

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Time Horizon and Cooperation in Continuous Time

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Time Horizon and Cooperation in Continuous Time / Bigoni, Maria; Casari, Marco; Skrzypacz, Andrzej; Spagnolo, Giancarlo. - In: *ECONOMETRICA*. - ISSN 0012-9682. - STAMPA. - 83:2(2015), pp. 587-616. [10.3982/ECTA11380]

Availability:

This version is available at: <https://hdl.handle.net/11585/514727> since: 2020-02-29

Published:

DOI: <http://doi.org/10.3982/ECTA11380>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the peer reviewed version of the following article:

Bigoni, M., Casari, M., Skrzypacz, A., Spagnolo, G., 2015. Time Horizon and Cooperation in Continuous Time. *Econometrica*. DOI:10.3982/ECTA11380

which has been published in final form at <https://doi.org/10.3982/ECTA11380>.

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

TIME HORIZON AND COOPERATION IN CONTINUOUS TIME*

MARIA BIGONI[†] MARCO CASARI[‡] ANDRZEJ SKRZYPACZ[§]
GIANCARLO SPAGNOLO[¶]

Abstract

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

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

Keywords: Prisoner's dilemma, Experiment, Folk Theorem, Backward induction

*We thank Pedro Dal Bó, Gary Charness, David Cooper, Dan Friedman, Drew Fudenberg, Rosemarie Nagel, In-Uk Park, Justin Rao, Al Roth, Yossi Spiegel, Simon Weidenholzer, Nathaniel Wilcox, three referees, the editor and, in particular, Guillaume Fréchette for insightful comments and advice. This paper also benefited from comments received by participants at the Luxembourg ESA meeting, the Nordic Workshop in Experimental and Behavioral Economics, the Paris Workshop on Extensive Form Games in the Lab, the Goteborg meeting of the Econometric Society, the Alberobello Workshop on Industrial Organization, CRESSE 2014, and seminar participants at University of Arizona - Tucson, University of Bologna, Chapman University, UC Berkeley, UC San Francisco, UCLA, Düsseldorf Institute for Competition Economics, University of East Anglia, University of Mannheim, and Ewebe in Bertinoro. The usual disclaimer applies. Casari gratefully acknowledges financial support from the Starting ERC grant no. 241196. Spagnolo thanks the Haas School of Business for hospitality and Handelbanken for financial support.

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

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

[§]Graduate School of Business, Stanford University. E-mail: *skrz@stanford.edu*

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

1 Introduction

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

Our experiment combines elements of the design of Dal Bó (2005) and of Friedman et al. (2012). In particular, we study cooperation patterns in (almost) continuously repeated prisoners' dilemmas by comparing deterministic vs. stochastic termination rules for different lengths of the horizon. Our first main finding is that in short games players cooperate more when the horizon is deterministic than when it is stochastic with the same expected duration. This pattern is opposite to that observed by Dal Bó (2005) in discretely repeated games. The second main finding is that while in games with deterministic horizon there is a clear end-game effect of the form that cooperation drops significantly in the last few seconds of each supergame, with experience subjects cooperate longer. This is also in contrast to what has been documented in experiments with discrete rounds, where with experience subjects learn to defect sooner (see Embrey et al., 2013 for a comprehensive review of this literature and latest evidence).

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

While the inherent difference between continuous-time and discrete-time games has been recognized in theoretical work, we offer the first

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

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

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

To shed light on the causes of this unexpected behavior we analyze individual strategies and their evolution across supergames. In all treatments, subjects change their strategies from defection to conditional cooperation as they gain experience. However, we find that the speed of this transition

differs across treatments and is particularly slow in short stochastic treatments, which justifies the overall difference in cooperation rates. We then show that such pattern of behavior is consistent with a simple reinforcement learning model: due to the unpredictability and variability of game duration, subjects receive particularly noisy signals about the relative profits of these types of strategies in short games with stochastic ending and hence it takes longer for the low-payoff strategies to be abandoned.

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

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

2 Related Literature

The repeated (or 'iterated') Prisoner's Dilemma with perfect monitoring has probably been the most important setup in which the question of what leads people to cooperate has been explored experimentally since the early work of Rapoport and Chammah (1965). A central and highly debated issue has been the role played by the time horizon, sometimes called the 'termination rule'. The experimental literature has shown that the theoretical prediction that backward induction should apply to finitely repeated games with the features of a Prisoner's Dilemma often does not hold in the laboratory.¹ In field situations, the moment at which a relationship

¹See e.g. Selten and Stoecker (1986), Andreoni and Miller (1993), Cooper et al. (1996), Hauk and Nagel (2001) and Bereby-Meyer and Roth (2006).

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 cooperation rates.²

Friedman and Oprea (2012) is the first experiment on a Prisoner’s Dilemma in continuous time. In their experiment subjects can switch actions frequently (with latency times on the order of 0.02 seconds) and the game has a deterministic horizon of 60 seconds. Like in our experiments, they observe high cooperation rates (a median between 81% and 93% after subjects gained experience) and, within a supergame, cooperation is typically sustained until the very last seconds of the game, when a short

²Notice however that Normann and Wallace do not allow subjects to gain experience, as they only play one supergame. See also Palfrey and Rosenthal (1994), who compared contributions to a public good in one shot 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. Aoyagi and Fréchette (2009) instead report results for a treatment with continuation probability of 0.9 and one with one-shot games, and report large differences in cooperation.

but drastic end-game effect takes place. They also compare cooperation rates for different frequencies of moves varying from a one-shot game, to eight-moves-per-supergame, to one-move-per-second. They find that games with higher frequency of moves have higher cooperation rates and that this effect is quantitatively large: the eight-moves-per-supergame treatments achieve less than half of the cooperation levels observed in the continuous-time treatment, providing support for theories predicting that a smaller per-period gain from defection (relative to continuation payoffs) stabilizes cooperation.³

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

Our work is also related to experimental studies of finitely repeated games played in discrete time at low frequency that, among other things, investigate whether experience induces subjects to cooperate more and to properly apply backward induction. A consistent finding in this literature, including Selten and Stoecker (1986), Andreoni and Miller (1993), Hauk and Nagel (2001) and Bereby-Meyer and Roth (2006), among others, is that subjects learn to cooperate more at early stages of the finitely repeated game and less at the final stages. Subjects also seemed to learn to

³Charness 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.

anticipate the end game effect, consistent with a partial learning backward induction story. The recent meta analysis of previous work and additional experiments by Embrey et al. (2013) confirm that this is a consistent result in discretely repeated games experiments with deterministic horizon: subjects tend to learn to anticipate the beginning of the end game phase. In our continuous-time environment we find precisely the opposite result.

Finally, the experimental literature on games in continuous time has blossomed during the last few years, so there are several less related studies focusing on strategic situations that are quite different from a Prisoner’s Dilemma, such as games of network formation (Berninghaus et al., 2006, 2007), minimum effort games (Deck and Nikiforakis, 2012), and hawk-doves games (Oprea et al., 2011).

3 Experimental Design

The experiment has five treatments. The two treatment variables are the expected duration of each supergame and the termination rule.⁴

In all treatments, subjects played a series of (quasi) continuous-time Prisoner’s Dilemma games with stage-game payoffs as in Table 1.⁵

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

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

Table 1: Stage game payoffs

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

⁴Table B.1 in the on-line appendix B summarizes the characteristics of each treatment.

⁵As 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. Notice that in the instructions, supergames were referred to as “periods”.

matching).⁶

In all treatments, the stage game was as follows. Each subject had to select an initial action for the supergame, either Cooperate (green) or Defect (orange). When all subjects were done, the supergame began. Within a supergame, subjects could switch action up to six or seven times per second. More precisely, there was a tick every 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.⁷

Earnings for all possible combinations of actions were visible on the screen at all times. The payoff matrix showed earnings in tokens per second. The subject's current action was always highlighted in yellow in the payoff matrix. Moreover, every subject could observe her cumulative earnings on a continuously updated graph. Subjects' earnings in every supergame included an initial endowment (see Table B.1), and could stay constant, increase, or decrease over time, depending on the choices of the pair. The graph showed these patterns of earnings as a flat, increasing, or decreasing line, respectively. A steeper line indicated a faster accumulation or depletion. The line color was green or orange depending on the subject's own action. Hence, from the graph, subjects could unambiguously infer the action taken in any moment by their opponent. The progression of the earnings line marked the timing of the supergame for the subjects. They could observe at every 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 supergame by looking at the graph on the screen.

