EIEF WORKING PAPER SERIES

IEF

Einaudi Institute for Economics and Finance

EIEF Working Paper 16/11

November 2011

Time Horizon and Cooperation in Continuous Time

by

Maria Bigoni

(University of Bologna)

Marco Casari

(University of Bologna)

Andrzej Skrzypacz

(Stanford University, Graduate School of Business)

Giancarlo Spagnolo

(SITE, University of Rome "Tor Vergata", EIEF and CEPR)

Time Horizon and Cooperation in Continuous Time*

Maria Bigoni† Marco Casari‡ Andrzej Skrzypacz§
Giancarlo Spagnolo¶

January 18, 2013

Abstract

We study interactions with different durations and termination rules in a (quasi) continuous-time prisoner's dilemma experiment. We find that cooperation is easier to achieve and sustain with deterministic horizons than with stochastic ones; end-game effects emerge, but subjects postpone them with experience; longer duration helps cooperation. Static theories for continuous-time games cannot simultaneously account for these findings and miss the evolution of behavior across supergames. We propose a simple model – based on the replicator dynamics – that proves consistent with this evidence. The analysis of strategies and an additional treatment lend further support to the proposed explanation.

JEL Classification Numbers: C72, C73, C91, C92, D74

Keywords: Backward induction, Cut-off strategy, Experiment, Folk Theorem, Learning

*We thank Pedro Dal Bó, Gary Charness, David Cooper, Dan Friedman, Drew Fudenberg, Rosemarie Nagel, Justin Rao, Al Roth, and Nathaniel Wilcox for insightful comments. This paper also benefited from comments received by participants at the ESA meeting in Luxembourg and at the Nordic Workshop in Experimental and Behavioral Economics, and seminar participants at Chapman University, UCLA, University of Bologna, Düsseldorf Institute for Competition Economics, University of East Anglia, University of Mannheim, and Ewebe in Bertinoro. The usual disclaimer applies. We gratefully acknowledge financial support from Konkurrensverket and from the ERC grant Strangers 241196.

†Department of Economics, University of Bologna E-mail: *maria.bigoni@unibo.it*

‡Department of Economics, University of Bologna.E-mail: *marco.casari@unibo.it*

§Graduate School of Business, Stanford University. E-mail: *andy@gsb.stanford.edu*

¶SITE-SSE, Tor Vergata & CEPR. E-mail: *giancarlo.spagnolo@hhs.se.*

1 Introduction

In many field situations, actors can change actions frequently and asynchronously, a different and less well understood strategic environment than discretely repeated games. Examples include firms posting prices on the Internet or via a centralized and transparent marketplace (as airlines), workers choosing effort in a plant, nearby restaurants choosing menus, and spouses sharing everyday chores. In these situations, the trade-off between short-run gains from defection and long-term losses from punishment, centered on the discount factor, tends to lose relevance because reactions are almost immediate. Other factors may then be more important as determinants of cooperative behavior, and the question is whether the usual comparative statics results for discretely repeated games also hold in continuous time.

The present paper reports results from laboratory experiments on Prisoner's Dilemma games played in (almost) continuous time, shedding new light on subjects' behavior in high frequency environments. Besides studying the comparative statics, this paper also examines which theories best describe behavior in continuous-time games. Different theoretical frameworks may be applied to these field and laboratory situations. We can think of them as standard repeated games played very quickly (or with very low discounting); we can adopt theories that model the game directly in continuous time (Simon and Stinchcombe, 1989; Bergin and MacLeod, 1993); or we can model them as limits of perturbed discrete-time games (Kreps et al., 1982; Radner, 1986; Friedman and Oprea, 2012). These theoretical approaches generate different predictions regarding the existence of a fully cooperative equilibrium, the uniqueness of the equilibrium, the individual strategies that support equilibrium outcomes and the presence and timing of an end-game effect. Running experiments in (almost) continuous time offers a clean and controlled way to empirically investigate what is different in these high frequency environments and which theoretical approach fits them best.

Most previous experimental studies of social dilemmas compared situations where the game is repeated with relatively low frequency. This path has only been broken recently by the experiment of Friedman and Oprea

(2012), which has shown that continuous time does make a substantial difference. They found that when actions in a Prisoner's Dilemma can be changed at a very high frequency, so as to approximate a continuous-time game, very high levels of cooperation are sustained with a deterministic time horizon and relatively modest duration (60 seconds).

Here we study the comparative statics of cooperation levels in similar high frequency interactions when changing key parameters such as period duration – Long (60 seconds), Short (20 seconds), or Variable – and termination rule – deterministic or stochastic horizon. We parallel the experimental study conducted by Dal Bó (2005) in discrete time to study the comparative statics of games in continuous time.

While obtaining cooperation rates similar to Friedman and Oprea (2012) in a comparable treatment, we uncover a puzzling qualitative difference from the typical findings for discretely repeated games. In the short duration treatments, cooperation rates are significantly *higher* with a deterministic horizon than with a stochastic one (in the long duration they are similar). In experiments with discretely repeated games, the opposite effect of time horizon on cooperation rates is typically observed. The result appears robust and contrasts both with predictions from standard discrete-time repeated game theory and with previous experimental results on discrete-time Prisoner's Dilemma games.

A second puzzling result is that with deterministic duration, an *end-game effect exists but does not unravel cooperation*. As subjects gained experience, the end-game effect became less pronounced instead of the subjects learning to apply backward induction and defect sooner. This result also speaks against treating these situations like standard discretely repeated games, and contrasts with experimental findings from discretely repeated games where, with finite horizon, the end-game effect is stable or strengthened by experience. It is also not entirely consistent with theories of continuous-time games, as these do not predict an end-game effect (but it is consistent with the third class of theories based on perturbations).

The third empirical regularity we observe is that cooperation rates are significantly higher in treatments with longer expected duration. This finding is in line with predictions based on standard, discretely repeated games, where the shadow of the future is the main determinant of the decision to cooperate. It is particularly strong in treatments with a stochastic ending and contrasts with predictions of continuous-time models.

How do these empirical findings compare with predictions of the three classes of theoretical models? The most promising approach appears to be the one based on a perturbation of discrete-time games (Kreps et al., 1982; Radner, 1986; Friedman and Oprea, 2012), which predicts high cooperation rates under deterministic duration and an end-game effect that persists without leading to a complete unraveling of cooperation, with possible treatment-specific patterns. However, none of the mentioned theoretical approaches can simultaneously account for the three findings. In particular, we are not aware of any equilibrium model which could explain the considerably higher cooperation rates with a deterministic horizon than with a stochastic one in the short duration treatments.

We hypothesize that a possible explanation of this puzzle is linked to differences in agents' learning patterns in the different treatments. To this end, we introduce a simple evolutionary model based on the replicator dynamics. We show that this model would predict that cooperation rates increase with experience in all treatments, but that the increase should be slower when the horizon is shorter and when it is stochastic rather than deterministic. Hence, this model can account for cooperation rates being higher in both deterministic horizon treatments and in longer expected duration ones (something that none of the theoretical approaches discussed earlier could explain in terms of equilibrium predictions). The model also predicts a wider use of cut-off strategies – that prescribe cooperation up to a given time and defection afterwards – when the horizon is deterministic rather than stochastic. Empirical analysis of individual strategy adoption brings consistent results: the use of cut-off strategies evolves to dominate all treatments with a deterministic horizon (and only those).

To shed additional light on why learning appears slower in the Short-Stochastic treatment, we run an additional treatment with periods of pre-announced but variable duration. Period durations exactly replicate those that occurred randomly in the Short-Stochastic treatment, but are deterministic and announced in advance. This allows us to disentangle the two differences between the Short-Deterministic and Short-Stochastic treatments: that in the latter period durations vary and are unknown at the beginning. Both factors have been suggested by different theories to have a potential impact on learning. Our experimental results show that both factors contribute to the faster learning to cooperate in the deterministic treatment: the cooperation rates in the additional Variable-Deterministic treatment are in between the other two short duration treatments, albeit the median cooperation rates are closer to the Short-Deterministic one.

Finally, we find that subjects' reaction time – while decreasing with experience – remains consistently and significantly smaller in short duration treatments than in long duration ones. This is suggestive of theories of limited attention budgets (Simon, 1971; Kahneman, 1973; Sims, 2003) as subjects seem to save and spread their limited attention energies along the game's duration. We leave to future, purposely designed experiments the task of further exploring in this interesting direction.

The next section discusses the theoretical background in more detail; Section 3 describes the experimental design; Section 4 presents our first set of results; Section 5 contains our learning model; Section 6 presents additional results and robustness checks and Section 7 reviews the related literature. Section 8 briefly concludes.

2 Theoretical Considerations

The theory of games in continuous time is less developed than its counterpart in discrete time. The topic can be approached from different perspectives; here we sketch three of them that apply to social dilemma games.

A candidate theoretical approach is to view continuous-time games as the smooth limit of standard discrete-time games. When a game is repeated in discrete time, theory predicts that behavior under deterministic vs. stochastic time horizons should be quite different. Under the standard assumptions of full rationality and self-regarding preferences, cooperation cannot be sustained in equilibrium because of the standard backward induction argument. In contrast, following the Folk theorems, if future interactions loom sufficiently large, agents can support full cooperation under a stochastic horizon. This approach would therefore predict significantly higher rates of cooperation with stochastic rather than deterministic horizons and with longer expected duration.

A second possible approach is to model the games directly in continuous time, which entails that deviations can be punished immediately (Simon and Stinchcombe, 1989; Bergin and MacLeod, 1993). In continuous-time games, the backward induction argument breaks down, as the real line is not well ordered and a last period cannot be identified, even under a deterministic horizon. In other words, in continuous time 'there is always another period' in which a deviation can be punished and backward induction cannot be used. Moreover, the immediate gain from deviation is negligible, since it can be punished with no delay. This setting leads to the prediction that cooperation is an equilibrium regardless of the type of stopping rule or of the length of the interaction, and (taken literally) that no end-game effect should take place.

A third possible approach considers discrete-time games with a perturbation, which can take several forms (Kreps et al., 1982; Radner, 1986; Friedman and Oprea, 2012). The continuous-time games can be modeled as the limit of perturbed discrete-time games (as done in Friedman and Oprea, 2012). These models predict cooperation in the beginning of the game and defections in the end, with the end-game effects depending on the beliefs subjects have about the reaction times and the possible perturbations.

Table 1 summarizes the theoretical predictions of the above models, which we briefly describe in what follows. Simon and Stinchcombe (1989) build a general model of games with a finite number of actions and players. They define the game on a discrete grid in a finite interval, then let the grid interval go to zero, and assume that each strategy admits a uniformly bounded number of moves in the game. Reasoning by backward induction, they obtain that cooperation is typically sustainable in subgame perfect equilibrium and that for the Prisoner's Dilemma, full cooperation is the unique equilibrium surviving iterated deletion of weakly dominated strategies. The intuition behind this result is that no player would ever switch from defection to cooperation when she has only one move left. So if both players can react with a delay that tends to zero and can switch action at least once in the game, the game will never end with one player defecting and another one cooperating. Strictly speaking, in a continuous-time game this theory suggests that under a deterministic horizon, there will be no sizable end-game effects.

	Discrete-time games	**Limit of perturbed discrete time games**		**Continuous-time games**	
		Limit of Kreps et al. (1982)	Radner (1986), Friedman & Oprea (2012)	Bergin & McLeod (1993)	Simon & Stinchcombe (1989)
Deterministic horizon					
Full initial cooperation is an equilibrium	No	Yes	Yes	Yes	Yes
End-of-period effect	No	Yes	Yes	No/Yes†	No
Period length can impact cooperation rates	No	Yes	Yes	No	No
Stochastic horizon helps initial cooperation to be sustained in equilibrium	Yes	No	No	No	No
Type of equilibrium concept	SPNE	PBE	ϵ-equilibrium	ϵ-SPNE	Iterated dominance
Overall empirical support	weak	moderate/strong	moderate/strong	moderate	moderate

Notes: SPNE: Subgame Perfect Nash Equilibrium; PBE: Perfect Bayesian Equilibrium.
† With Folk Theorem equilibria with end-game effects also exist.

