kappa}{T}))$. The "sucker's" payoff from playing a conditionally cooperative strategy against Always Defect is simply given by $-\bar{\pi}(\kappa, H)$.

treatments.[29] More generally, the ordering across treatments following a basin-of-attraction criterion is consistent with the pattern of the frequency of cooperators depicted in Figure 3.

## 7. ADDITIONAL EMPIRICAL EVIDENCE

The discussion in the previous section showed that in an environment in which cooperators and defectors coexist, cooperating is more profitable than defecting only in supergames that are sufficiently long. Our learning model suggests that subjects tend to cooperate more when they have experienced past cooperation to be profitable, which happens less frequently in the Short-Stochastic treatment, where the incidence of very short supergames is relatively high. However, one may wonder whether it is the *variability* of supergame duration per se and the presence of many short supergames that induces this result, or it is the *unpredictability* of a stochastic horizon, which makes it impossible for subjects to condition their strategy on the duration of the supergame.

To shed further light on the puzzle of the lower cooperation rates in the Short-Stochastic treatment and deepen our understanding of the determinants of the dynamics across supergames, we designed and ran an additional treatment. In this treatment, called Variable-Deterministic, supergame ending is deterministic but supergame duration is variable. The sequence of supergame durations is calibrated to match exactly the realized durations in the Short-Stochastic treatment so as to allow for a direct comparison. If the main factor that slows down the evolution of cooperation is the variability of supergame durations and the many short realizations, then we should expect similarly low rates of cooperation and a slow adjustment in the Variable-Deterministic treatment as in the Short-Stochastic treatment. If, instead, unpredictability of duration is the main obstacle to the evolution toward cooperation, then we should observe higher rates of cooperation in the Variable-Deterministic treatment.

RESULT 5: *Cooperation levels in the Variable-Deterministic treatment are in between the Short-Deterministic and the Short-Stochastic treatments in terms of initial, mean, and median rates, but they are closer to the Short-Deterministic treatment.*

Initial and mean cooperation rates are higher in the Variable-Deterministic than in the Short-Stochastic treatment (77.9 vs. 65.9 and 57.1 vs. 52.3, respectively).[30] A deterministic horizon facilitates cooperation relative to a stochastic

---

[29]Dal Bó and Fréchette (2011) and Blonski, Ockenfels, and Spagnolo (2011) showed that these concepts capture coordination problems well in discretely repeated games experiments.

[30]The difference is significant for both comparisons ($p$-value $< 0.05$ in both cases) according to a linear regression with one observation per subject and bootstrapped standard errors. Regression results are reported in Table A.II in Appendix A, and in Tables B.IX and B.X in the Supplemental Material. Figure B.3 in the Supplemental Material presents the time profile of the mean share of cooperators, taken across supergames. The five lines correspond to supergames of different duration.
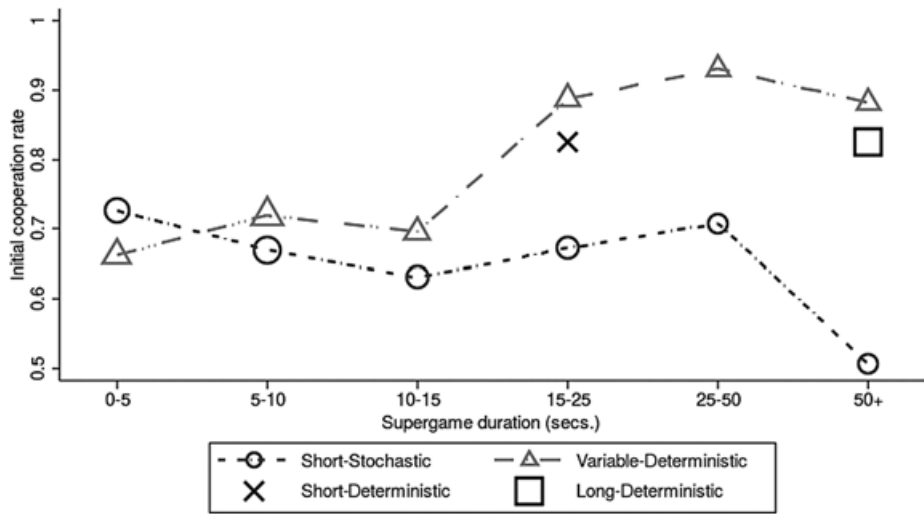
FIGURE 4.—Rate of initial cooperation by supergame duration. *Notes*: The sizes of line markers are proportional to the number of observations.

horizon, even with identical realized durations, further confirming the robustness of our main result.