In the Long-Deterministic treatment, a supergame always lasted 60 seconds. In the Long-Stochastic treatment, a supergame lasted in expectation 60 seconds. Similarly for the short treatments, where the expected duration was 20 seconds. In the stochastic treatments, the exact duration was selected at random supergame by supergame. As explained in the instructions for the Long(Short)-Stochastic treatment, the supergame duration

⁶In the Short-Deterministic session run on February 2, 2011, due to a technical problem, subjects met again their opponents of supergame 1 in supergame 23. All reported results hold even if supergame 23 in that session is dropped.

⁷Figure B.1 in the on-line appendix B presents a screen-shot of the computer interface.

depended on a random 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 supergame (which is called a “period” in the instructions) lasted beyond 60 seconds, the scale of the horizontal axis of the graph automatically shifted forward. The experimenters had no direct control, nor any ex-ante information on the realized supergames’ durations.

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

Stage-game payoffs are such that cooperation should be easily achieved (at least in the stochastic ending treatments). In continuous time, cooperation is always sustainable because the instantaneous discount factor is 1: then a grim trigger strategy should, in theory, support cooperative play as an equilibrium, 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 Table 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.⁸ Subjects had the opportunity to ask questions, which were answered in private, and then went through three practice supergames with a robot opponent that was programmed to switch action in the middle of the supergame. After each practice supergame, subjects had to guess the actions taken by the robot, and then 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.⁹ The session ended with a questionnaire.

The experiment involved 240 subjects, mostly students at the University of Bologna, who took part in only one of the ten sessions and were assigned through an online recruitment software (Greiner, 2004). The experiment was run in the Bologna Laboratory for Experiments in Social Sciences using z-Tree (Fischbacher, 2007). Subjects were seated at visually isolated computer terminals and could not communicate. A session lasted, on average, 2 hours for the Long treatments and 1 hour and 20 minutes for the Short and Variable ones. At the end of the sessions, points were converted into Euors at a rate of 1€every 150 points in the Long treatments, and at a rate of 1€every 50 points in 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 supergame 13 on in Friedman and Oprea (2012) ranges between 81% and 93%, depending on the treatment, and in our Long-Deterministic treatment it is 91%.

⁸An English translation of the experimental instructions is available in the on-line appendix C.

⁹In the three practice supergames, 71% of the subjects always made correct guesses about the sequence of actions taken by the robots. In answering the four control questions about the instructions, 53.8% of the subjects made at most one mistake.

This provides a robustness check of their findings for different procedures, subject pools, and payoff levels. The novelty of this study, however, stems from the comparison across our treatments.

Cooperation rates are higher in supergames of longer (expected) duration, with the difference being large for stochastic ending and small for deterministic ending (Tables 2 and 3). The impact of duration on cooperation rates is highly significant in the stochastic treatments, while being only marginally significant in the deterministic treatments. The unit of observation is the cooperation rate, which is defined as the fraction of time $R_{i,s}$ a subject i spends cooperating within supergame s . Given that these observations are not independent, to assess the significance of the observed differences we take the average cooperation rate by subject across all supergames, and run a linear regression with bootstrapped standard errors. Results are reported in Table 3.¹⁰ The outcome of this regression indicates

Duration:		Termination rule:		
		Deterministic		Stochastic
Long	mean	65.5	~	66.9
	<i>median</i>	<i>84.0</i>		<i>84.8</i>
		√**		√***
Short	mean	63.3	>***	52.3
	<i>median</i>	<i>79.2</i>		<i>47.0</i>

Notes: The mean cooperation rate of a session is the average across all 23 supergames and all 24 subjects. The unit of observation is a subject per supergame. For every treatment there are two sessions and 1104 observations. In this and in the following tables, symbols *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively. Significance levels are derived from the regression presented in Table 3.

Table 2: Cooperation rates by treatment

these differences are significant when controlling for individual characteris-

¹⁰We obtain similar results with a panel regression with random effects at the subject level, where the unit of observation is the cooperation rate of a subject in a supergame, and standard errors are robust for heteroschedasticity (see Table B.2 in the on-line appendix B). As a further robustness check we also ran linear regressions with standard errors robust for clustering at the subject and pair level. The same treatment effects emerge if we compare the rates of mutual cooperation (Tables B.3 and B.4), or the average profits per second (Tables B.5 and B.6). Notice however that the difference between the Long- and Short-deterministic treatment ceases to be statistically significant if we drop controls for individual characteristics, if we focus only on the last five supergames, or if we consider only the first 2/3 of each supergame.

tics.¹¹

Dependent variable: cooperation rate		
	Coefficient	(s.e.)
Short-Deterministic	-5.537**	(2.760)
Long-Stochastic	1.682	(2.577)
Short-Stochastic	-17.799***	(3.830)
Constant	63.798***	(14.128)
Controls for individual characteristics	Yes	
N	192	
R-squared	0.220	

Notes: Linear regression with bootstrapped standard errors. To take care of the potential correlation between observations coming from the same session, we take sessions as resampling clusters. The unit of observation is the fraction of time a subject spends cooperating within a supergame, averaged by subject across all supergames. Default treatment: Long-Deterministic. See footnote 11 for a list of controls of individual characteristics.

Table 3: Linear regression on cooperation rates

Let us now turn to our main result.

Result 1 *With deterministic duration, cooperation rates are equal or higher than with stochastic duration.*

Support for Result 1 comes from Tables 2 and 3. In the long duration treatments, cooperation rates are statistically indistinguishable between stochastic and deterministic horizons (p-value > 0.1 , see Table 3). 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 3). 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

¹¹In 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. A full version of this regression, including the estimated coefficients for all the controls, is reported in Table B.2 in the on-line appendix B.

games in discrete time, where cooperation is typically higher with stochastic than with deterministic duration.¹²

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

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

Support for Result 2 comes from Figure 1 and Table 4. Figure 1 presents the time profile of the mean share of cooperators, taken across supergames and sessions. A subject can change action every 0.16 seconds. Our unit of observation is the share of cooperators S_{tp} over time t within a supergame p .

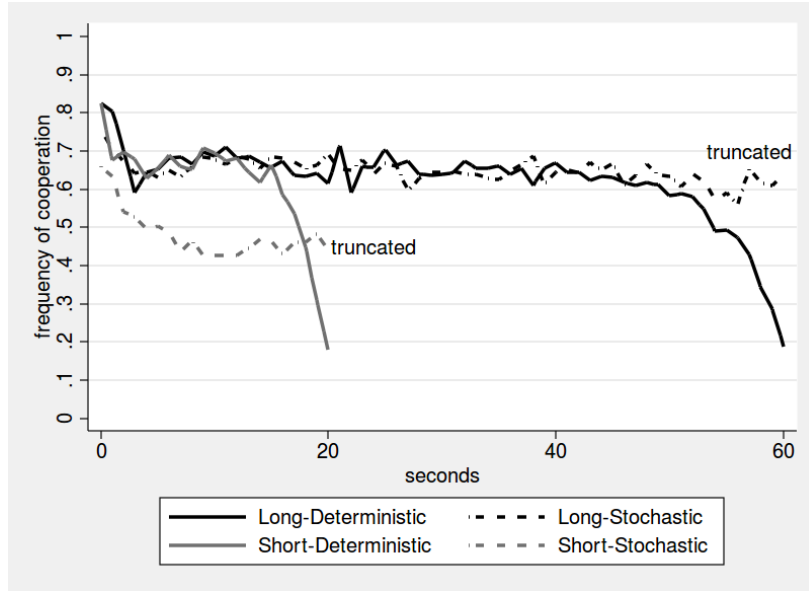
In both the Short- and Long-Deterministic treatments, there is a clear end-game effect: the share of cooperators suddenly drops a few seconds before the end of the supergame (Figure 1).¹³ There are, of course, many ways to quantitatively measure the timing of this switch from cooperation to permanent defection. We measured it by focusing on all pairs that at some point during a supergame reached simultaneous cooperation, CC, and then switched to defection before the end of the supergame, i.e. CD, DC, or DD.¹⁴

¹²For example, in a repeated game with short expected duration, Dal Bó (2005) finds that, “for every round, [...] the percentage of cooperation in infinitely repeated games [...] is greater than in finitely repeated games of the same expected length [...], with p-values of less than 0.01.” (the expected number of action choices is 125 in our short treatments, 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.