Table 1: Main theoretical predictions of different models.

Bergin and MacLeod (1993) build a related model that includes a degree of inertia in changing actions as interactions are structured in a sequence of intervals from t to $t+\epsilon$. They characterize the set of ϵ-subgame-perfect-equilibria and then let ϵ go to zero. This leads to a full Folk Theorem for the continuous-time Prisoner's Dilemma that holds for both deterministic and stochastic horizons. The intuition behind these predictions is that if players adopt a trigger strategy that punishes a defection after a time interval of size ϵ, the magnitude of the gains of defection is also of the order of ϵ. Thus, as ϵ approaches zero, the incentive to deviate also vanishes. Because of the multiplicity of equilibria, this theory has a weaker predictive power than Simon and Stinchcombe (1989) and is consistent with a large number of equilibrium paths observed in the lab.

Radner (1986) puts forward a theory of bounded rationality in discretely repeated games with a finite horizon based on $\epsilon-$equilibria (recently adopted and extended by Friedman and Oprea, 2012, to explain their results). He predicts full initial cooperation, as long as there is a small probability that the opponent plays a "cooperative" dynamic behavioral strategy. His behavioral restriction is to a class of strategies of the form "cooperate until period k or until the other player defects and defect otherwise," so-called cut-off strategies. He notes that if the players can react swiftly to a defection of the other player, the losses that a player may incur using a cut-off strategy with a very large k are bounded to be very small, while the same strategy allows large gains from prolonged cooperation if the opponent uses a cut-off strategy with a large k. The best response strategy, defect at $k-1$ if the other player waits until k, leads to backward induction and unraveling. Relative to the safe but low non-cooperation payoffs obtained using best reply and the induction argument they trigger, the cooperative strategies become more and more attractive when the number of repetitions grows. This implies that cooperation can be sustained in deterministic horizon games with many periods or frequent actions, if subjects realize that continuing cooperating rather than defecting produces large expected benefits compared to the risk of small losses one is exposed to. This argument applies, of course, to a stronger extent to continuous-time games, as stressed by Friedman and Oprea (2012). It is also consistent with an end-game effect at the end of the finite horizon. As shown in Appendix B.1, this model predicts the time distance between the switch to permanent defection and the end of the game to be proportional to subjects' reaction time and otherwise independent of games' duration.

Within the same approach, one can also include the 'gang of four' paper for discrete-time games under a deterministic horizon (Kreps et al., 1982). As we discuss in Appendix B.2, the main predictions of this model are observationally equivalent to those of Friedman and Oprea (2012). Taking the periods to be short, Kreps et al. (1982) also predict high rates of cooperation in the beginning of the game and an end-game effect. The timing of this effect should be independent of the game's duration but a function of subjects' reaction time. There are subtle differences between the predictions in Kreps et al. (1982) and Friedman and Oprea (2012) since the models are built around slightly different perturbations of rational behavior. Our data are not well-suited to discriminate between these models.

The models in Kreps et al. (1982), Radner (1986), and Friedman and Oprea (2012) are built for a deterministic horizon, hence their application to a stochastic horizon is less straightforward. That said, it appears that the presence of types that play cooperative strategies, once appropriately adjusted for the infinite horizon, should not make cooperation more difficult to sustain as an equilibrium outcome.

3 Experimental Design

The experiment has five treatments. The two treatment variables are the expected duration of each period and the termination rule. Table 2 summarizes the characteristics of each treatment.

In all treatments, subjects played a series of (quasi) continuous-time Prisoner's Dilemma games with stage-game payoffs as in Table 3.[1]

Each session comprised a non-overlapping group of 24 subjects, who interacted in pairs for 23 periods. Pairs were formed so that each subject

[1]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 16 hundredths 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.

	Termination rule	
	Deterministic	**Stochastic**
Short (20 secs.)	N=48 Period endowment: 15 pts. Conversion rate: 50 pts.=1 € - January 24, 2011 - February 4, 2011	N=48 Period endowment: 15 pts. Conversion rate: 50 pts.=1 € Average realized duration: 22.6" - February 2, 2011 - February 4, 2011
Long (60 secs.)	N=48 Period endowment: 50 pts. Conversion rate: 150 pts.=1 € - October 21, 2010 - October 28, 2010	N=48 Period endowment: 50 pts. Conversion rate: 150 pts.=1 € Average realized duration: 68.3" - October 22, 2010 - October 28, 2010
Variable	N=48 Period endowment: 15 pts. Conversion rate: 50 pts.=1 € Same realized durations as in Short-Stochastic - April 4, 2012 - April 4, 2012	

Table 2: Treatments and sessions

	coop.	**defect**
coop.	1, 1	-2, 2
defect	2, -2	0, 0

Notes: The numbers in each cell represent the payoff per second.

Table 3: Stage game payoffs

met all the others once and only once in a session (perfect strangers).[2]

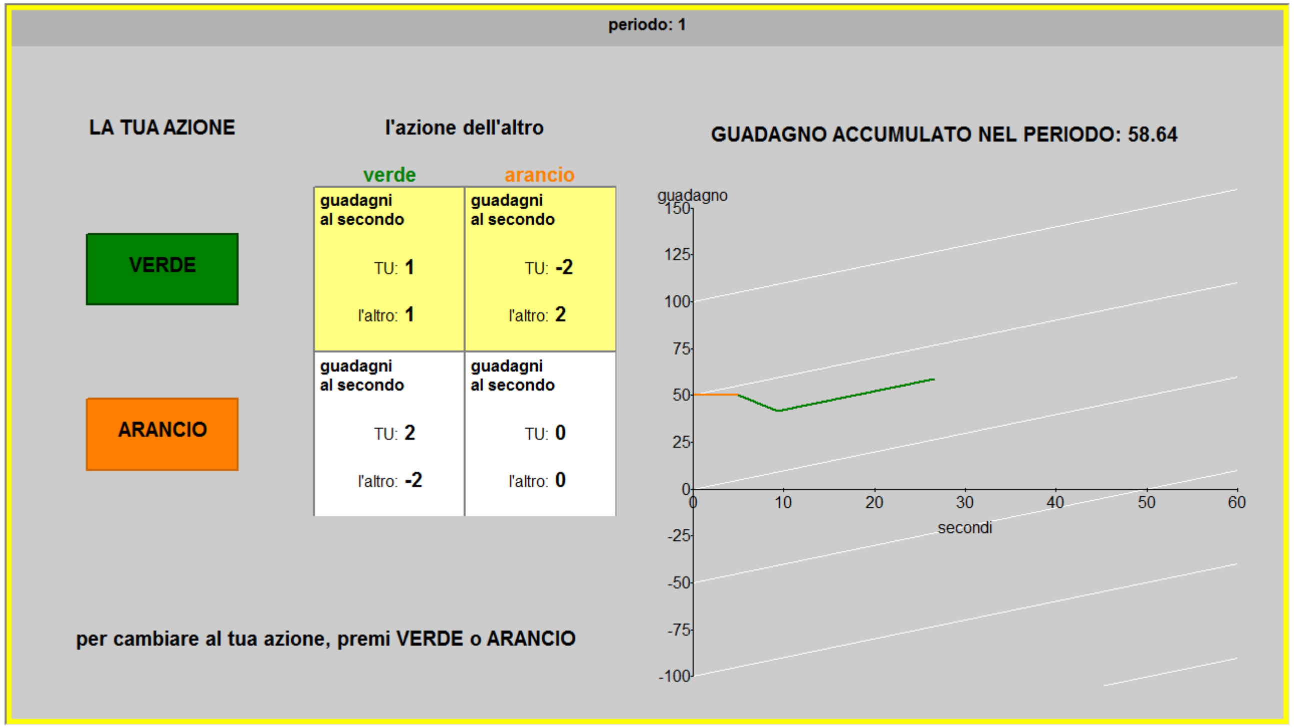
In all treatments, the stage game was as follows. Each subject had to select an initial action for the period, either Cooperate (green) or Defect (orange). When all subjects were done, the period began. Within a period, subjects could switch action up to six or seven times per second. More precisely, there was a tick every 16/100th 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.

Earnings for all possible combinations of actions were visible on the screen at all time (Figure 1). 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 (Figure 1). Subjects' earnings in every period included an initial endowment (see Table 2), 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 period for the subjects. They could observe at every instant 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 period by looking at the graph on the screen.

In the Long-Deterministic treatment, a period always lasted 60 seconds. In the Long-Stochastic treatment, a period 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 period by period. As explained in the instructions for the Long(Short)-Stochastic treatment, the period duration depended on a ran-

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

Figure 1: Screenshot of the stage game for the Long treatments



Notes: VERDE = green, ARANCIO = orange, l'azione dell'altro = your opponent's action, guadagno = earnings.

dom draw. "Imagine a box with 10,000 (1000) balls, of which 9,973 (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 16 hundredths 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 period lasted beyond 60 seconds, the the scale of the horizontal axis of the graph automatically shifted forward.

The Variable-Deterministic treatment was designed as a modification of the Short-Stochastic treatment. Period durations were variable and pre-announced. At the beginning of each period, the current period duration was disclosed to the subjects both numerically – in terms of seconds – and graphically – through a vertical line drawn in the chart of Figure 1. To facilitate comparisons, the actual period durations in Variable-Deterministic sessions replicated by design the random draws employed in the Short-Stochastic sessions.

Stage-game payoffs are such that cooperation should be easily achieved (at least in the stochastic ending treatments). In continuous time, cooperation is always supportable because the instantaneous discount factor is 1: then a grim trigger strategy should, in theory, always support cooperative play as an equilibrium, no matter the arrival rate of the end of the game. But even if agents perceived the game to be played discretely, e.g. because of minimal human reaction time, cooperation should be easily sustained with our parameterization. For example, if subjects react with 1 second delay and treat it as a time interval length of 1 second, then, given our stage game payoffs (see Figure 1), cooperation can be sustained with 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. Subjects had the opportunity to ask questions, which were answered in private, and then went through three practice periods with a robot opponent that was programmed to switch action in the middle of the period. After each practice period, subjects had to guess the actions taken by the robot, and then completed 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.[3] The session ended with a questionnaire.

The experiment involved 240 subjects, mostly students at the University of Bologna, participated to one of the ten sessions and 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. Subjects earned, on average, 16.72 Euros and 15.47 Euros, respectively, which include a show-up fee of 3 Euros.

4 Results

With our Long-Deterministic treatment, we replicate the results reported in Friedman and Oprea (2012). The median cooperation rate from period 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.

Result 1 *Cooperation rates are higher in periods of longer (expected) duration.*

Support for Result 1 comes from Tables 4 and 5. The impact of duration on cooperation rates is significant both in the deterministic and stochastic treatments. The unit of observation is the cooperation rate, which is defined as the fraction of time R_{ip} a subject i spends cooperating within period p. Given that these observations are not independent, to assess the significance

[3]In the three practice periods, 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.

of the observed differences we take the average cooperation rate by subject across all periods, and run a linear regression with bootstrapped standard errors. Results are reported in Table 5.[4] The outcome of the regression in Table 5 indicates that Result 1 holds when controlling for individual characteristics.[5]

	Termination rule		
Duration	**Deterministic**		**Stochastic**
Long	65.5	$\sim$	66.9
	(84.0)		(84.8)
	$\vee^{*}$		$\vee^{***}$
Short	63.3	$>^{***}$	52.3
	(79.2)		(47.0)

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

Table 4: Mean (and median) cooperation rates

Result 2 *With deterministic duration, cooperation rates are equal or higher than with stochastic duration.*

Support for Result 2 comes from Tables 4 and 5. In the long duration treatments, cooperation rates are statistically indistinguishable between

[4] 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 period, and standard errors are robust for heteroschedasticity (see Table A.1 in Appendix A). 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 A.2 and A.3), or the average profits per second (Tables A.4 and A.5).