When the horizon is deterministic, supergame duration matters for cooperation. In the experiment, initial cooperation rates monotonically increase with supergame duration (Figure 4).[31] As a benchmark, Figure 4 also presents the initial cooperation rate in the Short-Stochastic treatment, where subjects cannot condition behavior on the actual supergame duration.[32]

The data from the Variable-Deterministic treatment suggest that both the variability and the unpredictability of supergame durations had an impact on cooperation, although the unpredictability of the stochastic horizon was the stronger force (Result 5). This last finding is in line with the Bereby-Meyer and Roth's (2006) result that introducing randomness in strategic situations, while holding expected payoffs constant, hinders the subjects ability to adapt.

---

[31]There is a kink around 10–15 seconds, which makes the relation nonlinear. The presence of the kink suggests that when the horizon is deterministic, most subjects cooperate only when the end is far enough away, while in very short supergames, they defect from the start.

[32]By contrast, we observe that in the Short-Stochastic treatment, the realized duration of the *previous* supergame weakly affects subjects' decisions whether to cooperate in the subsequent game, which is in line with Dal Bó and Fréchette (2011). In the Variable-Deterministic treatment, instead, this is not the case (Table B.XI in the Supplemental Material). One possible interpretation of this is that when the horizon is stochastic, subjects update their beliefs about the expected value of cooperation. The expected value depends not only on what they think others do, but also on the expected length of the supergame: the longer is the supergame, the higher is the gain from cooperation. When the horizon is deterministic, there is no need for such an update.

## 8. DISCUSSION

We now briefly discuss the theoretical literature related to our experiment and results.

Standard theory of repeated games in discrete time predicts that by subgame perfection, there should be no cooperation in prisoner's dilemma games with a deterministic horizon. While it is well known that this prediction is not confirmed in experiments, partial empirical support for this theory comes from two results documented in the literature: with experience, subjects typically start defecting sooner; in short games, cooperation rates are higher with a stochastic rather than as deterministic horizon (Dal Bó (2005)). The main result of this paper is that we observe the opposite patterns in a continuous-time experiment. Among the theories that predict no unraveling and cooperation (or at least the possibility of high cooperation) in games with a deterministic horizon, we can identify two sets.

The first set of theories puts some restrictions on the strategy space or on how players form expectations on others' strategies. It includes Neyman (1985) and Jehiel (2005) on discrete-time games and Simon and Stinchcombe (1989) and Bergin and MacLeod (1993) on continuous-time games. Simon and Stinchcombe (1989) obtain mutual cooperation as the unique equilibrium of the continuous-time limit, while Bergin and MacLeod (1993) present a folk theorem for continuous-time games, independently of their duration. In discrete time, Neyman (1985) obtains a cooperative equilibrium for finitely repeated games when there is a bound on the complexity of the strategies players can adopt. Jehiel (2005) presents a model with coarse analogy groupings (strategies depend on a coarse set partition of all possible histories) that admits both unconditional defection and conditional cooperation to form an equilibrium of the game.

Theories in the second set perturb the repeated game by introducing some incomplete information. They include Kreps, Milgrom, Roberts, and Wilson (1982) and Radner (1986) for discretely repeated games, and Friedman and Oprea (2012) and Park (2013) for continuous-time games. In these models, players are not certain about the behavior or preferences of their opponents. For example, in Radner (1986) and Friedman and Oprea (2012), players assign some probability to their opponent playing a Cut-Off strategy: play Grim-trigger until some time $t$ and then defect. In Kreps et al. (1982), players assign positive probability to their opponents playing a Grim-trigger strategy, and in Park (2013), there is, additionally, private information about reaction times. These models predict equilibria (often unique) with cooperation at the beginning of the game and defection close to the end.

Overall, both sets of theories admit high cooperation rates as one of the equilibrium outcomes in continuous-time or frequently repeated games, no matter what the termination rule is. Moreover, both sets of theories offer no reason for unraveling in deterministic-horizon games. In these dimensions, our results

support these models in contrast to unperturbed discrete-time models with unrestricted strategies that got stronger support from discrete-time experiments.