¹³Friedman 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.”

¹⁴This calculation includes the lion’s share of the observations. Out of a total of 552 subject-supergame observations per treatment, we have 468 and 460 in the Long and Short treatment, respectively, such that both subjects cooperated simultaneously at least once in that supergame. Of these, some (54/468 and 51/460, respectively) kept on cooperating until the end of the supergame, while in the other cases (414/468 and 409/460, respectively) at least one of the subjects in the pair switched to permanent defection.

Figure 1: Time profile of the share of cooperative actions



Notes: The unit of observation is the share of cooperative actions in every second of a supergame. A subject could switch action every 0.16 of a second. All subjects and all supergames are included for the first 20 or 60 seconds. In the Long-Stochastic treatments, 45.7% of supergames lasted more than 60 seconds. In the Short-Stochastic treatments, 30.4% of supergames lasted more than 20 seconds.

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

Notes: average distance (in seconds) between the end of the supergame, and the time of a permanent switch of a pair from mutual cooperation (CC) to defection – i.e. to CD, DC, or DD.

Table 4: Timing of the end-game effect

It is evident from Table 4 that in the Short-Deterministic treatment, the end-game effect kicks in significantly later than in the Long-Deterministic treatment: about 5 seconds from the end in contrast to about 12 seconds (Table 4, $p\text{-value} < 0.001$).¹⁵ It is also evident that the end-game effect kicks in later and later, as subjects gain experience (3.6 to 10.6 seconds later, Table 4). The impact of experience is significant both in the Long-Deterministic ($p\text{-value} < 0.001$) and in the Short-Deterministic treatment ($p\text{-value} < 0.001$).

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

The correlation between reaction times and timing of the end-game effect, however, is not-significant (Table B.7 in the on-line appendix B). When controlling for the average reaction time in the regression, the decrease in the duration of the end-game effect across supergames is still significant (see Table A.1 in Appendix).

5 Identification of individual strategies

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

The adopted approach to strategy estimation is based on the assumption that subjects (i) adopt one of the 20 strategies presented in Table 5, (ii) may change strategy across supergames, and (iii) may make errors on actions, i.e., with some probability may choose an action that is not recommended by the strategy.¹⁶ We considered all strategies analyzed in

¹⁵The p -values reported in this paragraph are obtained from panel regressions with random effects at the session level. The unit of observation is a session in a supergame. Regression's results are reported in Table A.1 in Appendix.

¹⁶Table A.3 in Appendix presents a description of the 20 strategies.

Fudenberg et al. (2012), which include in particular three sets. Fully cooperative strategies, which prescribe cooperation at all points in time if matched with each other (A); Fully defective strategies, which prescribe permanent defection if matched with each other (B); and other strategies that would prescribe a combination of cooperation and defection (C). Notice that Table 5 does not include a Cut-Off strategy, i.e. a strategy that prescribes to play Grim-trigger until close to the end of the supergame, then switch to permanent defection. The Cut-Off strategy is particularly relevant in deterministic treatments, as it could account for the aforementioned end-game effect (Table 4). The reason why we exclude it is that including it would have required to arbitrarily set the cut-off time. As a consequence, subjects following these strategies will be most likely classified as Grim-trigger, because the two strategies differ by at most one action (the first defection). For strategy estimation, we adopted a methodology introduced in Dal Bó and Fréchette (2011), for repeated games in discrete time.¹⁷ In order to apply this methodology to our continuous-time setup, we considered subjects' actions in a sequence of specific instants of play (every second) and disregarded the rest of the data.¹⁸

For each of the strategies, we define the sequence of actions prescribed by the strategy to the subject, given the sequence of the subject's and the opponent's previous actions. The maximum likelihood procedure is based on the comparison between the actual sequence of actions taken by the subject, and the 20 sequences corresponding to the 20 strategies. It returns an estimate of the frequency with which each strategy is played, in the population.¹⁹

To study how strategy adoption evolved with experience, we carried out

¹⁷The same estimation approach was also used, among others, by Fudenberg et al. (2012); Camera et al. (2012); Dal Bó and Fréchette (2012); Fréchette and Yuksel (2013); Rand et al. (2013); Vespa (2013).

¹⁸Using the whole sequence of data, i.e. one observation every 0.16 seconds, would have implied a computational burden which goes beyond the limits the ICT resources available to us. We chose not to use the modal action in a second, because the mode does not necessarily preserve the temporal sequence of actions, which is crucial for strategy estimation. As a robustness check we repeated the estimation taking one action every 2, and every 3 seconds. Our main results are robust to these changes. We thank Guillaume Fréchette for kindly providing the original code for strategy estimation.

¹⁹We refer the reader to the on-line appendix of Dal Bó and Fréchette (2011) for a more detailed explanation of the estimation approach.

the maximum likelihood estimation over four blocks of supergames: 1-6, 7-12, 13-18, 19-23 (Figure 2).

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

Support for Result 3 comes from Figure 2 and Table 5: the share of subjects adopting Fully defective strategies (B) is initially high (between 25% and 41%) and sharply decreases with experience. The decrease, however, is slower for the Stochastic treatments, and in particular for the Short-Stochastic one. When considering only experienced play (supergames 19-23, Table 5) the estimated frequency of Fully defective strategies is substantially higher in the Short-Stochastic than in the other treatments (15% vs. 1%-7%).

Table 5 also reveals that, in the last 5 supergames, Grim trigger is by far the most frequent strategy in the Deterministic treatments (50% in the Long-Deterministic, and 61% in the Short-Deterministic, Table 5). As previously mentioned, this probably indicates that in these treatments most subjects play Cut-Off strategies, which is in line with the analysis in Friedman and Oprea (2012).

Taken together, these results suggest that the difference in cooperation rates across treatments is to be found in a different dynamics of behavior, in particular in the relatively slow reduction in the share of defective strategies in the Short-Stochastic treatment. In the next section we put forward a reinforcement learning model that can explain why the take-off of cooperation is slower in the Short-Stochastic treatment than in the other three treatments.²⁰

²⁰Contrary to what Figure 2 might suggest, initial and average cooperation rates in the first supergame are higher in the Short-Stochastic treatment (60.4% and 40.7%, respectively) than in the other three treatments (50% and 25.4% in the Long-Deterministic, 54.2% and 35.4% in the Long-Stochastic, 52.1% and 24% in the Short-Deterministic). Hence, we believe that the higher prevalence of defective strategies estimated for the first block of six supergames in the Short-Stochastic treatment is determined by how play evolves with experience, rather than by the initial inclinations of the subjects.