[5] 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 one for subjects who have previously followed courses of economics, statistics, and game theory, respectively; (iv) non-incentivized questionnaire measures: risk attitudes, level of generalized trust, and two IQ-type questions.

Dependent variable: cooperation rate		
	Coefficient	(s.e.)
Short-Deterministic	-6.082*	(3.265)
Long-Stochastic	1.998	(4.153)
Short-Stochastic	-17.755***	(3.744)
Constant	62.600***	(12.508)
Controls for individual characteristics	Yes	
N	192	
R-squared	0.223	

Notes: Linear regression with bootstrapped standard errors. The unit of observation is the fraction of time a subject spends cooperating within a period, averaged by subject across all periods. Default treatment: Long-Deterministic.

Table 5: Linear regression on cooperation rates

stochastic and deterministic horizons (p-value > 0.1, see Table 5). 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 5). 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.[6]

The next result shifts the focus on the dynamics within each period, 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 apply backward induction.

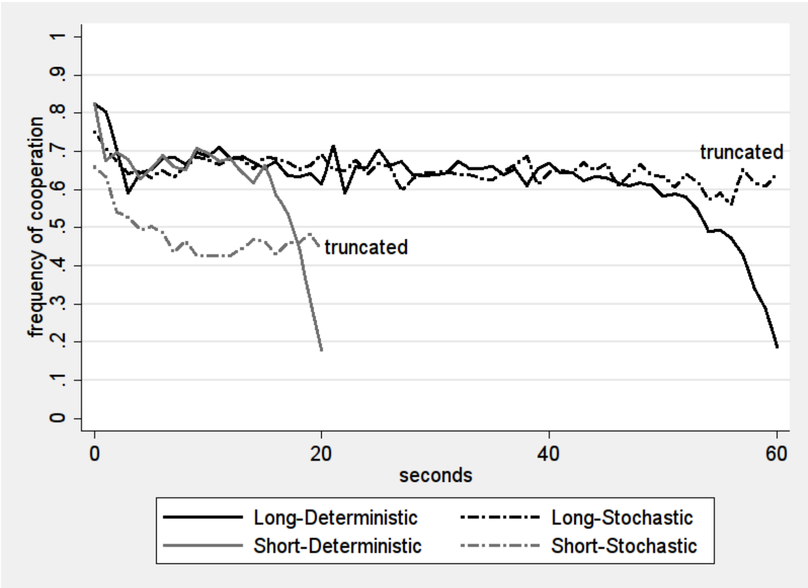
Result 3 *With deterministic duration, end-game effects exist, do not unravel cooperation, and appear later with experience.*

Support for Result 3 comes from Figure 2, which presents the time profile of the mean share of cooperators, taken across periods and sessions.

[6]For example, in a repeated game with short expected duration, Dal Bó 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, 375 in our long treatments, while it ranges between 2 and 4 in his treatments). More specifically, when the expected duration is 2 (4) periods, the average cooperation rate is 28.3% (35.2%) with stochastic ending and 12.5% (24.8%) with deterministic ending.

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 period p.

Figure 2: Time profile of the share of cooperators



Notes: The unit of observation is the share of cooperators in every second of a period. A subject could switch action every 0.16 of a second. All subjects and all periods are included for the first 20 or 60 seconds. In the Long-Stochastic treatments, 45.7% of periods lasted more than 60 seconds. In the Short-Stochastic treatments, 30.4% of periods 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 period (Figure 2). With deterministic duration, this end-game effect kicked in, on average, 8.4 seconds before the end of the period (Table 6).[7] 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 a period reached simultaneous cooperation, CC, and then switched to defection

[7]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."

before the end of the period, i.e. CD, DC, or DD.[8]

	Periods				
Treatment	1-6	7-12	13-18	19-23	Overall
Long-Deterministic	17.7	11.5	11.4	7.1	11.9
	N=95	*N=110*	*N=110*	*N=99*	*N=414*
Short-Deterministic	7.4	5.3	4.1	3.8	4.9
	N=77	*N=100*	*N=120*	*N=112*	*N=409*

Notes: the table reports the average number of seconds before the end of the period when a pair in CC permanently switches to defection, i.e. either CD, DC, or DD.

Table 6: Timing of the end-game effect

We observe that the end-game effect kicks in later and later, as subjects gain experience (3.6 to 10.6 seconds later, Table 6). 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). In addition, in the Short-Deterministic treatment, the end-game effect kicks in significantly later than in the Long-Deterministic treatment (Table 6, p-value < 0.01).[9]

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, we observe that reaction time – measured as the time interval between a deviation from mutual cooperation and the beginning of the punishment phase – decreases across periods. The correlation between reaction times and timing of the end-game effect, however, is not-significant (Table A.7 in Appendix A). When controlling for the average reaction time in the regression, the decrease in the duration of the end-game effect across periods is still significant (see Table A.6 in Appendix A). These findings show that the end-game effect does not unravel cooperation.

[8]This calculation includes the lion's share of the observations. Out of a total of 552 subject-period 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 period. Of these, some (54/468 and 51/460, respectively) kept on cooperating until the end of the period, while in the other cases (414/468 and 409/468, respectively) at least one of the subjects in the pair switched to permanent defection.

[9]The p-values reported in this paragraph are obtained from linear regressions with bootstrapped standard errors. The unit of observation is a session in a period. Regression's results are reported in Table A.6 in Appendix A.

Taken together, Results 1-3 offer a puzzling picture, because neither the theories of games in discrete time nor those in continuous time can provide a coherent explanation. In our treatments with stochastic horizon, cooperation rates are higher when the horizon is longer (Result 1), which is in line with theoretical and empirical findings for discretely repeated games, but contrasts with all the theories of play in continuous time we considered in Section 2.

Conversely, Result 2 suggests that when the frequency of interaction becomes very high, the shadow of the future is not the main factor driving subjects' behavior. We find that cooperation can be achieved and sustained even when the horizon is deterministic, a result which is in contrast with the theory and (some of) the experimental evidence on discretely repeated games (Dal Bó, 2005), and is instead in line with theories of games in continuous time. In addition, Result 3 indicates that an end-game effect emerges when the horizon is deterministic. This provides support for those theories of games in continuous time that predict the presence of this effect, such as the limit version of Kreps et al. (1982) and the theory developed by Friedman and Oprea (2012) from the model by Radner (1986).

Another more serious puzzle is the low cooperation rate in the Short-Stochastic treatment, which sets it apart from the other three treatments (Result 2, see Tables 4 and 5, and Figure 2). None of the theories in discrete time nor continuous time in Section 2 would predict such a result.

In the next section we introduce a new element to interpret the above results. We conjecture that what changes across treatments is not the equilibrium prediction, but the path of convergence to this equilibrium.

5 The dynamics across supergames

In this section, we develop an evolutionary model of how individuals change their strategies across supergames. In particular, we model the path of convergence to a stationary state through an evolutionary process across periods (supergames) based on the replicator dynamics.[10]

[10]Börgers and Sarin (1997) have shown that strong analogies exist between the replicator dynamics and reinforcement learning. We are grateful to George Mailath and Jörgen Weibull, who both suggested to look at the replicator dynamics.

Consider a large population of individuals programmed to play pure strategies. Let $\mathbf{x}$ represent the population state, x_i the population share of individuals playing pure strategy i, and e_i the vector having the i-th element equal to 1 and all other elements equal to 0. In each period, individuals are randomly matched to play the continuous-time Prisoner's Dilemma presented in Table 3, for a period of expected duration equal to T. The average payoff $u(e^i, \mathbf{x})$ of a strategy i in a population state $\mathbf{x}$ measures its "fitness". The basic assumption of the replicator dynamics is that fitter strategies replicate faster:

$$\dot{x_i} = [u(e^i, \mathbf{x}) - u(\mathbf{x}, \mathbf{x})]x_i.$$

Consequently, the growth rate $\dot{x_i}/x_i$ of the population share using strategy i equals the difference between the strategy's current payoff and the current average payoff in the population, $u(\mathbf{x}, \mathbf{x})$.

Here we consider a population of individuals programmed to play one of the following strategies: Always-cooperate (AC), Always-defect (AD), Grim-trigger (GT), or a Cut-off strategy (CO). In line with the assumptions of Radner (1986)'s and Friedman and Oprea (2012)'s models, strategy CO prescribes to play Grim-trigger at every $t < \hat{t}$, with $\hat{t} < T$, where T is the expected duration of the period, and play Always-defect at every $t \geq \hat{t}$. We use existing equilibrium models to link $\hat{t}$ to the reaction time of players.

The expected period payoffs for a pair of players are as follows. Let R be the expected payoff if both players cooperate throughout the whole period, r the expected payoff if they both adopt the cut-off strategy, and d the expected loss emerging from a delay in reacting to a defection, which is proportional to reaction time. We assume that $0 < d < r < R$ and $R - r > d$. If we consider each period of our experiment as a "generation", in which our players can choose only among the aforementioned four strategies, we can write the matrix of the per-period expected payoffs as in Table 7.

We first identify which population states are asymptotically stable, then discuss the impact of the reaction time in each treatment, and finally state two propositions. If population state $\mathbf{x}$ is asymptotically stable in the replicator dynamics, it must correspond to a (mixed or pure) trembling-hand perfect Nash equilibrium strategy of the stage game (Weibull, 1995,

Player 1/Player 2	Always-defect	Grim-trigger	Cut-off	Always-cooperate
Always-defect	0, 0	d, -d	d, -d	2R,-2R
Grim-trigger	-d, d	R, R	r-d, r+d	R, R
Cut-off	-d, d	r+d, r-d	r, r	2R-r, 3r-2R
Always-cooperate	-2R, 2R	R, R	3r-2R, 2R-r	R, R

Table 7: Expected period payoffs in a pair with four possible strategies

Proposition 3.9). As a consequence, it cannot contain a positive fraction of individuals playing any weakly dominated strategy. For this reason, we can exclude Always-cooperate, and focus on the remaining three strategies.

Since the replicator dynamics is invariant under positive affine transformations of payoffs (Weibull, 1995, p.73), we can rewrite the payoff matrix as in Table 8, where $a = r/R$ and $b = d/R$.

	Always-defect	Grim-trigger	Cut-off
Always-defect	0, 0	b, -b	b, -b
Grim-trigger	-b, b	1, 1	a-b, a+b
Cut-off	-b, b	a+b, a-b	a, a

Table 8: Expected period payoffs in a pair with three possible strategies

The game in Table 8 has seven Nash equilibria, three in pure strategies, and four in mixed strategies.[11] To find whether these equilibria are asymptotically stable, we study the properties of the following system of differential equations:

$$\begin{cases} \dot{x}_{gt} = x_{gt}\left(-b + x_{gt} - ax_{gt} + bx_{gt} - x_{gt}^2 + ax_{gt}^2 + ax_{ad} - ax_{ad}^2\right) \\ \dot{x}_{ad} = -x_{ad}\left(a - b + x_{gt}^2 - ax_{gt}^2 - 2ax_{ad} + bx_{ad} + ax_{ad}^2\right). \end{cases}$$

It can be shown that the four equilibria in mixed strategies are not asymptotically stable, while the three equilibria in pure strategies are (locally) asymptotically stable (Figure 3).