While our experiment shows qualitative differences that speak against the approximation of continuous-time games with models of standard repeated games in discrete time, it cannot discriminate among the two sets of theories discussed above. We do observe clear end-game effects predicted by only some of these theories, but we recommend caution in declaring them winners. In particular, they do not predict that with experience, defections would happen later and later in the game. The pattern we observe could be consistent with subjects converging to mutual cooperation for the whole duration of the game. In other words, depending on the theoretical model taken as reference, the dynamic across supergames could be interpreted as a convergence toward equilibrium play or a gradual shift from less cooperative equilibria toward more cooperative equilibria, or both. For the latter interpretations, one would also need a theory of equilibrium selection. We cannot tell these possibilities apart. In general, new experimental designs are needed to discriminate between these models.

Some of our results, however, are hard to reconcile with any of these theories. First, none of these equilibrium theories predicts higher cooperation levels in the deterministic than in the stochastic treatments. Second, they all seem to predict that initial cooperation rates should be independent of the termination rule or the duration of the game. Our results contradict both these predictions. Our best explanation of the data is that the observed patterns are an effect of differences in the evolution of behavior with experience rather than differences in the set of equilibria in these games. In particular, in Section 6 we have shown that a model based on reinforcement learning captures well the differences across treatments in the growth of cooperation.

## 9. CONCLUSIONS

We studied repeated prisoner's dilemma games of different durations in (quasi-) continuous time, under deterministic and stochastic horizons. The experiment showed how behavior in continuous-time games is qualitatively different from what is typically observed in discrete-time games. With games of short duration, cooperation rates in continuous time were significantly higher in deterministic than in stochastic horizon treatments (with long duration, they were similar). Moreover, while with a deterministic horizon there was a drop in cooperation levels toward the end of the game, in our experiments, subjects learned to postpone more and more this end-game effect, contrary to the evidence in discrete time.

Two main economic forces influence whether agents can successfully cooperate in repeated games with perfect monitoring: (i) incentives, that is, the trade-off between immediate gains and the shadow of future punishments; (ii) coordination, that is, the ability to reach a cooperative equilibrium. These

forces are relevant both in theory and in the experimental evidence. Our results suggest that in situations where agents can react quickly to moves by others, the first force is second order. Instead, the second force—the ability to make use of experience for coordinating on future play—appears to be a first-order determinant of cooperation levels.

We explicitly considered the evolution of behavior with experience as a potential explanation of our results. We found that a simple reinforcement learning model can account for most of the observed patterns. In particular, the model predicts that experience is an important driver that increases cooperation in all treatments, but that its impact is weaker in games with a stochastic horizon because, in very short supergames, defecting from the start is more profitable than cooperating.

An additional treatment with deterministic duration and supergames of variable length matching the realized durations of the stochastic treatment confirmed that a deterministic horizon facilitates cooperation. It also showed that the presence of uncertainty in supergame duration slows down the convergence to cooperation more than the variability in supergame length.

The reported findings may have important implications for a variety of field applications. People facing social dilemmas in which they can react swiftly, as in many productive, labor, sporting, and military activities, can easily overcome the challenge of achieving mutual cooperation, irrespective of the deterministic or stochastic horizon of the interaction, even for short duration activities. In those situations, a deterministic horizon is not an impediment to cooperation and may even facilitate it.

On collusion practices, our results may explain why higher prices have been observed in oligopolies when the date of the last interaction is made public. Szymanski (1996), for example, noted that the two incumbent shipping companies in the English Channel increased prices substantially when the threat of the Eurotunnel taking the best part of their market became real. Assuming a monopolistic market, his model suggested that this happened because of the reduced fear of regulatory intervention, given its fixed costs, and the fact that the tunnel was expected to soon reduce prices dramatically anyway. However, he admitted that he could not explain how this theory could apply to the shipping duopoly that motivated his paper, that is, why competition among the duopolists did not drive prices down, given that the Eurotunnel limited the horizon of their interaction. Our results offer a plausible explanation. They also suggest that policies designed for discretely repeated interactions may be ineffective or counterproductive in high-frequency environments.

To draw implications from the experimental results, however, one should keep in mind that these activities must share some well defined features: they should involve a continuous-time effort by participants, as when together carrying a heavy object or jointly rowing in a boat, and participants must perfectly observe the action or effort taken by the opponent. Further work is needed to understand the domain of application of these results, for instance, with

respect to shorter supergame lengths or other details. In particular, the intro-
duction of imperfect monitoring of the opponent's action may limit, or remove
altogether, the possibility of sustaining a cooperative outcome when actions
are chosen frequently (as in the theoretical results in Sannikov and Skrzypacz
(2007)).