Strategy	Long-Deterministic		Long-Stochastic		Short-Deterministic		Short-Stochastic	
	coeff.	s.d.	coeff.	s.d.	coeff.	s.d.	coeff.	s.d.
A. Fully cooperative strategies								
Always cooperate	0.963		0.888		0.987		0.814	
<i>Conditionally cooperative strategies (A1 + A2)</i>	0.043**	0.025	0.079*	0.061	0.000	0.000	0.026*	0.020
<i>A1. Forgiving cooperative strategies</i>	<i>0.920</i>		<i>0.809</i>		<i>0.987</i>		<i>0.788</i>	
Tit-for-tat	<i>0.338</i>		<i>0.659</i>		<i>0.247</i>		<i>0.282</i>	
win-stay-lose-shift	0.119***	0.038	0.203***	0.041	0.126*	0.084	0.000	0.024
T2	0.007	0.017	0.034	0.039	0.000	0.011	0.000	0.006
Tit-for-2-tats	0.068***	0.029	0.049*	0.038	0.063**	0.036	0.039**	0.023
Tit-for-3-tats	0.007	0.014	0.105**	0.056	0.015	0.019	0.041***	0.016
2-Tits-for-tat	0.042*	0.028	0.064**	0.029	0.000	0.006	0.027*	0.019
2-tits-for-2-tats	0.037*	0.026	0.047	0.048	0.042	0.040	0.087**	0.046
Tit-for-tat-2	0.060**	0.032	0.126**	0.068	0.000	0.024	0.070	0.068
<i>A2. Unforgiving cooperative strategies</i>	0.000	0.018	0.031**	0.016	0.000	0.000	0.018	0.017
Grim trigger	<i>0.582</i>		<i>0.150</i>		<i>0.740</i>		<i>0.505</i>	
Lenient grim 2	0.495***	0.065	0.069**	0.030	0.609***	0.158	0.350**	0.211
Lenient grim 3	0.036**	0.020	0.041	0.047	0.105***	0.034	0.105***	0.042
B. Fully defective strategies	0.051***	0.018	0.040***	0.016	0.026**	0.016	0.050**	0.027
Always defect	0.034		0.071		0.005		0.152	
Suspicious Tit-for-tat	0.027	0.025	0.043***	0.016	0.000	0.006	0.075***	0.021
C. Other strategies	0.008	0.009	0.028**	0.014	0.005	0.012	0.077	0.075
False cooperator	0.002		0.041		0.008		0.034	
Suspicious Tit-for-2-tats	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Suspicious Tit-for-3-tats	0.000	0.000	0.021	0.033	0.000	0.002	0.000	0.000
Suspicious lenient grim2	0.000	0.000	0.020	0.020	0.008	0.008	0.034	0.029
Suspicious lenient grim3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
D-alternator	0.002	0.006	0.000	0.000	0.000	0.000	0.000	0.000
gamma	0.000		0.000		0.000		0.000	
beta	0.308***	0.022	0.348***	0.012	0.358***	0.028	0.335***	0.035
	0.962		0.946		0.942		0.952	

Notes: the table reports the estimated frequency of each strategy in the population. In bold, we present the aggregate frequency for relevant families of strategies. The estimation methodology is described in the main text, while Table A.3 in Appendix A presents a description of the strategies.

Table 5: Estimated strategies in supergames 19-23

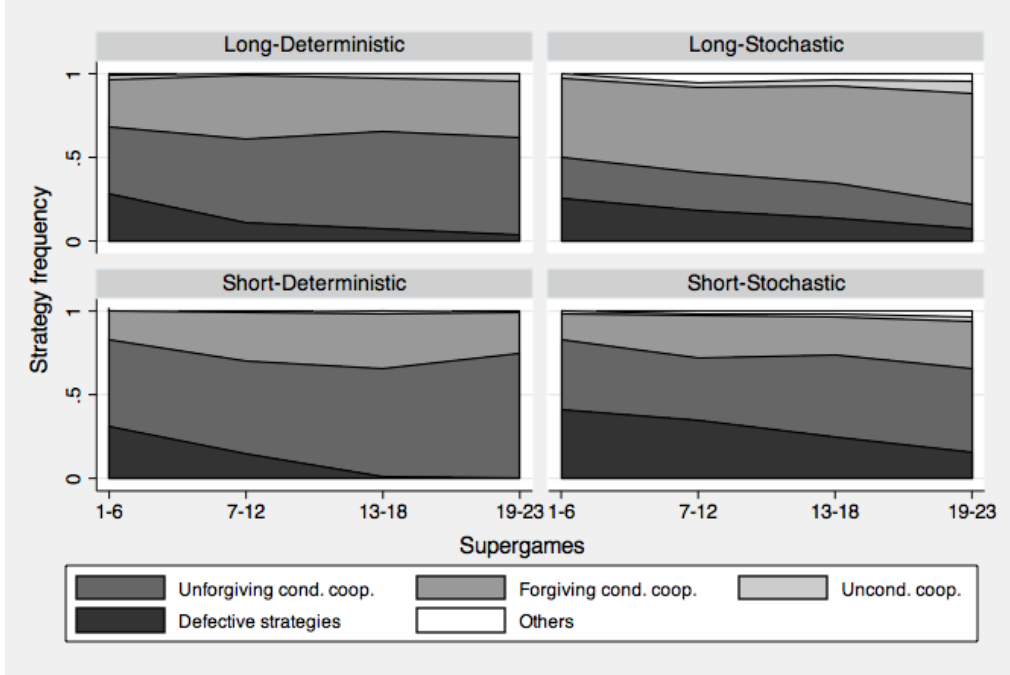


Figure 2: Evolution of strategies across supergames.

6 Dynamics across supergames

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

Results in Section 5 reveal that the vast majority of subjects played either Always defect or one of the conditionally cooperative strategies (Result 3). The empirical approach adopted there estimates the prevalence of each strategy in the population; it does not associate a specific individual

to a strategy. As a consequence, it does not allow to assess the empirical profitability of each strategy. Such an assessment is crucial to estimate a reinforcement learning model which can explain the evolution of strategy adoption. To this purpose, here we employ the following simple approach. We classify all individuals into two non-overlapping categories: “cooperators” and “defectors.” The former category includes all subjects who start cooperating, and never defect before their opponents or defect only during the last fifth of the (expected) duration of the supergame.²¹ The latter category includes all other subjects. This classification is broadly consistent with the estimates in Table 5: in supergames 19-23 cooperators account for 73%-90% of the population, depending on the treatment.²²

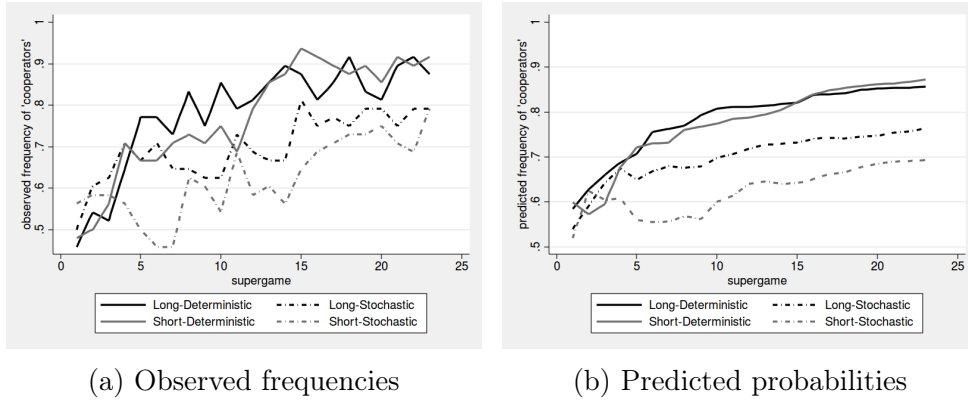


Figure 3: Frequency of “cooperators”

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

Support for Result 4 comes from Figure 3a, which reports the frequency of “cooperators.” An upward trend emerges in all treatments, but this trend is weakest in the Short-Stochastic treatment.²³ Our results show that, when

²¹The category of “cooperators” hence comprises also subjects adopting a Cut-Off strategy. A finer partition of the “cooperators” category is empirically troublesome because the same individual might be classified under multiple strategies. As Camera et al. (2012) show, cooperative strategies are frequently observationally equivalent.

²²In Bigoni et al. (2013), we present a similar reinforcement learning model based on initial actions, rather than on this strategy classification. Results are qualitatively the same.

²³A similar trend characterizes initial and average cooperation rates (Figure B.2 in the on-line appendix B.)