[11] The equilibrium strategies are $\mathbf{x_1} = (1,0,0)$, $\mathbf{x_2} = (0,1,0)$, $\mathbf{x_3} = (0,0,1)$, $\mathbf{x_4} = (0, 1 - \frac{b}{a}, \frac{b}{a})$, $\mathbf{x_5} = (b, 1-b, 0)$, $\mathbf{x_6} = (\frac{b^2}{a(1-a)+b^2}, 1 - \frac{(1-a)b}{a(1-a)+b^2}, \frac{(1-a)b-b^2}{a(1-a)+b^2})$, and $\mathbf{x_7} = (\frac{b}{1-a}, 0, 1 - \frac{b}{1-a})$, where the first element of the triple corresponds to Grim-trigger, the second element to Always-defect, and the third one to the Cut-off strategy.

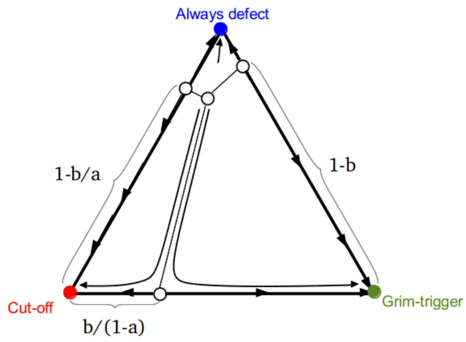


Figure 3: Replicator dynamics

We now relate the parameters of this model to the reaction time and the four treatments. Recall that the relative loss for a cooperator when the opponent switches from cooperation to defection, b, depends on her reaction time; recall also that the relative payoff when both payers follow a Cut-off strategy, a, depends on the timing of the cut-off. Given that $0 < d < r < R$, it follows that $0 < b < a < 1$. Let τ be the relative reaction time (i.e. the reaction time as a fraction of the duration of the supergame). With deterministic horizon, we have:

$$b_{Det} = 2 \cdot \tau = 2 \cdot \frac{\text{reaction time}}{T}$$

If we assume that reaction time does not depend on the duration of the supergame, or decreases less than proportionally to the duration, we have that parameter b is larger for the Short-Deterministic treatment than for the Long-Deterministic treatment.

We propose to use existing theories as guidance about parameter a. According to Friedman and Oprea (2012)'s extension of Radner (1986)'s model and to our extension of Kreps et al. (1982)'s model, if the horizon is deterministic, then the end-game effect will emerge at a time $\hat{t}$, inversely proportional to the reaction time (See Appendices B.1 and B.2):

$$\frac{\hat{t}}{T} = 1 - k \cdot \tau, \ k > 1$$

Therefore, with a deterministic horizon

$$a_{Det} = \frac{\hat{t}}{T} = 1 - k \cdot \tau$$

according to both models (albeit, k is model-specific), so a_{Det} is decreasing in τ.

With a stochastic horizon, the profits from defection depend on the realized duration of the period. If the period lasts shorter than the reaction time, then they are equal to 2 times the period duration, otherwise they are equal to 2 times the reaction time. Hence, the expected value of b in the stochastic treatments is:

$$b_{Stoch} = 2 \cdot \left(\int_0^{\tau} xe^{-x} dx + \tau \int_{\tau}^{\infty} e^{-x} dx \right) = 2(1 - e^{-\tau}) \leq 2\tau \tag{1}$$

Similarly, if all individuals played a Cut-off strategy, they would earn a profit equal to the period duration in all periods lasting less than $\hat{t}$, and a profit equal to $\hat{t}$ in longer periods. As a consequence, the expected value of a in the stochastic treatments is:

$$\begin{aligned} a_{Stoch} &= \int_0^{1-k\cdot\tau} xe^{-x} dx + (1 - k \cdot \tau) \int_{1-k\cdot\tau}^{\infty} e^{-x} dx \\ &= 1 - e^{k\cdot\tau - 1} \leq 1 - k \cdot \tau \end{aligned} \tag{2}$$

Therefore, also with a stochastic horizon, a_{Stoch} is decreasing in τ, and b_{Stoch} is increasing in τ.

From the above equations, it follows that $\frac{\partial(b/a)}{\partial\tau} > 0$. Hence, if the absolute reaction time is independent of the duration, or if it decreases less than proportionally to duration, $\frac{b}{a}$ is larger for the Short than for the Long treatments. This implies:

Proposition 1 *If the relative reaction time is smaller in the longer treatment, the basin of attraction of the equilibrium where all individuals play Always-defect is larger in the Short than in the Long treatments.*

Equations (1) and (2) also show that moving from a deterministic to a stochastic horizon, holding period duration constant, induces both a and b to be smaller. One can then show that for small relative reaction times (i.e.

if the reaction time is small compared to the duration of the supergame), moving from a deterministic to a stochastic horizon implies an increase in the ratio $\frac{b}{a}$, and a decrease in $\frac{b}{1-a}$. Indeed:

$$\frac{b_{Stoch}}{a_{Stoch}} > \frac{b_{Det}}{a_{Det}} \Leftrightarrow \frac{2(1-e^{-\tau})}{1-e^{k\tau-1}} > \frac{2\tau}{1-k\tau} \Leftrightarrow \frac{1-e^{-\tau}}{\tau} > \frac{1-e^{k\tau-1}}{1-k\tau}$$

Let $F(x) = \frac{1-e^{-x}}{x}$. Since $F'(x) < 0$, it follows that

$$\frac{b_{Stoch}}{a_{Stoch}} > \frac{b_{Det}}{a_{Det}} \Leftrightarrow \tau < \frac{1}{1+k}$$

Given that $b_{Stoch} \leq b_{Det}$ from equation (1), and $a_{Stoch} \leq a_{Det}$ from equation (2), it also follows that

$$\frac{b_{Stoch}}{1-a_{Stoch}} < \frac{b_{Det}}{1-a_{Det}}$$

Since $b_{Det} > b_{Stoch}$ but $\frac{b_{Stoch}}{a_{Stoch}} > \frac{b_{Det}}{a_{Det}}$ it is not a priori clear whether the area of the basin of attraction of the equilibrium where all individuals play Always-defect is larger in the stochastic than in the deterministic treatments. This relation, however, can be shown to be true by direct computation (see Appendix C). As a consequence, we can state the following proposition.

Proposition 2 *For small relative reaction times, the basin of attraction of the equilibrium where all individuals play Always-defect is larger in the stochastic than in the deterministic treatments, while the basin of attraction of the equilibrium where all individuals play the Cut-off strategy is smaller.*

Taken together, Propositions 1 and 2 can explain why cooperation rates are particularly low in the Short-Stochastic treatment, which was our major empirical puzzle. When the horizon is longer and when it is deterministic, it should be easier for subjects to evolve towards adopting a cooperative strategy (either Grim-trigger or Cut-off), rather than playing Always-defect. These predictions can account for Results 1 and 2 together.

Moreover, Proposition 2 provides a reason for why we observe an end-game effect in the deterministic but not in the stochastic treatments and, most importantly, why this end-game effect does not lead cooperation to unravel with experience (Result 3). The explanation hinges on the adoption of a cut-off strategy, which is progressively more widespread. Section 6 addresses this point from an empirical standpoint.

To conclude, the replicator dynamics can explain Results 1-3 through modeling convergence and strategy adoption and makes additional predictions about the behavior in continuous-time games. We now turn to comparing these additional predictions to what we observed in the experiments.

6 Discussion and additional results

Here we shift the emphasis away from the equilibrium predictions of the models and study the patterns of learning in continuous-time games. This section discusses the individual strategy adoption (Result 4), the observed evolution of cooperation rates across supergames as subjects gain experience (Result 5), and an additional treatment aimed at the identification of empirical drivers of learning (Result 6). These results are mostly consistent with the replicator dynamics model, but their aim is to report the impact of experience more than to carry out a formal test of the model.

6.1 Empirical identification of individual strategies

When taking the four strategies considered in the replicator dynamics, one can classify the majority of individuals in the last five periods of the experiment, i.e. between 52.5% and 78.3%, depending on the treatment (Table 9). The unclassified subjects follow a pattern of actions that is incompatible with all of the four strategies considered. They are overwhelmingly "rabbit"-type subjects, i.e. subjects who switched more than twice within the period and are hard to assign to any specific strategy category.[12] In short, our focus on these four strategies followed from suggestions of theoretical models and when bringing them to the data, we find broad support.

[12] For instance, considering "reverse cut-off" and "reverse-grim trigger" strategies, one captures very few additional subjects.

Result 4 *With stochastic duration, Grim-trigger strategies prevail. With deterministic duration, instead, Cut-off strategies are the most widely adopted.*

		Stochastic		Deterministic		
		Short	Long	Short	Long	Variable
Classified		**0.679**	**0.525**	**0.713**	**0.662**	**0.783**
Cut-off	U.B.	0.158	0.033	0.621	0.492	0.583
& defect before the opponent	L.B.	*0.050*	*0.000*	*0.279*	*0.196*	*0.208*
Grim-trigger	U.B.	0.529	0.492	0.429	0.446	0.446
if the opponent defects	L.B.	*0.108*	*0.033*	*0.342*	*0.296*	*0.375*
Always-cooperate	U.B.	0.438	0.471	0.087	0.150	0.071
if the opponent defects	L.B.	*0.017*	*0.013*	*0.000*	*0.000*	*0.000*
Always-defect	U.B.	0.083	0.021	0.004	0.021	0.129
Unclassified		**0.321**	**0.475**	**0.287**	**0.338**	**0.217**
switch once		0.050	0.025	0.000	0.008	0.004
switch twice		0.063	0.075	0.017	0.021	0.042
switch more than twice		0.208	0.375	0.271	0.308	0.171
N		**240**	**240**	**240**	**240**	**240**

Notes: We classify the behavior of each subject in periods 19-23. A subject is classified as compatible with Always-cooperate if she cooperates throughout the whole period. Such a subject would also be considered compatible with Grim-trigger, if her opponent cooperates for the whole period too. Hence, there is a possible overlap in the classification. We report in italics the fraction of subjects whose behavior is compatible with Always-cooperate, but not with Grim-trigger, as the opponent defects but the subject keeps on cooperating until the end of the period.

Table 9: Subjects' strategies in the last 5 periods

Support for Result 4 comes from Table 9, which shows the share of individuals compatible with the behavior prescribed by a given strategy. There can be overlap between categories, hence Table 9 also supplies a classification conditional on a specific behavior of the opponent. In the table, we denote these two figures "upper bound" (U.B) and "lower bound" (L.B.), respectively. For instance, the strategy Always-cooperate classifies 47% of individuals in the Long-Stochastic treatment, but only 1% when restricting to cases where the opponent defects at some point in the period. More generally, very few subjects always cooperate despite meeting an opponent who defects, which suggests that Always-cooperate is not widespread.

The Cut-off strategy can account for up to 49-62% (L.B. 20-28%) of individuals in the Deterministic treatments versus 3-16% (L.B. 0-5%) in the Short- and Long-Stochastic treatments. The Grim-trigger strategy can account for up to 43-45% (L.B. 30-37%) of individuals in the Deterministic treatments versus 49-53% (L.B. 3-11%) in the Stochastic treatments.

6.2 Evolution of cooperation with experience

The evolution toward cooperation occurs at different speeds depending on the treatment, which can explain the observed differentials in average cooperation rates.