# APPENDIX A

### TABLE A.I

PANEL REGRESSION ON THE AVERAGE TIMING OF THE END-GAME EFFECT[a]

| *Dependent Variable: Average Timing of the End-Game Effect* | Coefficient | (s.e.) |
|---|---|---|
| Short-Deterministic | −7.182* | (3.718) |
| Supergame | −1.178*** | (0.342) |
| Supergame$^2$ | 0.027** | (0.014) |
| Supergame × Short-Deterministic | 0.609 | (0.479) |
| Supergame$^2$ × Short-Deterministic | −0.012 | (0.019) |
| Reaction time | 4.135*** | (1.153) |
| Reaction time × Short-Deterministic | −2.639 | (1.775) |
| Constant | 14.956*** | (2.876) |
| $N$ | 92 | |
| $R$-squared overall | 0.680 | |
| $R$-squared between | 0.933 | |
| $R$-squared within | 0.514 | |

[a]Panel regression with random effects at the session level. The unit of observation is the
session average of timing of the end-game effect (in seconds) in each supergame of the Short-
Deterministic and Long-Deterministic treatments.

### TABLE A.II

LINEAR REGRESSION ON MEAN AND INITIAL COOPERATION RATES[a]

| | Mean Cooperation Rate | | Initial Cooperation Rate | |
|---|---|---|---|---|
| Short-Deterministic | 11.574*** | (2.359) | 17.051*** | (6.164) |
| Variable-Deterministic | 6.966** | (3.094) | 13.489** | (6.751) |
| Constant | 39.488*** | (13.287) | 65.719*** | (16.057) |
| Controls for individual characteristics | Yes | | Yes | |
| $N$ | 144 | | 144 | |
| $R$-squared | 0.155 | | 0.216 | |

[a]Linear regressions with bootstrapped standard errors. To take care of the potential correlation between obser-
vations coming from the same session, we take sessions as resampling clusters. The unit of observation is the mean
(initial) cooperation rate in a section in a supergame. Default treatment: Short-Stochastic. Standard errors are re-
ported in parentheses. A full version of this regression, including the estimated coefficients for all the controls, is
reported in Table B.IX in the Supplemental Material.

TABLE A.III

SET OF STRATEGIES USED IN THE MAXIMUM LIKELIHOOD ESTIMATION

| Strategy | Description |
|---|---|
| A. Fully cooperative strategies | |
| Unconditional | |
| Always Cooperate | Always play C |
| Conditional | |
| Forgiving | |
| Tit-for-tat | Play C unless the opponent played D at $t-1$ |
| T2 | Play C until either player plays D; then play D for 2 seconds and revert to C |
| Win-stay-lose-shift | Play C unless there was a mismatch between the subject's and the opponent's action in $t-1$; otherwise play D |
| Tit-for-2-tats | Play C unless the opponent played D at $t-1$ and at $t-2$ |
| Tit-for-3-tats | Play C unless the opponent played D at $t-1$, $t-2$, and $t-3$ |
| 2-tits-for-1-tat | Play C unless the opponent played D either in $t-1$ or $t-2$ |
| 2-tits-for-2-tats | Play C unless the opponent played D either in $t-1$ and $t-2$ or in $t-2$ and $t-3$ |
| Tit-for-tat 2 | Play C unless there was a mismatch between the subject's and the opponent's actions either in $t-1$ or $t-2$, or if both subjects played D in $t-2$ and C in $t-1$ |
| Unforgiving | |
| Grim-trigger | Play C until either the subject or the opponent plays D; then always play D |
| Lenient grim 2 | Play C until either the subject or the opponent plays D twice in a row; then always play D |
| Lenient grim 3 | Play C until either the subject or the opponent plays D three times in a row; then always play D |
| B. Fully defective strategies | |
| Always Defect | Always play D |
| Suspicious tit-for-tat | Play D as initial action; then play tit-for-tat |
| C. Other strategies | |
| False cooperator | Play C as initial action; then always play D |
| Suspicious tit-for-2-tats | Play D as initial action; then tit-for-2-tats |
| Suspicious tit-for-3-tats | Play D as initial action; then tit-for-3-tats |
| Suspicious lenient grim 2 | Play D as initial action; then lenient grim 2 |
| Suspicious lenient grim 3 | Play D as initial action; then lenient grim 3 |
| D alternator | Play D as initial action; then alternate between C and D |