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

We now show that the observed dynamics of strategy adoption across supergames are consistent with a reinforcement learning model. Along the lines of Mookherjee and Sopher (1994, eq.2.1), we estimate the probability p_i^t that subject i is a “cooperator” in supergame t in the following model:

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

where $f(\cdot)$ denotes the logistic function; $A_{i,C}^{t-1}$ is the average payoff per second obtained by subject i up to supergame $t - 1$ when adopting a cooperative strategy; $A_{i,D}^{t-1}$ is the average payoff per second obtained by subject i up to supergame $t - 1$ when adopting a defective strategy;²⁴ \mathbf{x}_i is a vector of three dummy variables corresponding to treatments Short-Deterministic, Long-Stochastic, and Short-Stochastic; and \mathbf{z}_i is a vector of individual characteristics.²⁵

Results estimating the model in equation (1) via a logit regression are presented in Table 6 (Model 1). Figure 3b presents the predicted probability of adopting a cooperative strategy across treatments and supergames, estimated through Model 1. This reinforcement learning model fits well our data, the estimated coefficients ϕ and χ have the expected sign, and are both highly significant.

Table 6 (Model 2) also presents results from an alternative model, where the only independent variables are the individual characteristics and the treatment dummies ($\phi = \chi = 0$). The comparison between the two models in Table 6 reveals that the observed differences in the rates of adoption of cooperative strategies across treatments are almost fully captured by the reinforcement learning terms ($A_{i,C}^{t-1}$ and $A_{i,D}^{t-1}$), whose inclusion into the model substantially improves the estimated Log-likelihood.²⁶

²⁴Variables $A_{i,C}^0$ and $A_{i,D}^0$ are set equal to 0.

²⁵To allow for correlation between observations belonging to the same subject, this vector also includes a dummy taking value one if in the first supergame the subject adopted a cooperative strategy, zero otherwise. We thank an anonymous referee for the suggestion.

²⁶In Model 1 there is no statistically significant difference between the coefficients of the Short-Stochastic and Short-Deterministic treatment dummies (p-value > 0.1). In

Dependent variable: subject's strategy (1= cooperator, 0= defector)				
	Model 1		Model 2	
	Coefficient	(s.e.)	Coefficient	(s.e.)
$A_{i,C}^{t-1}$	3.043***	(0.262)		
$A_{i,D}^{t-1}$	-0.753**	(0.311)		
Short-Deterministic	-0.000	(0.315)	-0.419	(0.350)
Long-Stochastic	-0.072	(0.275)	-0.645**	(0.286)
Short-Stochastic	-0.383	(0.463)	-1.599***	(0.430)
Constant	-1.670*	(1.008)	0.332	(0.667)
Controls for individual characteristics	Yes		Yes	
N	4416		4416	
Log-likelihood	-1859.435		-2081.673	

Notes: panel logit regression with random effects at the subjects' level and standard errors robust for heteroschedasticity. Default treatment: Long-Deterministic. Standard errors are reported in parentheses. The unit of obs. is a subject's strategy (1 = "cooperator", 0 = "defector") in a supergame. See footnote 11 for a list of controls of individual characteristics. For a full version of this regression, see Table B.8 in the on-line appendix.

Table 6: Panel regression on subjects' strategy adoption.

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

Consider for simplicity a situation where all subjects play – with no errors – either Always Defect or one of the conditionally cooperative strategies (described in Table 5). No matter what is the specific strategy, when two conditional cooperators meet, they cooperate at all points in time, while if a conditional cooperator meets a defector, he will cooperate at the beginning, and eventually switch to permanent defection.

There exist both empirical and theoretical reasons to consider a restriction to these two types of strategies meaningful. On the empirical side, the analysis of individual strategies in Section 5 shows that for all levels of experience and in all treatments, the combined frequency of these strategies is above 80% (Result 3). On the theoretical side, justifications for restricting the focus to this class of strategies may come from the model of Jehiel (2005), or from the one of Neyman (1985). Albeit different in the underlying assumptions, the logic behind these two models can be applied to high

Model 2, this difference is highly significant (p-value < 0.001).

frequency/continuous time games and can lead to very similar predictions. Jehiel (2005) shows that conditional cooperation can be an equilibrium strategy if strategies depend on a coarse set partition of all possible histories. For example, consider the following partition: set A contains all histories in which opponent never defected (including the empty history) and set B all other histories. If both players use such two coarse analogies, the following strategies form equilibria for long games: i) both players play always defect and ii) both players follow a grim trigger, i.e., cooperate if history is in set A and defect otherwise. Neyman (1985), instead, shows that a cooperative equilibrium exists in finitely repeated prisoner's dilemmas, if subjects are bound to adopt strategies of limited complexity. More specifically, if players are restricted to use two-state automata (for instance Always Defect and Grim Trigger strategies), then they can sustain a cooperative equilibrium (unless the game is very short).

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

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

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

where T represents the (expected) total duration of the supergame, and $Ei(-\frac{\kappa}{T})$ is the exponential integral function.²⁷

²⁷For the stochastic treatments, this profit is equal to 2 in all supergames that last less than κ , which have probability $\int_0^\kappa \lambda e^{-x\lambda} dx$, where $\lambda = \frac{1}{T}$ is the parameter of the geometric distribution from which supergames' duration is drawn. For supergames whose realized duration τ is longer than κ , instead, the per-second profit from playing Always Defect against a conditionally cooperative strategy is $\frac{2\kappa}{\tau}$. Hence, the expected per-second profit from defection is $2 \int_0^\kappa \lambda e^{-x\lambda} dx + \int_\kappa^\infty \frac{2\kappa}{x} \lambda e^{-x\lambda} dx = 2(1 - e^{-\frac{\kappa}{T}}) + \frac{2\kappa}{T} (-Ei(-\frac{\kappa}{T}))$. The "sucker's" payoff from playing a conditionally cooperative strategy against Always

<i>Strategy of the column player/ Strategy of the row player</i>	Always defect	Conditional cooperation
Always defect	0, 0	$\bar{\pi}(\kappa, H), -\bar{\pi}(\kappa, H)$
Conditional cooperation	$-\bar{\pi}(\kappa, H), \bar{\pi}(\kappa, H)$	1, 1

Table 7: Matrix of the expected profits per second.

The length of κ depends on the leniency of the specific conditional strategy adopted by the cooperator (it can be chosen strategically or simply driven by reaction time). We computed an empirical measure of κ , by considering those pairs in the experiment which initiated a supergame with one subject cooperating and the other defecting, and then directly transitioned to mutual defection (i.e. the cooperator switched to defection before the end of the supergame). The average duration of this phase was 1.75 seconds in the Long-Deterministic, 1.51 seconds in the Long-Stochastic, 1.40 seconds in the Short-Deterministic, and 1.61 seconds in the Short-Stochastic treatment. Plugging the empirical measures of κ into equations (2) and (3) yields the payoff matrix presented in Table 7, where $\bar{\pi}(\kappa, d) = 0.06$ in the Long-Deterministic treatment, $\bar{\pi}(\kappa, d) = 0.14$ in the Short-Deterministic, $\bar{\pi}(\kappa, s) = 0.21$ in the Long-Stochastic, and $\bar{\pi}(\kappa, s) = 0.48$ in the Short-Stochastic.

For these values of $\bar{\pi}(\kappa, H)$, the set of beliefs that make a conditionally cooperative strategy optimal (i.e. its “basin of attraction”, see Myerson, 1991, chapter 7.11) is much smaller and the riskiness of cooperating (Blonski and Spagnolo, forth.) much larger in the Short-Stochastic than in the other three treatments.²⁸ More generally, the ordering across treatments following a basin-of-attraction criterion is consistent with the pattern of the frequency of “cooperators” depicted in Figure 3.

7 Additional empirical evidence

The discussion in the previous section showed that, in an environment in which cooperators and defectors coexist, cooperating is more profitable

defect is simply given by $-\bar{\pi}(\kappa, H)$.