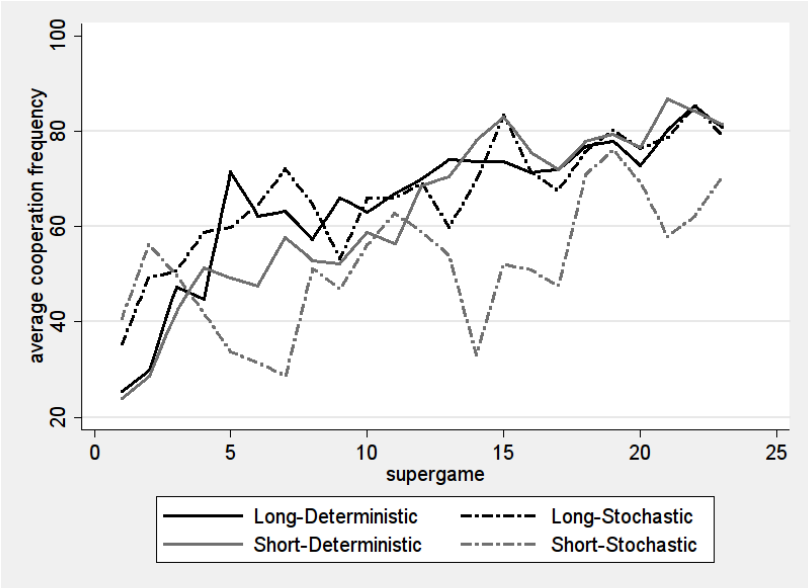
Result 5 *Cooperation rates increase with experience in all treatments, but the increase is slowest in the Short-Stochastic treatment.*

Support for Result 5 comes from Figure 4, which reports the average of the cooperation rate R_{ip} of subject i in period p across subjects. In all treatments, there is an upward trend in cooperation, but this trend is weakest in the Short-Stochastic treatment. More detailed evidence comes from the regressions reported in the Appendix (Tables A.8 and A.9). A similar trend also emerges from the evolution of the fraction of subjects choosing defection as their initial action of the period (Figure A.1 in Appendix A). This fraction declines in all treatments, but the decline is the slowest in the Short-Stochastic treatment, as predicted by the replicator dynamics.

Our results on the impact of experience on cooperation levels are consistent with the findings of Dal Bó and Fréchette (2011) for discretely repeated games. When playing repeated games, the amount of experience is a critical determinant of outcomes and it takes more than ten repetitions to settle on a stable level of cooperation.

6.3 The Variable-Deterministic treatment

The data document a strong dynamic of learning in these continuous-time games. We want to further clarify the differences in such dynamics across treatments and provide further evidence for Results 1-4. In the Short-Stochastic treatment, very short periods are frequent and there is a high variability in period duration. The variability in period duration has an impact on profits. The reason is that in very short periods, the immediate



Notes: The unit of observation is the fraction of time a subject spends cooperating in a period.

Figure 4: Evolution of cooperation

gains from defection may be larger than the foregone profits from cooperation, whereas defecting from the start in longer periods yields small gains in comparison to the large foregone profit from cooperation. According to the reinforcement learning model studied in Bereby-Meyer and Roth (2006), learning is faster in environments where the same action is rewarded in the same way under all circumstances, rather than being rewarded only in a variable and unpredictable way.[13] These factors typical of a stochastic horizon may therefore also have caused slower learning in the Short-Stochastic treatment. One could also think that learning is slower because in very short periods, players do not have time to think about what is happening. Another plausible explanation for differences in cooperation rates across treatments could be linked to issues of coordination. For example, one could think that a commonly known finite horizon helps subjects to coordinate on a specific cooperative strategy, such as a Cut-off

[13]The authors compare games with deterministic or probabilistic payoffs having the same expected value, and show that the pace at which subjects learn to cooperate is strongly affected by the variance in payoffs, to the point that, in very noisy environments, cooperation may fail to emerge altogether.

strategy prescribing to cooperate until close to the end.

In order to deepen our understanding of how time horizon and period length drive learning, we designed and ran an additional treatment. In this treatment, called Variable-Deterministic, period ending is deterministic but period duration is variable. The sequence of period durations was calibrated to match exactly the realized durations in the Short-Stochastic treatment, in order to allow for a tight comparison. Stochastic treatments are different from Deterministic treatments both because period ending is random and because period durations are variable. The specific goal of this additional treatment is to understand which one of these factors has more impact on the speed of learning.

Testing alternative learning models would require a specific design and is beyond the scope of the present work. However, our additional treatment can shed some light on which of the above possible explanations is more likely. If the lower cooperation rates in the Short-Stochastic treatment are caused by the variability of the period lengths and the presence of many short periods, then we should expect similarly low rates and slow learning in the Variable-Deterministic treatment. If, instead, the slower learning is caused by the unpredictability introduced with the stochastic horizon, then we should observe higher rates of cooperation in the Variable-Deterministic treatment.

Result 6 *Initial, mean, and median cooperation rates in the Variable-Deterministic treatment are closer to the Short-Deterministic than to the Short-Stochastic treatment. The same can be said for the individual strategies adopted.*

Support for Result 6 comes from Table 10. Initial and mean cooperation rates are significantly higher in the Variable-Deterministic than in the Short-Stochastic treatment (p-value < 0.05 and p-value < 0.10, respectively, according to a linear regression with one observation per subject, and bootstrapped standard errors).[14] Hence, having a stochastic rather than a deterministic horizon substantially changes behavior. Result 6 shows that also under a variable period duration – and despite the short durations

[14]Regression results are reported in Tables A.10 and A.11, in Appendix A.

of many periods – a deterministic horizon helps cooperation more than a stochastic horizon. This finding complements Result 2 and it reinforces the behavioral contrast between continuous-time games and discrete-time games.

	Cooperation rates			
Treatment	**Average**	**Median**	**Initial**	**Final**
Short-Deterministic	63.3	79.2	82.6	15.8
Short-Stochastic	52.3	47.0	65.9	46.8
Variable-Deterministic	57.1	72.7	77.9	18.6

Table 10: Cooperation rates in the Variable-Deterministic treatment

The data from the Variable-Deterministic treatment suggest that both the variability and the unpredictability of period durations had an impact on learning and cooperation, but the unpredictability of the stochastic horizon was the stronger force (Result 6). In the Variable-Deterministic treatment, subjects seem to learn to cooperate faster than in the Short-Stochastic. This evidence is compatible with models of reinforcement learning or replicator dynamics, and corroborates the explanation given in Section 5 for the low cooperation rates in the Short-Stochastic treatment.

Other explanations received less support. Subjects had the same amount of time to think about what was happening in Short-Stochastic and Variable-Deterministic treatments, but cooperation rates were different. On another front, the issue of coordination is compatible with different (mean) cooperation rates across treatments, but if it was the main driver, it should leave unaffected the evolution of cooperation with experience. More specifically, if the lower cooperation rates in the Short-Stochastic treatment are caused by impediments to coordination generated by the uncertain duration, then we should observe a difference in the initial cooperation rates with the Variable-Deterministic treatment, at the beginning of the first period, when behavior cannot be affected by learning from past experience and is mainly determined by introspection. The evidence does not quite support this interpretation, as the initial cooperation rates in the first period of the Short-Stochastic treatment (60.4%) are similar to those in the Short-Deterministic (52.1%) and in the Variable-Deterministic (62.5%) ones (no significant difference emerges according to pairwise z-tests; p-value

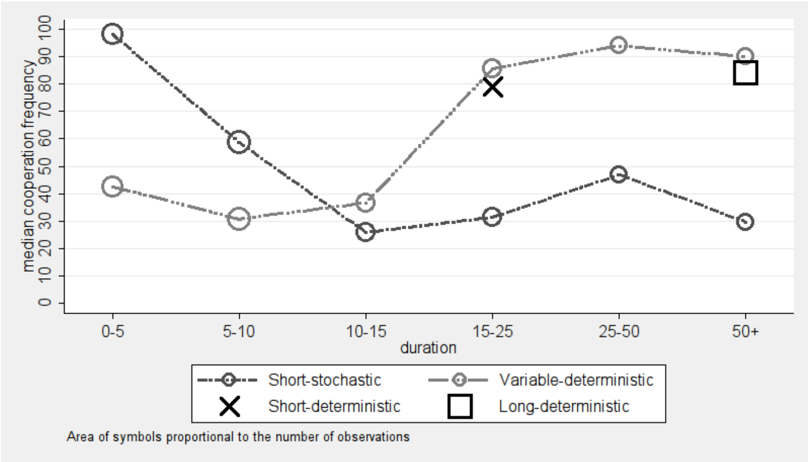


Figure 5: Median cooperation rate, by period duration

> 0.10 in all three cases).

The Variable-Deterministic treatment brings another contribution to our understanding of cooperation and period duration in continuous time, as one can map the relation between period length and observed cooperation rates. In the Variable-Deterministic treatment, the shorter the period the lower the average cooperation (Figure 5). We knew this from comparing the Short- vs. Long-Deterministic treatments (Result 1), but there the difference in cooperation rates was quite small. Given the wide range of period durations available, one can clearly see a monotonic relation between initial cooperation rates and period duration (with a kink around 10-15 seconds that makes it non-linear).[15] This finding complements and reinforces Result 1. In terms of individual strategies employed, the Variable-Deterministic treatment shows a marked difference with the Short-Stochastic, in a direction that confirms and reinforces Result 4. The Cut-off strategy is as widespread in the Variable-Deterministic treatment (58%, Table 9) as it is in the other two deterministic treatments, providing additional support to the theories of continuous-time games, which predict the adoption of this kind of strategy when the horizon is deterministic.

[15]The presence of the kink suggests that when the horizon is deterministic, most subjects cooperate only when the end is far enough, while in very short periods they defect from the start.

7 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.[16] 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, uncertain horizon in the lab under a stochastic continuation/termination rule for the repeated game. Selten et al. (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

[16] See e.g. Selten and Stoecker (1986), Andreoni and Miller (1993), Cooper et al. (1996), Hauk and Nagel (2001) and Bereby-Meyer and Roth (2006).

cooperation rates.[17]

Closest to our work is Friedman and Oprea (2012), where subjects play a symmetric Prisoner's Dilemma in which they could switch actions with latency times on the order of 0.02 seconds (for a total period length of 60 seconds), after which the interaction stops with certainty and subjects are rematched to play another continuous-time supergame. Observed rates of cooperation after some experience reach a median between 81% and 93% and cooperation is typically sustained until the very last seconds of the game, when a short but drastic end-game effect takes place.[18]

The present study differs from Friedman and Oprea (2012) because it considers continuous-time Prisoner's Dilemma games both under a deterministic and a stochastic time horizon. Moreover, it does so for games of different durations (60 and 20 seconds in expectation). It also includes a treatment where the horizon is deterministic and the game duration varies across periods. There are additional differences in other dimensions: the stage-game payoffs; the protocol to match subjects across supergames; 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 subjects learn with experience to 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 close to the end, cooperation rates fall more the more subjects gain experience, the opposite pattern than the one we observe in continuous time.

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 et al., 2006,

[17]See also Palfrey and Rosenthal (1994), who compared contributions to a public good in one shot vs. indefinitely repeated games. Engle-Warnick and Slonim (2004) report little difference when comparing a trust game repeated exactly five times vs. repeated with a continuation probability of 0.8.

[18]Charness et al. (2011) ran a 4-person public good experiment in continuous time and report a somewhat lower impact of continuous-time interaction on cooperation.

2007), minimum effort games (Deck and Nikiforakis, 2012), and hawk-doves games (Oprea et al., 2011).

8 Conclusions

We studied Prisoner's Dilemma games of different durations in continuous time, under deterministic and stochastic horizons. The experiment showed how behavior in (quasi) continuous-time games is qualitatively different from standard repeated games with discrete periods. The main findings are as follows. For long duration treatments, cooperation rates were on a similar, high level in the deterministic and stochastic horizon treatments. With short duration, cooperation rates were significantly higher in deterministic than in stochastic horizon treatments. As subjects gained experience, cooperation rates grew in all treatments. Moreover, in treatments with a deterministic horizon we observed an end-game effect, which got shorter and shorter as subjects acquired experience. The time horizon also had significant impact on the strategies employed. In deterministic horizon treatments, subjects widely employed cut-off strategies such as "Cooperate until time T and then defect forever."