REFERENCES

ANDREONI, J., AND J. H. MILLER (1993): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *The Economic Journal*, 103 (418), 570–585. [589, 591]

AOYAGI, M., AND G. FRÉCHETTE (2009): "Collusion as Public Monitoring Becomes Noisy: Experimental Evidence," *Journal of Economic Theory*, 144 (3), 1135–1165. [590]

BEREBY-MEYER, Y., AND A. E. ROTH (2006): "The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation," *American Economic Review*, 96 (4), 1029–1042. [589,591,608]

BERGIN, J., AND W. B. MACLEOD (1993): "Continuous Time Repeated Games," *International Economic Review*, 34 (1), 21–37. [609]

BERNINGHAUS, S., K.-M. EHRHART, AND M. OTT (2006): "A Network Experiment in Continuous Time: The Influence of Link Costs," *Experimental Economics*, 9, 237–251. [591]

BERNINGHAUS, S., K.-M. EHRHART, M. OTT, AND B. VOGT (2007): "Evolution of Networks—An Experimental Analysis," *Journal of Evolutionary Economics*, 17, 317–347. [591]

BIGONI, M., M. CASARI, A. SKRZYPACZ, AND G. SPAGNOLO (2013): "Time Horizon and Cooperation in Continuous Time," Report. [602]

———— (2015): "Supplement to 'Time Horizon and Cooperation in Continuous Time'," *Econometrica Supplemental Material*, 83, http://dx.doi.org/10.3982/ECTA11380. [591]

BLONSKI, M., AND G. SPAGNOLO (2015): "Prisoners' Other Dilemma," *International Journal of Game Theory* (forthcoming). [606]

BLONSKI, M., P. OCKENFELS, AND G. SPAGNOLO (2011): "Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence," *American Economic Journal: Microeconomics*, 3 (3), 164–192. [607]

CAMERA, G., M. CASARI, AND M. BIGONI (2012): "Cooperative Strategies in Anonymous Economies: An Experiment," *Games and Economic Behavior*, 75 (2), 570–586. [599,602]

CHARNESS, G., D. FRIEDMAN, AND R. OPREA (2011): "Continuous Time and Communication in a Public-Goods Experiment," Report. [590]

COOPER, R., D. V. DEJONG, R. FORSYTHE, AND T. W. ROSS (1996): "Cooperation Without Reputation: Experimental Evidence From Prisoner's Dilemma Games," *Games and Economic Behavior*, 12 (2), 187–218. [589]

DAL BÓ, P. (2005): "Cooperation Under the Shadow of the Future: Experimental Evidence From Infinitely Repeated Games," *American Economic Review*, 95 (5), 1591–1604. [587,588,590,591, 596,609]

DAL BÓ, P., AND G. R. FRÉCHETTE (2011): "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review*, 101 (1), 411–429. [599,603,607, 608]

———— (2012): "Strategy Choice in the Infinitely Repeated Prisoners' Dilemma," Report. [599]

DECK, C., AND N. NIKIFORAKIS (2012): "Perfect and Imperfect Real-Time Monitoring in a Minimum-Effort Game," *Experimental Economics*, 15, 71–88. [591]

EMBREY, M., G. FRÉCHETTE, AND S. YÜKSEL (2013): "Backward Induction in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," Report. [588,591]

ENGLE-WARNICK, J., AND R. L. SLONIM (2004): "The Evolution of Strategies in a Repeated Trust Game," *Journal of Economic Behavior & Organization*, 55 (4), 553–573. [590]

FISCHBACHER, U. (2007): "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," *Experimental Economics*, 10 (2), 171–178. [594]

FRÉCHETTE, G. R., AND S. YUKSEL (2013): "Infinitely Repeated Games in the Laboratory: Four Perspectives on Discounting and Random Termination," available at http://ssrn.com/abstract= 2225331. [599]

FRIEDMAN, D., AND R. OPREA (2012): "A Continuous Dilemma," *American Economic Review*, 102 (1), 337–363. [587,590,591,594,597,601,609]