²⁸Dal Bó and Fréchette (2011) and Blonski et al. (2011) showed that these concepts capture well coordination problems in discretely repeated games experiments.

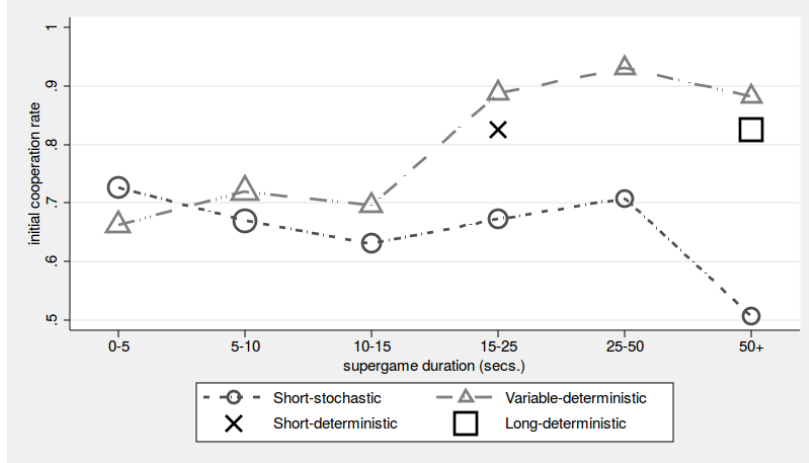
than defecting only in supergames that are sufficiently long. Our learning model suggests that subjects tend to cooperate more when they have experienced past cooperation to be profitable, which happens less frequently in the Short-Stochastic treatment, where the incidence of very short supergames is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

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

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

Initial and mean cooperation rates are higher in the Variable-Deterministic than in the Short-Stochastic treatment (77.9 vs. 65.9 and 57.1 vs. 52.3, respectively).²⁹ A deterministic horizon facilitates cooperation relative to a

²⁹The difference is significant for both comparisons (p-value < 0.05 in both cases, according to a linear regression with one observation per subject, and bootstrapped standard errors. Regression results are reported in Table A.2 in Appendix, and in Tables B.9 and B.10 in the on-line appendix B. Figure B.3 in the on-line appendix presents the time profile of the mean share of cooperators, taken across supergames. The five lines correspond to supergames of different duration.



Notes: The size of line markers are proportional to the number of observations.

Figure 4: Rate of initial cooperation, by supergames' duration

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

When the horizon is deterministic, supergames' duration matters for cooperation. In the experiment, initial cooperation rates monotonically increase with supergames' duration (Figure 4).³⁰ As a benchmark, Figure 4 also presents the initial cooperation rate in the Short-Stochastic treatment, where subjects cannot condition behavior on the actual supergames' duration.³¹

The data from the Variable-Deterministic treatment suggest that both the variability and the unpredictability of supergames' durations had an impact on cooperation, although the unpredictability of the stochastic horizon was the stronger force (Result 5). This last finding is in line with Bereby-Meyer and Roth's (2006) result that introducing randomness in strategic

³⁰There is a kink around 10-15 seconds, which makes the relation non-linear. 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 supergames they defect from the start.

³¹By contrast, we observe that in the Short-Stochastic treatment, the realized duration of the *previous* supergame weakly affects subjects' decision whether to cooperate in the subsequent one, which is in line with Dal Bó and Fréchette (2011). In the Variable-Deterministic treatment, instead, this is not the case (Table B.11 in the on-line appendix B). One possible interpretation of this is that when the horizon is stochastic subjects update their beliefs about the expected value of cooperation. The expected value depends not only on what they think others do, but also on the expected length of the supergame: the longer the supergame, the higher the gains from cooperation. When the horizon is deterministic, there is no need for such an update.

situations, while holding expected payoffs constant, hinders subjects ability to adapt.

8 Discussion

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

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

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

Theories in the second set perturb the repeated game by introducing some incomplete information. They include Kreps et al. (1982) and Radner (1986) for discretely-repeated games, and Friedman and Oprea (2012) and Park (2013) for continuous-time ones. In these models, players are not

certain about the behavior or preferences of their opponents. For example, in Radner (1986) and Friedman and Oprea (2012), players assign some probability to their opponent playing a cutoff strategy: play grim-trigger till some time t and then defect. In Kreps et al. (1982), players assign positive probability to their opponents playing a grim trigger strategy and in Park (2013) there is additionally private information about reaction times. These models predict equilibria (often unique) with cooperation at the beginning of the game and defection close to the end.

Overall, both sets of theories admit high cooperation rates as one of the equilibrium outcomes in continuous-time or frequently repeated games, no matter what the termination rule. Moreover, both sets of theories offer no reason for unraveling in deterministic horizon games. In these dimensions our results support these models in contrast to unperturbed discrete time models with unrestricted strategies that got a stronger support from discrete-time experiments.

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

Some of our results, however, are hard to reconcile with any of these theories. First, none of these equilibrium theories predicts higher cooperation levels in the deterministic than stochastic treatments. Second, they all seem to predict that initial cooperation rates should be independent of the termination rule or duration of the game. Our results contradict both

these predictions. Our best explanation of the data is that the observed patterns are an effect of differences in the evolution of behavior with experience rather than differences in the set of equilibria in these games. In particular, in Section 6 we have shown that a model based on reinforcement learning captures well the differences across treatments in the growth of cooperation.

9 Conclusions

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

Two main economic forces influence whether agents can successfully cooperate in repeated games with perfect monitoring: incentives, i.e. the tradeoff between immediate gains and the shadow of future punishments; and coordination, i.e. the ability to reach a cooperative equilibrium. These forces are relevant both in theory and in the experimental evidence. Our results suggest that – in situations where agents can react quickly to moves by others – the first force is second-order. Instead, the second, the ability to make use of experience for coordinating on future play appears to be a first-order determinant of cooperation levels.

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

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

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 supergames' 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).

A Appendix - for publication

Dependent variable: average timing of the end-game effect		
	Coefficient	(s.e.)
Short-Deterministic	-7.182*	(3.718)
Supergame	-1.178***	(0.342)
Supergame ²	0.027**	(0.014)
Supergame \times Short-Deterministic	0.609	(0.479)
Supergame ² \times Short-Deterministic	-0.012	(0.019)
Reaction time	4.135***	(1.153)
Reaction time \times Short-Deterministic	-2.639	(1.775)
Constant	14.956***	(2.876)
N	92	
R-squared overall	0.680	
R-squared between	0.933	
R-squared within	0.514	

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

Table A.1: Panel regression on the average timing of the end-game effect

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

Notes: Linear regressions with bootstrapped standard errors. To take care of the potential correlation between observations coming from the same session, we take sessions as resampling clusters. The unit of observation is the mean (initial) cooperation rate in a section in a supergame. Default treatment: Short-Stochastic. Standard errors are reported in parentheses. A full version of this regression, including the estimated coefficients for all the controls, is reported in Table B.9 in the on-line appendix B.

Table A.2: Linear regression on mean and initial cooperation rates

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

Table A.3: Set of strategies used in the maximum likelihood estimation.

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.
- AOYAGI, M. AND G. FRÉCHETTE (2009): “Collusion as public monitoring becomes noisy: Experimental evidence,” *Journal of Economic Theory*, 144, 1135 – 1165.
- 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.
- BIGONI, M., M. CASARI, A. SKRZYPACZ, AND G. SPAGNOLO (2013): “Time Horizon and Cooperation in Continuous Time,” Mimeo.
- BLONSKI, M., P. OCKENFELS, AND G. SPAGNOLO (2011): “Equilibrium Selection in the Repeated Prisoner’s Dilemma: Axiomatic Approach and Experimental Evidence,” *American Economic Journal: Microeconomics*, 3, 164–192.
- BLONSKI, M. AND G. SPAGNOLO (Forthcoming): “Prisoners’ other Dilemma,” *International Journal of Game Theory*, 1–21.
- CAMERA, G., M. CASARI, AND M. BIGONI (2012): “Cooperative strategies in anonymous economies: An experiment,” *Games and Economic Behavior*, 75, 570 – 586.