Taken together, the observed patterns are not entirely consistent with existing equilibrium-based theoretical frameworks, whether they model the situation directly in continuous time, as the limit of standard games in discrete time, or as the limit of perturbed games in discrete time. The latter models (Kreps et al., 1982; Radner, 1986, and Friedman and Oprea, 2012), however, appear the most promising theories for continuous-time behavior (Table 1), especially if we were to focus on behavior in the last few supergames.

Since none of these equilibrium-based models found full empirical support, given the dataset as a whole, we turned attention to learning as a potential explanation. Our aim was to highlight the overwhelming impact of experience on behavior in continuous-time games, rather than carrying out a horse race among different learning models. We argued that a simple evolutionary model based on the replicator dynamics can account for most of the observed patterns. In particular, the model predicts that experience is an important drive toward cooperation in all treatments, but that its impact is weaker in games with a stochastic horizon. An additional treatment with deterministic duration and periods of variable length matching the realized durations of the stochastic treatment confirmed that the presence of uncertainty in period duration slows down the convergence to cooperation more than the variability in period lengths.

Theoretical and experimental analyses of repeated games in discrete time have identified two economic forces that influence whether agents can successfully cooperate: the backward induction reasoning in finitely repeated games and the tradeoff between immediate gains and the shadow of future punishments. Our results suggest that – in situations where agents can react quickly to moves by others – both these forces are second-order. Instead, the ability to learn from past successes of cooperative strategies appears to be a first-order determinant of cooperation levels.

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 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, i.e. 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 carrying together 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 period lengths or other details. In particular, the introduction 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 Additional tables

Dependent variable: cooperation rate		
	Coefficient	(s.e.)
Short-Deterministic	-6.082**	(2.950)
Long-Stochastic	1.998	(3.632)
Short-Stochastic	-17.755***	(3.781)
Constant	62.600***	(9.984)
Controls for individual characteristics	Yes	
N	4416	
R-squared overall	0.047	
R-squared between	0.223	
R-squared within	0.000	

Notes: Panel regression with random effects at the subjects' level and standard errors robust for heteroschedasticity. The unit of obs. is the fraction of time a subject spends cooperating within a period. Default treatment: Long-Deterministic. The difference between coefficients for the Short-Stochastic and Short-Deterministic treatments is significant at any standard significance level (p-value < 0.001).

Table A.1: Panel regression on cooperation rates

	Termination rule		
Duration	**Deterministic**		**Stochastic**
Long	0.613	$\sim$	0.608
	(0.791)		(0.769)
	$\vee^{***}$		$\vee^{***}$
Short	0.557	$>^{***}$	0.447
	(0.700)		(0.283)

Notes: The mean rate of mutual cooperation of a session is the average across all 23 periods and all 12 groups in each period. Median rates of mutual cooperation are reported in parentheses. Significance levels are derived from the regression presented in Model 1 of Table A.3.

Table A.2: Average rate of mutual cooperation per second

Dependent variable: rate of mutual cooperation				
	Model 1		Model 2	
	Coefficient	(s.e.)	Coefficient	(s.e.)
Short-Deterministic	-0.091***	(0.032)	-0.091***	(0.029)
Long-Stochastic	0.007	(0.041)	0.007	(0.036)
Short-Stochastic	-0.206***	(0.038)	-0.206***	(0.038)
Constant	0.613***	(0.123)	0.613***	(0.099)
Controls for individual characteristics	Yes		Yes	
N	192		4416	
R-squared	0.259			
R-squared overall			0.048	
R-squared between			0.259	
R-squared within			0.000	

Notes: Model 1 presents results from a linear regression with bootstrapped standard errors. The unit of observation is the average fraction of time a pair of subjects coordinate on cooperation within a period, across all periods. Model 2 presents results from a panel regression with random effects at the subjects' level and standard errors robust for heteroschedasticity. The unit of observation is the fraction of a period duration in which both subjects in a pair cooperate. Default treatment: Long-Deterministic.

Table A.3: Linear regression on rates of mutual cooperation

	Termination rule		
Duration	**Deterministic**		**Stochastic**
Long	0.614	$\sim$	0.608
	(0.792)		(0.793)
	$\vee^{**}$		$\vee^{***}$
Short	0.570	$>^{***}$	0.447
	(0.696)		(0.455)

Notes: The mean profit per second of a session is the average across all 23 periods and all 24 subjects. Median profits are reported in parentheses. Significance levels are derived from the regression presented in Model 1 of Table A.5.

Table A.4: Average profits per second

Dependent variable: average profit per second				
	Model 1		Model 2	
	Coefficient	(s.e.)	Coefficient	(s.e.)
Short-Deterministic	-0.062**	(0.030)	-0.062**	(0.025)
Long-Stochastic	0.019	(0.031)	0.019	(0.029)
Short-Stochastic	-0.188***	(0.030)	-0.188***	(0.030)
Constant	0.685***	(0.096)	0.685***	(0.086)
Controls for individual characteristics	Yes		Yes	
N	192		4416	
R-squared	0.353			
R-squared overall			0.034	
R-squared between			0.353	
R-squared within			0.000	

Notes: Model 1 presents results from a linear regression with bootstrapped standard errors. The unit of observation is the average profit per second, across all periods. Model 2 presents results from a panel regression with random effects at the subjects' level and standard errors robust for heteroschedasticity. The unit of observation is a subject's profit per second, in a period. Default treatment: Long-Deterministic.

Table A.5: Linear regression on average profits per second

Dependent variable: average timing of the end-game effect		
	Coefficient	(s.e.)
Short-Deterministic	-7.182	(4.413)
Period	-1.178**	(0.576)
Period^2	0.027	(0.021)
Period $\times$ Short-Deterministic	0.609	(0.609)
Period$^2\times$ Short-Deterministic	-0.012	(0.023)
Reaction time	4.135	(2.764)
Reaction time$\times$ Short-Deterministic	-2.639	(2.842)
Constant	14.956***	(4.414)
N	92	
R-squared	0.680	

Notes: Linear regression with bootstrapped standard errors. The unit of observation is the session average of timing of the end-game effect (in seconds) in each period. Default treatment: Long-Deterministic.

Table A.6: Linear regression on the average timing of the end-game effect

	Periods				
Treatment	1-6	7-12	13-18	19-23	Overall
Long-Deterministic	1.73	1.38	1.52	1.44	1.52
	N=236	*N=262*	*N=244*	*N=212*	*N=954*
Short-Deterministic	1.06	1.08	0.95	0.90	0.99
	N=164	*N=186*	*N=224*	*N=192*	*N=766*

Table A.7: Average reaction time across periods.

Dep. variable: cooperation rate	**Treatment**			
	Long-Det.	Long-Stoc.	Short-Det.	Short-Stoc.
Period	4.877***	2.615***	4.731***	0.417
	(1.103)	(0.678)	(0.713)	(1.053)
Period^2	-0.122***	-0.045*	-0.094***	0.038
	(0.044)	(0.026)	(0.028)	(0.040)
Period's duration		-0.066***		-0.228**
		(0.013)		(0.106)
Duration of the previous period		0.047*		0.113*
		(0.026)		(0.062)
Constant	29.963***	45.438***	24.259***	42.541***
	(6.396)	(4.268)	(3.743)	(6.148)
N	46	46	46	46
R-squared	0.725	0.695	0.829	0.415

Notes: Linear regression with bootstrapped standard errors. The unit of observation is the average cooperation rate in a session, by period.

Table A.8: Linear regression on cooperation and experience

Dep. variable: cooperation rate	**Treatment**			
	Long-Det.	Long-Stoc.	Short-Det.	Short-Stoc.
Period	4.877***	2.615***	4.731***	0.355
	(0.774)	(0.650)	(0.581)	(0.919)
Period^2	-0.122***	-0.045*	-0.094***	0.041
	(0.031)	(0.026)	(0.024)	(0.032)
period's duration		-0.066***		-0.200***
		(0.017)		(0.050)
duration of the previous period		0.046***		0.140***
		(0.012)		(0.037)
Constant	29.963***	45.466***	24.259***	41.576***
	(3.961)	(4.772)	(3.176)	(5.844)
N	1104	1104	1104	1104
R-squared overall	0.144	0.094	0.244	0.066
R-squared between	0.000	0.006	0.000	0.037
R-squared within	0.000	0.125	0.000	0.081

Notes: Panel regression with random effects at the subjects' level. Standard errors robust for heteroschedasticity. The unit of observation is the individual cooperation rate of a subject in a period.

Table A.9: Panel regression on cooperation and experience

	Mean coop. rate	**Initial coop. rate**
Short-Deterministic	10.802***	15.757***
	(4.060)	(5.425)
Variable-Deterministic	6.353*	12.581**
	(3.703)	(5.281)
Constant	37.996**	65.389***
	(17.023)	(19.930)
Controls for individual characteristics	Yes	Yes
N	144	144
R-squared	0.169	0.226

Notes: Linear regressions with bootstrapped standard errors. The unit of observation is the mean (initial) cooperation rate in a section in a period. Default treatment: Short-Stochastic. Standard errors are reported in parentheses.

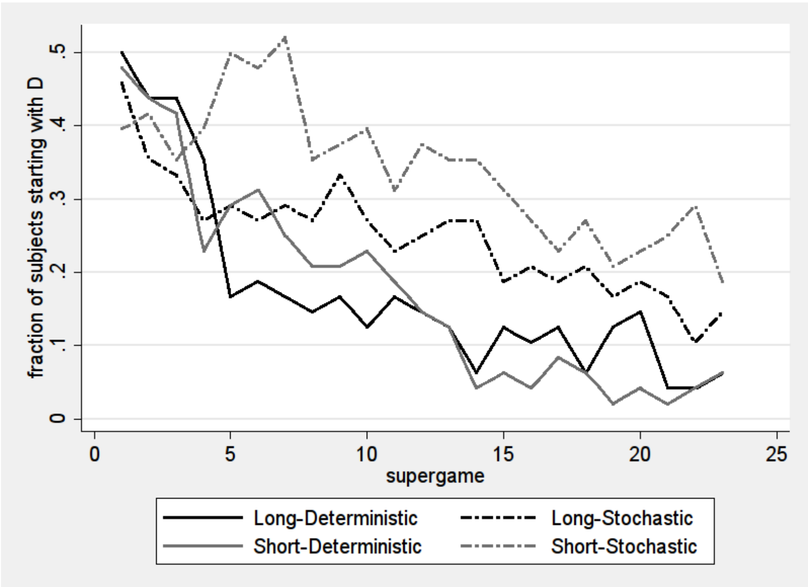
Table A.10: Linear regression on mean and initial cooperation rates

	Mean coop. rate	**Initial coop. rate**
Short-Deterministic	10.802***	0.681***
	(3.398)	(0.148)
Variable-Deterministic	6.353	0.409***
	(3.928)	(0.154)
Constant	37.996**	-0.795**
	(14.892)	(0.390)
Controls for individual characteristics	Yes	Yes
N	3312	3312
R-squared overall	0.032	
R-squared between	0.169	
R-squared within	0.000	
Log-likelihood		-1517.3

Notes: Panel regression with random effects at the subjects' level and standard errors robust for heteroschedasticity. The unit of observation is a subject's cooperation rate/initial action a period. Since initial cooperation is a binary variable, in the last column of this table we present results from a panel logit regression with random effects at the subjects' level and standard errors robust for heteroschedasticity. Default treatment: Short-Stochastic. Standard errors are reported in parentheses.