FUDENBERG, D., D. G. RAND, AND A. DREBER (2012): "Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World," *American Economic Review*, 102 (2), 720–749. [598,599]

GREINER, B. (2004): "The Online Recruitment System ORSEE 2.0—A Guide for the Organization of Experiments in Economics," Working Paper Series in Economics 10, University of Cologne. [594]

HAUK, E., AND R. NAGEL (2001): "Choice of Partners in Multiple Two-Person Prisoner's Dilemma Games: An Experimental Study," *Journal of Conflict Resolution*, 45 (6), 770–793. [589, 591]

JEHIEL, P. (2005): "Analogy-Based Expectation Equilibrium," *Journal of Economic Theory*, 123 (2), 81–104. [605,609]

KREPS, D. M., P. MILGROM, J. ROBERTS, AND R. WILSON (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27 (2), 245–252. [609]

MOOKHERJEE, D., AND B. SOPHER (1994): "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, 7 (1), 62–91. [603]

MURNIGHAN, J. K., AND A. E. ROTH (1983): "Expecting Continued Play in Prisoner's Dilemma Games: A Test of Several Models," *Journal of Conflict Resolution*, 27 (2), 279–300. [589]

MYERSON, R. B. (1991): *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press. [606]

NEYMAN, A. (1985): "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma," *Economics Letters*, 19 (3), 227–229. [605,609]

NORMANN, H.-T., AND B. WALLACE (2012): "The Impact of the Termination Rule on Cooperation in a Prisoner's Dilemma Experiment," *International Journal of Game Theory*, 41, 707–718. [590]

OPREA, R., K. HENWOOD, AND D. FRIEDMAN (2011): "Separating the Hawks From the Doves: Evidence From Continuous Time Laboratory Games," *Journal of Economic Theory*, 146 (6), 2206–2225. [591]

PALFREY, T. R., AND H. ROSENTHAL (1994): "Repeated Play, Cooperation and Coordination: An Experimental Study," *Review of Economic Studies*, 61 (3), 545–565. [590]

PARK, I.-U. (2013): "Cooperation in Continuous Dilemma by Privately Informed Players," Report. [609]

RADNER, R. (1986): "Can Bounded Rationality Resolve at the Prisoner's Dilemma?" in *Contributions to Mathematical Economics*, ed. by A. Mas-Colell and W. Hildenbrand. Amsterdam: North-Holland, 387–399. [609]

RAND, D., D. FUDENBERG, AND A. DREBER (2013): "It's the Thought That Counts: The Role of Intentions in Reciprocal Altruism," available at http://ssrn.com/abstract=2259407. [599]

RAPOPORT, A., AND A. M. CHAMMAH (1965): *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor, MI: University of Michigan Press. [589]

ROTH, A. E., AND J. K. MURNIGHAN (1978): "Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma," *Journal of Mathematical Psychology*, 17 (2), 189–198. [589]

SANNIKOV, Y., AND A. SKRZYPACZ (2007): "Impossibility of Collusion Under Imperfect Monitoring With Flexible Production," *American Economic Review*, 97 (5), 1794–1823. [612]

SELTEN, R., AND R. STOECKER (1986): "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach," *Journal of Economic Behavior & Organization*, 7 (1), 47–70. [589-591]

SELTEN, R., M. MITZKEWITZ, AND G. R. UHLICH (1997): "Duopoly Strategies Programmed by Experienced Players," *Econometrica*, 65 (3), 517–555. [590]

SIMON, L. K., AND M. B. STINCHCOMBE (1989): "Extensive Form Games in Continuous Time: Pure Strategies," *Econometrica*, 57 (5), 1171–1214. [609]

SZYMANSKI, S. (1996): "Making Hay While the Sun Shines," *Journal of Industrial Economics*, 44 (1), 1–16. [611]

VESPA, E. (2013): "Cooperation in Dynamic Games: An Experimental Investigation," Report. [599]

*Dept. of Economics, University of Bologna, Bologna, Italy; maria.bigoni@unibo.it,*

*Dept. of Economics, University of Bologna, Bologna, Italy and IZA; marco.casari@unibo.it,*

*Graduate School of Business, Stanford University, Stanford, CA 94305, U.S.A.; skrz@stanford.edu,*

*and*

*SITE–SSE, Stockholm, Sweden, University of Rome Tor Vergata, EIEF, and CEPR; giancarlo.spagnolo@hhs.se.*