- 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.
- DAL BÓ, P. AND G. R. FRÉCHETTE (2012): “Strategy Choice in the Infinitely Repeated Prisoners Dilemma,” Mimeo.
- 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.
- EMBREY, M., G. FRÉCHETTE, AND S. YÜKSEL (2013): “Backward Induction in the Finitely Repeated Prisoner’s Dilemma: Experimental Evidence,” Mimeo.
- 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.
- FRÉCHETTE, G. R. AND S. YUKSEL (2013): “Infinitely Repeated Games in the Laboratory: Four Perspectives on Discounting and Random Termination,” Available at SSRN: <http://ssrn.com/abstract=2225331>.
- FRIEDMAN, D., S. HUCK, R. OPREA, AND S. WEIDENHOLZER (2012): “From imitation to collusion: Long-run learning in a low-information en-

- vironment,” Discussion Paper, Wissenschaftszentrum Berlin für Sozialforschung, Abteilung Ökonomik des Wandels, Schwerpunkt Märkte und Politik.
- FRIEDMAN, D. AND R. OPREA (2012): “A Continuous Dilemma,” *American Economic Review*, 102, 337–63.
- FUDENBERG, D., D. G. RAND, AND A. DREBER (2012): “Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World,” *American Economic Review*, 102, 720–749.
- 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.
- JEHIEL, P. (2005): “Analogy-based expectation equilibrium,” *Journal of Economic Theory*, 123, 81 – 104.
- 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.
- MOOKHERJEE, D. AND B. SOPHER (1994): “Learning behavior in an experimental matching pennies game,” *Games and Economic Behavior*, 7, 62–91.
- 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.
- MYERSON, R. B. (1991): *Game Theory: Analysis of Conflict*, Harvard University Press.
- NEYMAN, A. (1985): “Bounded Complexity Justifies Cooperation In The Finitely Repeated Prisoners’ Dilemma,” *Economics Letters*, 19, 227–229.

- 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.
- PARK, I.-U. (2013): “Cooperation in Continuous Dilemma by Privately Informed Players,” Mimeo.
- RADNER, R. (1986): “Can Bounded Rationality Resolve at the Prisoner’s Dilemma?” in *Contributions to Mathematical Economics*, ed. by A. Mas-Colell and W. Hildenbrand, North-Holland, Amsterdam, 387–399.
- RAND, D., D. FUDENBERG, AND A. DREBER (2013): “It’s the Thought that Counts: The Role of Intentions in Reciprocal Altruism,” Available at SSRN 2259407.
- 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, L. K. AND M. B. STINCHCOMBE (1989): “Extensive Form Games in Continuous Time: Pure Strategies,” *Econometrica*, 57, 1171–1214.
- SZYMANSKI, S. (1996): “Making Hay While the Sun Shines,” *The Journal of Industrial Economics*, 44, 1–16.
- VESPA, E. (2013): “Cooperation in Dynamic Games: An Experimental Investigation,” Mimeo.

Online appendix

B Additional tables and figures

	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 B.1: Treatments and sessions

Dependent variable: cooperation rate				
	Model 1		Model 2	
	Coefficient	(s.e.)	Coefficient	(s.e.)
Short-Deterministic	-5.537**	(2.760)	-5.537**	(2.801)
Long-Stochastic	1.682	(2.577)	1.682	(3.565)
Short-Stochastic	-17.799***	(3.830)	-17.799***	(3.769)
<i>Demographics</i>				
Age	-0.140	(0.373)	-0.140	(0.247)
Male	3.201	(3.188)	3.201	(2.307)
Italian	3.916	(3.711)	3.916	(4.594)
Education	1.140	(1.346)	1.140	(1.656)
Missing edu.	16.066***	(4.516)	16.066***	(6.226)
Non-student	1.239	(5.807)	1.239	(5.252)
<i>Field of study</i>				
Medicine, Sciences, Engineering	6.416*	(3.663)	6.416	(4.363)
Humanities	5.966	(5.604)	5.966	(5.114)
Other	2.452	(5.298)	2.452	(5.509)
<i>Task comprehension</i>				
Wrong answers	-0.074	(0.354)	-0.074	(0.256)
Response time	-0.010	(0.013)	-0.010	(0.013)
<i>Academic background</i>				
Economic classes	-2.841	(3.639)	-2.841	(3.060)
Statistics classes	0.454	(2.406)	0.454	(2.791)
Game theory classes	2.759	(4.962)	2.759	(3.362)
<i>Questionnaire measures</i>				
Risk attitudes	-0.834	(0.922)	-0.834	(0.621)
Generalized trust	-2.206	(1.718)	-2.206	(2.711)
Missing trust	1.177	(3.514)	1.177	(3.282)
IQ quiz 1	-0.163	(3.498)	-0.163	(4.133)
IQ quiz 2	6.626***	(1.871)	6.626***	(2.226)
Constant	63.798***	(14.128)	63.798***	(10.193)
N	192		4416	
R-squared	0.220			
R-squared overall			0.046	
R-squared between			0.220	
R-squared within			0.000	

Notes: Model 1 presents the full estimate of the regression in Table 3, while Model 2 presents a 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 supergame. Defaults: Long-Deterministic treatment, major in Economics and Business. The difference between coefficients for the Short-Stochastic and Short-Deterministic treatments is significant at any standard significance level (p-value < 0.001).

Table B.2: Regressions on cooperation rates

Duration	Termination rule	
	Deterministic	Stochastic
Long	0.613	~
	(0.791)	(0.769)
	√***	√***
Short	0.557	>***
	(0.700)	(0.283)

Notes: The mean rate of mutual cooperation of a session is the average across all 23 supergames and all 12 groups in each supergame. Median rates of mutual cooperation are reported in parentheses. Significance levels are derived from the regression presented in Model 1 of Table B.4.

Table B.3: 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	-0.086***	(0.026)	-0.086***	(0.028)
Long-Stochastic	0.002	(0.024)	0.002	(0.036)
Short-Stochastic	-0.207***	(0.045)	-0.207***	(0.038)
<i>Demographics</i>				
Age	-0.002	(0.004)	-0.002	(0.003)
Male	0.033	(0.033)	0.033	(0.023)
Italian	0.042	(0.036)	0.042	(0.045)
Education	0.015	(0.014)	0.015	(0.017)
Missing edu.	0.185***	(0.041)	0.185***	(0.060)
Non-student	-0.001	(0.057)	-0.001	(0.054)
<i>Field of study</i>				
Medicine, Sciences, Engineering	0.049	(0.036)	0.049	(0.042)
Humanities	0.045	(0.057)	0.045	(0.049)
Other	0.021	(0.055)	0.021	(0.055)
<i>Task comprehension</i>				
Wrong answers	-0.001	(0.004)	-0.001	(0.003)
Response time	-0.000	(0.000)	-0.000	(0.000)
<i>Academic background</i>				
Economic classes	-0.017	(0.034)	-0.017	(0.030)
Statistics classes	-0.003	(0.024)	-0.003	(0.028)
Game theory classes	0.023	(0.051)	0.023	(0.033)
<i>Questionnaire measures</i>				
Risk attitudes	-0.009	(0.009)	-0.009	(0.006)
Generalized trust	-0.027	(0.020)	-0.027	(0.028)
Missing trust	0.010	(0.034)	0.010	(0.030)
IQ quiz 1	0.001	(0.035)	0.001	(0.040)
IQ quiz 2	0.069***	(0.019)	0.069***	(0.022)
Constant	0.619***	(0.154)	0.619***	(0.101)
N	192		4416	
R-squared	0.254			
R-squared overall			0.047	
R-squared between			0.254	
R-squared within			0.000	

Notes: Model 1 presents results from a linear regression with bootstrapped standard errors. To take care of the potential correlation between observations coming from the same session, we take sessions as resampling clusters. The unit of observation is the average fraction of time a pair of subjects coordinate on cooperation within a supergame, across all supergames. 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 supergame's duration in which both subjects in a pair cooperate. Defaults: Long-Deterministic treatment, major in Economics and Business.