Table A.11: Panel regression on mean and initial cooperation rates

Figure A.1: Initial defections



B Timing of the end-game effect

B.1 Friedman and Oprea (2012)

According to the model developed by Friedman and Oprea (2012, p.352), it can be shown that – under certain conditions – all strategies employed in a nearly dominant ϵ-equilibrium involve cut-off strategies. That is, strategies that start as grim-trigger strategies and, if there is no prior defection, switch to permanent defection no earlier than some fraction s_L of the total duration of the game.

$$s_L = 1 - \tau \frac{2x}{10 - y}$$

where τ denotes the response time as a fraction of the total period duration, and x and y are the temptation and the punishment payoffs in the following, generalized matrix for the prisoner's dilemma.

	C		**D**	
C	10	10	0	x
D	x	0	y	y

Table B.1: Generalized payoff matrix in Friedman and Oprea (2012)

In other words, this model predicts that the end-game effects would arise no sooner than $\frac{2x}{10-y}t$ seconds from the end, where t is the response time in absolute terms.

To obtain a specific prediction of the timing of the end-game effect in our setup, we map the payoff matrix adopted in our experiment to the matrix in Table B.1, by applying to each payoff π the following linear transformation: $\pi' = (\pi + 2)\frac{10}{3}$. This yields:

	C		**D**	
C	10	10	0	$\frac{40}{3}$
D	$\frac{40}{3}$	0	$\frac{20}{3}$	$\frac{20}{3}$

Table B.2: Payoff matrix adopted in our setup, mapped to Friedman and Oprea (2012)'s model.

Hence, in our set-up, $x = \frac{40}{3}$ and $y = \frac{20}{3}$, and the end-game effect should emerge no sooner than $8t$ seconds from the end.

To conclude, Friedman and Oprea (2012)'s model predicts that, if the game has a finite horizon, an end-game effect will emerge, and that the timing of this effect in terms of distance from the end of the period does not depend on the total length of the period, but only on subjects' reaction time.

B.2 Kreps et al. (1982)

Along the lines of Kreps et al. (1982), let us assume that one player, say the column player (COL) is not absolutely certain that the other (ROW) will play rationally, and assesses that, with probability δ, ROW is playing a Grim-trigger strategy that prescribes to cooperate as long as the opponent cooperates, and to switch to permanent defection as soon as possible after a deviation by his opponent. To incorporate the reaction time, we model the game as having a finite number of periods with a period length equal to τ.

If we restrict attention to sequential equilibria that are not Pareto-dominated by any other sequential equilibrium, we can show that both players should cooperate in all but the last "few" seconds before the end of the game, and that the duration of this final, non-cooperative phase does not depend on the total duration of the period but only on the response time τ and on the ex-ante expectations that the other player is not fully rational, δ.

To keep the same notation as in the previous subsection, let us normalize to 1 the total duration of the period, and let us refer to the generalized payoff matrix in Table B.1. In the following arguments, we stick as close as possible to Kreps et al. (1982), and simply translate their reasoning to our environment, and use the payoff matrix in Table B.1. Notice that the statement of each of the following steps but the last one should be preceded by: "In every sequential equilibrium...". All the statements refer to times that are on the grid of periods defined by τ and the length of the supergame (and we ignore integer problems; similar reasoning can be presented if the game was modeled as a continuous-time game with a response time τ but that would require additional notation and so we present this simpler reasoning).

Step 1: ...if at time $\hat{s} \in [0, 1]$ it is common knowledge that ROW is rational, both ROW and COL will defect at any time $s \in [\hat{s}, 1]$, and their payoffs from the reminder of the period are $(1 - \hat{s})y$.

This follows by a standard backward induction argument in finitely repeated games, since once there is common knowledge that the players are rational, beliefs do not change in a sequential equilibrium.

Step 2: ...if COL defects at time $\hat{s}$, then ROW will defect at any time $s \in [\hat{s} + \tau, 1]$, with probability one.

If ROW cooperated in $\hat{s} + \tau$, it would become common knowledge that he is rational, hence the continuation payoff from $\hat{s} + 2\tau$ on would be y. Because cooperating would not increase the continuation payoff, and would strictly decrease the instantaneous payoff at $\hat{s} + \tau$, defection does strictly better overall, and ROW will defect with probability 1.

Step 3: ...starting from any point $\hat{s} \in [0, 1]$ where COL assesses that ROW is a grim-trigger player with probability ρ and where both players cooperated at any time $s \in [0, \hat{s})$, if COL cooperates at $\hat{s}$ his expected payoff for the reminder of the game is no less than

$$\rho(1 - \hat{s})10 + (1 - \rho)(1 - \hat{s} - \tau)y$$

To see this, consider that if ROW is a grim-trigger player, COL's payoff will not be lower than $(1-\hat{s})10$, while if ROW is a rational player, the worst that can happen is that he defects at $\hat{s}$, thus revealing his type; hence COL will earn the sucker's payoff (0) as long as he does not switch to defection as well (which occurs with delay τ). Afterwards, both players will defect until the end of the period.

Step 4: ...starting from any point $\hat{s} \in [0, 1]$ where both players cooperated at any time $s \in [0, \hat{s})$, if COL defects at $\hat{s}$ his expected payoff for the reminder of the game is no more than

$$\tau x + (1 - \hat{s} - \tau)y$$

To see this, consider that in the best case scenario (from the point of view of COL), ROW is a grim-trigger player and will not defect at $\hat{s}$; hence COL will be able to get the temptation payoff x as long as ROW does not switch to defection, which will occur with delay τ.

Step 5: ... at $\hat{s} = 1 - \tau$ where COL assesses that ROW is a Grim-trigger player with probability ρ and where both players cooperated at any time $s \in [0, \hat{s})$, if COL cooperates at $\hat{s}$ his expected payoff for the reminder of the game is no more than

$$\rho\tau 10$$

while if COL defects his expected payoff is no less than

$$\rho\tau x + (1 - \rho)\tau y$$

Hence, a rational COL player should defect at $\hat{s} = 1 - \tau$, for any $\rho \in [0, 1]$.

Step 6: From Step 5, it results that a rational COL player should defect no later than

$$s_U = 1 - \tau$$

and from Steps 3 and 4 it results that he should defect no earlier than

$$s_L(\rho) = 1 - \tau \frac{x - \rho y}{\rho(10 - y)}$$

because, for any $s < s_L(\rho)$, the lowest possible expected payoff from cooperation (Step 3) is higher than the maximal payoff from defection (Step 4).

At the beginning of the game, the belief is $\rho = \delta$ and hence, if $s_L(\delta) > 0$, a rational COL player has strict incentives to cooperate as does the rational ROW player. That means the belief would remain equal to δ until the period τ and so on until time $s_L(\delta) > 0$. In other words, this model predicts that the end-game effects would arise no sooner than $\frac{x - \delta y}{\delta(10-y)} t$ seconds from the end, where t is the response time in absolute terms and what the equilibrium looks like after time $s_L(\delta)$ is independent of the duration of the game (it can depend on the duration indirectly if duration affects reaction time or beliefs). With our parameters, this implies that the end-game effect should not emerge earlier than $2t\left(\frac{2}{\delta} - 1\right)$ seconds from the end.

To conclude, as Friedman and Oprea (2012)'s as well as Kreps et al. (1982)'s models predict, if the game has a finite horizon, an end-game effect will emerge and the timing of this effect in terms of distance from the end of the period does not depend on the total length of the period, but only on subjects' reaction time. Differently from Friedman and Oprea (2012), Kreps et al. (1982)'s model also highlights the role of subjects' expectations on others' rationality.

C Proof of Proposition 2

The area of the basin of attraction of the strategy Always-defect is given by the sum of the areas of the two triangles whose vertices identify the following states.[19] $\mathbf{x_2} = (0, 1, 0)$,
$\mathbf{x_4} = (0, 1 - \frac{b}{a}, \frac{b}{a})$,
$\mathbf{x_5} = (b, 1 - b, 0)$, and
$\mathbf{x_6} = (\frac{b^2}{a(1-a)+b^2}, 1 - \frac{(1-a)b}{a(1-a)+b^2}, \frac{(1-a)b-b^2}{a(1-a)+b^2})$.

This area can be expressed as a function $A(a, b)$ of parameters a and b:

$$A(a,b) = \frac{(a-1)b^2 \left(a \sin\left(\frac{\pi(a+b-1)}{3(a-1)}\right) + \sin\left(\frac{\pi b}{3-3a}\right)\right)}{2a\left(a^2 - a - b^2\right)} \tag{3}$$

In Section 6, we computed the value of a and b as a function of the relative response time τ and of a parameter k, which characterizes the timing of the end-game effect.
For games with a deterministic horizon, we have:

$$a = 1 - k \cdot \tau \tag{4}$$
$$b = 2 \cdot \tau \tag{5}$$

Hence, replacing (4) and (5) in equation (3), we get the area of the basin of attraction of strategy Always-defect as a function of τ and k:

$$A_{Det}(k,\tau) = -\frac{2k\tau^2 \left((k\tau - 1)\cos\left(\frac{\pi(k+4)}{6k}\right) - \sin\left(\frac{2\pi}{3k}\right)\right)}{(k\tau - 1)\left(k^2\tau - k - 4\tau\right)}$$

For games with a stochastic horizon, we have:

$$a = 1 - e^{k\cdot\tau - 1} \tag{6}$$
$$b = 2(1 - e^{-\tau}) \tag{7}$$

Hence, replacing (6) and (7) in equation (3), we get the area of the basin

[19] In these triples, the first component represents the fraction of agents playing Grim-trigger, the second element the fraction of agents playing Always-defect, and the third element corresponds to the fraction of agents playing the Cut-off strategy, as explained in footnote 11, at page 21.

of attraction of the strategy Always-defect as a function of τ and k:

$$
\begin{aligned}
A_{Stoch}(k,\tau) =& 2\left(e^{\tau}-1\right)^{2} e^{k\tau+1} \times \\
&\left(\frac{\left(e^{k\tau}-e\right)\sin\left(\frac{1}{3}\pi\left(1-2\left(e^{\tau}-1\right)e^{1-(k+1)\tau}\right)\right)}{\left(e^{k\tau}-e\right)\left(-e^{2(k+1)\tau}+e^{(k+2)\tau+1}-8e^{\tau+2}+4e^{2\tau+2}+4e^{2}\right)} - \right.\\
&\left.\frac{e\sin\left(\frac{2}{3}\pi\left(e^{\tau}-1\right)e^{1-(k+1)\tau}\right)}{\left(e^{k\tau}-e\right)\left(-e^{2(k+1)\tau}+e^{(k+2)\tau+1}-8e^{\tau+2}+4e^{2\tau+2}+4e^{2}\right)}\right)
\end{aligned}
$$

Let $D(k,\tau) = A_{Det}(k,\tau) - A_{Stoch}(k,\tau)$. We have that:

$$
\begin{aligned}
D(k,0) &= 0 \ \forall k \\
\lim_{\tau\to 0} D'(k,0) &= 0 \ \forall k \\
\lim_{\tau\to 0} D''(k,0) &< 0 \ \forall k \geq 1
\end{aligned}
$$

We have thus proven, by direct computation, that $D(k,\tau) < 0$ for values of τ close enough to 0, and for any value of $k \geq 1$.

D Instructions

[Instructions for the Long-Stochastic treatment, translated from Italian. the parts that are different in the Long-Deterministic treatment are reported in *italics.*]

Welcome! This is a study about how people make economic decisions. This study is funded by the University of Bologna and other institutions. If you pay attention, the instructions will help you to make your decisions and earn a reasonable amount of money. The earnings will be calculated in points and then converted into euros.

For every 150 points you will receive 1 euro.

In addition, you will receive 3 euros for participation. Your earnings will be paid in cash at the end of today's session.

We ask that you turn off your phone now and do not communicate in any way with the people present in the room until the end of the study. If you have any questions, please raise your hand and we will assist you in private.

This study comprises **23 periods**. In each period, you will be paired with another person selected at random from those present in the room.

In every period you will be able to **repeatedly choose between a "GREEN" action and an "ORANGE" action**. The person matched with you will also be able to repeatedly choose between "green" and "orange" actions. As a consequence, there are four possible combinations: GREEN-green, ORANGE-orange, GREEN-orange, and ORANGE-green. For each combination of actions there is a corresponding cell in Figure D.1 below.

In each cell you can see the gains or losses during the period according to your action and the action of the other. Your action will determine the table row, while the action of the person matched with you will determine the table column.

The earnings described in Figure D.1 above represent earnings per second.

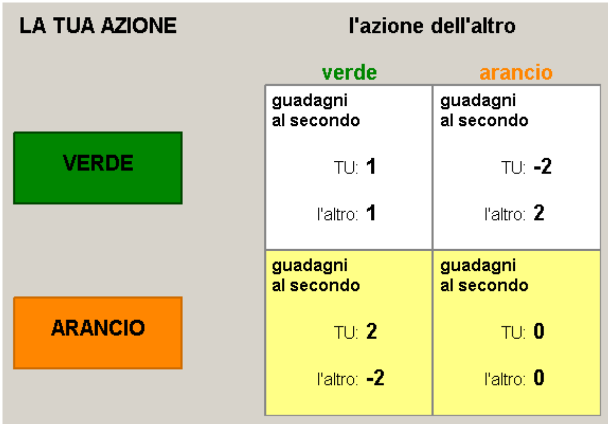


Figure D.1: Earnings table

For instance, suppose you choose "GREEN" and hold that choice over time: if the other chooses "green" and holds his choice in time, you earn 1 point per second and the other earns 1 point per second; if the other chooses "orange" and holds it, you lose 2 points per second and the other earns 2 points per second. And so on.

In each period, earnings depend on how much time you spend in each cell of Figure D.1. The more time you spend in a cell, the more your average earnings will approximate what is indicated in the cell. For instance, if you spend half of the period in the GREEN-green cell where you earn 1 and half of the period in the ORANGE-orange cell where you earn 0, your earnings will be 0.5 points per second. Are there any questions about how to read the table?

Who is the other person matched with me?
It could be anyone in this room. Your identity and hers will be kept confidential. Payments will also be made in private. There will be 23 periods. At the beginning of each period pairs will be changed. People will be recombined so that **you will never meet the same person twice**.

What should I do? In every period you choose an initial action and then you can decide every instant whether to keep or change that action. The person matched with you can do the same. During a period, both you and the other will be able to change action as many times as you like. Time flows through very fast ticks (16-hundredths of a second each); in practice there are between six and seven ticks per second, so if you want you can change the action six or seven times per second.

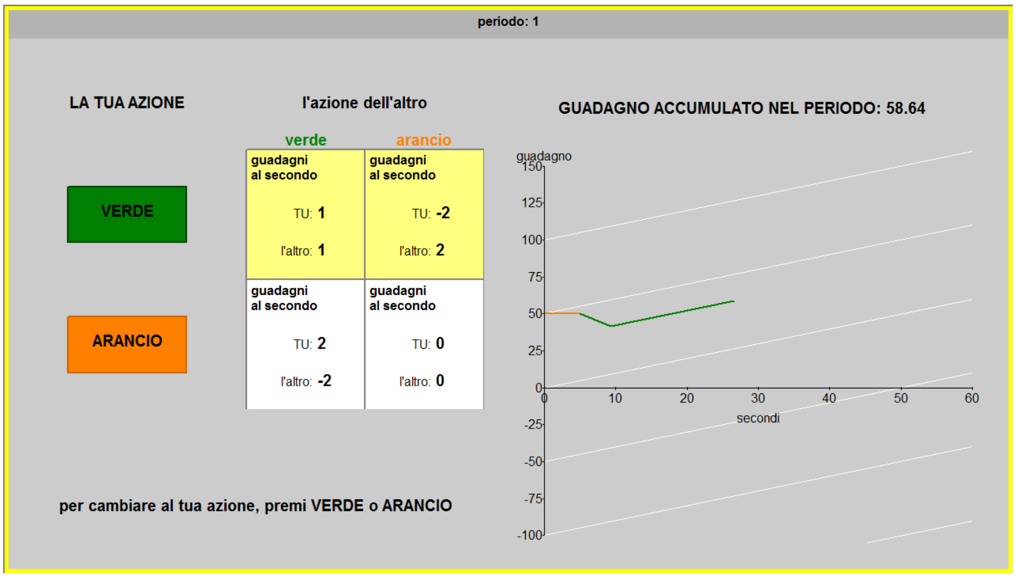


Figure D.2: Earnings table

Earnings

During the period you will receive information in real time on your earnings. In the screen pictured in Figure D.2 above, your cumulated earnings will appear in a graph as a line that will form at every tick of 16-hundredths of a second. In each period you will have an **initial endowment of 50 points** as cumulated earnings. **If, during the period, your earnings are zero, then the line will be flat. In case of losses, then the line will be declining. In the case of positive earnings, then the line will be increasing.** For instance, if you earn 1 point per second there will be an increasing line that is parallel to the graph grid. If you earn 2 points per second, the line will be increasing, but steeper. Looking at the earnings graph will give you information on the current action of the other person matched with you. Are there any questions?

To understand how to read the screen, **we will do a trial period**, without consequences on your earnings. For simplicity, the trial period will last 60 seconds and the other will be played by a robot. The robot will start with an action and then, halfway through the period, will change action. Now please look at the screen and follow the exact guidelines you are given. To start, choose the initial action. Press the screen with your finger on the button that you will be told to choose ("GREEN" or "ORANGE"). The robot will also choose its initial action ("green" or "orange"). **Everyone please choose "GREEN" now** as the initial action. The selected action will be highlighted in yellow on the table. The period will begin when everybody has chosen their initial action and pressed "OK". From this moment on, the time will begin to run. Then you will see that the graph line is green like your action. Now, please press your finger on the button "OK" to confirm. Does anyone need help? After 10 seconds, everyone please press the button "ORANGE." You will see that your action has changed because the line highlighted in yellow in the table will change and that indicates your current action. Moreover, the graph line will now be orange in color. After 30 seconds, everyone please press again the button "GREEN." Now we ask you to guess what actions the robot chose. Are there any questions?

We will do two more trial periods, without consequences on your earnings. For simplicity, the trial period will last 60 seconds and the other will be played by a robot. The robot will start with an action and then, halfway through the period, will change action. Now look at your screen. Choose the initial action that you prefer. When everyone has completed, you'll see the time running. You are free to change the action at any time. At the end of the period, we will ask you to guess what actions the robot chose.

Now we will do the last trial period. Go ahead and choose the action you want. Are there any questions?

For simplicity, in the trial periods the other was a robot and the duration was 60 seconds. However, in the coming periods, the other will be a person in this room *while the duration of each period will be variable and determined randomly. Each period will stop without notice and for everybody at the same moment, and the period duration could vary from less than a second to several minutes.*

How is a period duration established?

The period may stop at every tick of 16 hundredths of a second. This event depends on the result of a random draw. Imagine a box with 10,000 balls, of which 9,973 black and 27 white. It is as if, after every tick, a ball was drawn. If the ball drawn is white, the period ends. If the ball is black, the period continues and the ball is placed back into the box. At the next tick, a new ball is drawn at random. You have to imagine very rapid draws, that is one every tick of 16 hundredths of a second. We calculated that as a result of this, the periods will have an average duration of 60 seconds. There may be some short periods and some long periods. Are there any questions about this?
[DETERMINISTIC: The length of each period will be 60 seconds.]

Very well, then we can start.

References

Andreoni, J. and J. H. Miller (1993): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *The Economic Journal*, 103, 570–585.

Bereby-Meyer, Y. and A. E. Roth (2006): "The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation," *The American Economic Review*, 96, 1029–1042.

Bergin, J. and W. B. MacLeod (1993): "Continuous Time Repeated Games," *International Economic Review*, 34, 21–37.

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, 10.1007/s10683-006-9125-1.

Berninghaus, S., K.-M. Ehrhart, M. Ott, and B. Vogt (2007): "Evolution of networks—an experimental analysis," *Journal of Evolutionary Economics*, 17, 317–347, 10.1007/s00191-006-0050-4.

Börgers, T. and R. Sarin (1997): "Learning Through Reinforcement and Replicator Dynamics," *Journal of Economic Theory*, 77, 1 – 14.

Charness, G., D. Friedman, and R. Oprea (2011): "Continuous Time and Communication in a Public-goods Experiment," mimeo.

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, 187 – 218.

Dal Bó, P. (2005): "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games," *American Economic Review*, 95, 1591–1604.

Dal Bó, P. and G. Fréchette (2011): "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review*, 101, 411–429.

Deck, C. and N. Nikiforakis (2012): "Perfect and imperfect real-time monitoring in a minimum-effort game," *Experimental Economics*, 15, 71–88, 10.1007/s10683-011-9289-1.

Engle-Warnick, J. and R. L. Slonim (2004): "The Evolution of Strategies in a Repeated Trust Game," *Journal of Economic Behavior & Organization*, 55, 553 – 573, trust and Trustworthiness.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10, 171–178.

Friedman, D. and R. Oprea (2012): "A Continuous Dilemma," *American Economic Review*, 102, 337–63.

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

Hauk, E. and R. Nagel (2001): "Choice of Partners in Multiple Two-Person Prisoner's Dilemma Games: An Experimental Study," *The Journal of Conflict Resolution*, 45, 770–793.

Kahneman, D. (1973): *Attention and Effort*, Prentice Hall.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27, 245–252.

Murnighan, J. K. and A. E. Roth (1983): "Expecting Continued Play in Prisoner's Dilemma Games: A Test of Several Models," *The Journal of Conflict Resolution*, 27, 279–300.

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, 10.1007/s00182-012-0341-y.

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, 2206 – 2225.

Palfrey, T. R. and H. Rosenthal (1994): "Repeated Play, Cooperation and Coordination: An Experimental Study," *The Review of Economic Studies*, 61, 545–565.

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

Rapoport, A. and A. Chammah (1965): *Prisoner's Dilemma: a Study in Conflict and Cooperation*, Ann Arbor, University of Michigan Press.

Roth, A. E. and J. K. Murnighan (1978): "Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma," *Journal of Mathematical Psychology*, 17, 189 – 198.

Sannikov, Y. and A. Skrzypacz (2007): "Impossibility of Collusion under Imperfect Monitoring with Flexible Production," *American Economic Review*, 97, 1794–1823.

Selten, R., M. Mitzkewitz, and G. R. Uhlich (1997): "Duopoly Strategies Programmed by Experienced Players," *Econometrica*, 65, 517–556.

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, 47 – 70.

Simon, H. A. (1971): "Designing Organizations for an Information-Rich World," in *Computers, Communications, and the Public Interest*, ed. by M. Greenberger, Baltimore: The Johns Hopkins Press, 37–72.

Simon, L. K. and M. B. Stinchcombe (1989): "Extensive Form Games in Continuous Time: Pure Strategies," *Econometrica*, 57, 1171–1214.

Sims, C. (2003): "Implications of Rational Inattention," *Journal of monetary Economics*, 50, 665–690.

Szymanski, S. (1996): "Making Hay While the Sun Shines," *The Journal of Industrial Economics*, 44, 1–16.

Weibull, J. (1995): *Evolutionary Game Theory*, Cambridge, USA: MIT Press.