Table B.4: Linear regression on rates of mutual cooperation

Duration	Termination rule	
	Deterministic	Stochastic
Long	0.614 (0.792) √**	~ 0.608 (0.793) √***
Short	0.570 (0.696)	>*** 0.447 (0.455)

Notes: The mean profit per second of a session is the average across all 23 supergames and all 24 subjects. Median profits are reported in parentheses. Significance levels are derived from the regression presented in Model 1 of Table B.6.

Table B.5: Average profits per second

Dependent variable: average profit per second				
	Model 1		Model 2	
	Coefficient	(s.e.)	Coefficient	(s.e.)
Short-Deterministic	-0.060**	(0.026)	-0.060**	(0.025)
Long-Stochastic	0.013	(0.015)	0.013	(0.029)
Short-Stochastic	-0.189***	(0.041)	-0.189***	(0.030)
<i>Demographics</i>				
Age	-0.004*	(0.003)	-0.004*	(0.003)
Male	0.021	(0.023)	0.021	(0.019)
Italian	0.026	(0.032)	0.026	(0.035)
Education	0.012	(0.009)	0.012	(0.014)
Missing edu.	0.135***	(0.037)	0.135***	(0.048)
Non-student	0.007	(0.049)	0.007	(0.047)
<i>Field of study</i>				
Medicine, Sciences, Engineering	0.021	(0.027)	0.021	(0.030)
Humanities	0.009	(0.042)	0.009	(0.036)
Other	0.002	(0.048)	0.002	(0.043)
<i>Task comprehension</i>				
Wrong answers	-0.000	(0.003)	-0.000	(0.002)
Response time	-0.000	(0.000)	-0.000*	(0.000)
<i>Academic background</i>				
Economic classes	0.005	(0.019)	0.005	(0.022)
Statistics classes	-0.008	(0.021)	-0.008	(0.022)
Game theory classes	0.017	(0.038)	0.017	(0.026)
<i>Questionnaire measures</i>				
Risk attitudes	-0.006	(0.006)	-0.006	(0.005)
Generalized trust	-0.035**	(0.017)	-0.035	(0.023)
Missing trust	-0.011	(0.026)	-0.011	(0.031)
IQ quiz 1	0.013	(0.027)	0.013	(0.029)
IQ quiz 2	0.054***	(0.017)	0.054***	(0.018)
Constant	0.698***	(0.124)	0.698***	(0.088)
N	192		4416	
R-squared	0.344			
R-squared overall			0.034	
R-squared between			0.344	
R-squared within			0.000	

Notes: Model 1 presents results from a linear regression with bootstrapped standard errors. To take care of the potential correlation between observations coming from the same session, we take sessions as resampling clusters. The unit of observation is the average profit per second, across all supergames. 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 supergame. Defaults: Long-Deterministic treatment, major in Economics and Business.

Table B.6: Linear regression on average profits per second

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

Table B.7: Average reaction time across supergames.

Dependent variable: subjects' strategy (1= cooperator, 0= defector)				
	Model 1		Model 2	
	Coefficient	(s.e.)	Coefficient	(s.e.)
Short-Deterministic	-0.000	(0.315)	-0.419	(0.350)
Long-Stochastic	-0.072	(0.275)	-0.645**	(0.286)
Short-Stochastic	-0.383	(0.463)	-1.599***	(0.430)
$A_{i,C}^{t-1}$	3.043***	(0.262)		
$A_{i,D}^{t-1}$	-0.753**	(0.311)		
Cond. coop. supergame 1	1.517***	(0.201)	1.515***	(0.157)
<i>Demographics</i>				
Age	0.006	(0.019)	-0.028**	(0.013)
Male	0.136	(0.248)	0.423***	(0.149)
Italian	0.304	(0.443)	0.425*	(0.249)
Education	0.280**	(0.115)	0.203**	(0.096)
Missing edu.	1.179**	(0.555)	1.620***	(0.297)
Non-student	-0.062	(0.441)	-0.110	(0.250)
<i>Field of study</i>				
Medicine, Sciences, Engineer- ing	-0.045	(0.374)	0.372*	(0.214)
Humanities	0.587*	(0.302)	0.470	(0.286)
Other	0.292	(0.455)	0.065	(0.251)
<i>Task comprehension</i>				
Wrong answers	-0.024	(0.035)	-0.037***	(0.006)
Response time	0.000	(0.001)	0.001***	(0.000)
<i>Academic background</i>				
Economic classes	0.404*	(0.244)	0.240	(0.192)
Statistics classes	-0.006	(0.417)	-0.153	(0.201)
Game theory classes	-0.234	(0.165)	-0.199	(0.314)
<i>Questionnaire measures</i>				
Risk attitudes	-0.102*	(0.053)	-0.109***	(0.042)
Generalized trust	-0.331	(0.239)	-0.148	(0.221)
Missing trust	0.236	(0.252)	0.192	(0.198)
IQ quiz 1	0.236	(0.203)	0.498***	(0.162)
IQ quiz 2	0.699***	(0.190)	0.899***	(0.160)
Constant	-1.670*	(1.008)	0.332	(0.667)
N	4416		4416	
Log-likelihood	-1859.435		-2081.673	

Notes: panel logit regression with random effects at the subjects' level and standard errors robust for heteroschedasticity. Defaults: Long-Deterministic treatment, major in Economics and Business. Standard errors are reported in parentheses. The unit of obs. is a subject's strategy (1 = "cooperator", 0 = "defector"), in a supergame. See footnote 11 for a list of controls of individual characteristics.

Table B.8: Panel regression on subjects' strategies.

	Mean coop. rate		Initial coop. rate	
Short-Deterministic	11.574***	(2.359)	17.051***	(6.164)
Variable-Deterministic	6.966**	(3.094)	13.489**	(6.751)
<i>Demographics</i>				
Age	0.266	(0.453)	-0.192	(0.591)
Male	0.121	(3.091)	4.646***	(1.755)
Italian	6.637	(7.342)	8.536	(11.849)
Education	0.977	(1.325)	3.307	(2.304)
Missing edu.	16.371***	(4.432)	41.395***	(9.590)
Non-student	2.285	(6.308)	1.688	(9.312)
<i>Field of study</i>				
Medicine, Sciences, Engineering	2.537	(3.781)	5.486	(6.601)
Humanities	-1.831	(4.345)	-2.883	(8.032)
Other	-4.544	(5.749)	-4.458	(10.600)
<i>Task comprehension</i>				
Wrong answers	0.037	(0.569)	-0.205	(0.918)
Response time	-0.002	(0.023)	-0.006	(0.026)
<i>Academic background</i>				
Economic classes	-4.075	(2.707)	0.042	(4.312)
Statistics classes	3.470	(3.684)	-0.699	(5.364)
Game theory classes	0.037	(3.363)	-4.172	(4.542)
<i>Questionnaire measures</i>				
Risk attitudes	-0.715	(0.611)	-1.957***	(0.689)
Generalized trust	-0.599	(1.806)	-2.534	(3.338)
Missing trust	-1.432	(4.167)	-3.145	(9.401)
IQ quiz 1	-0.217	(2.929)	-3.332	(5.570)
IQ quiz 2	4.402*	(2.248)	5.801	(4.053)
Constant	39.488***	(13.287)	65.719***	(16.057)
N	144		144	
R-squared	0.155		0.216	

Notes: Linear regressions with bootstrapped standard errors. To take care of the potential correlation between observations coming from the same session, we take sessions as resampling clusters. The unit of observation is the mean (initial) cooperation rate in a section in a supergame. Defaults: Short-Stochastic treatment, major in Economics and Business. Standard errors are reported in parentheses.

Table B.9: Linear regression on mean and initial cooperation rates

	Mean coop. rate		Initial coop. rate	
Short-Deterministic	11.574***	(3.624)	1.239***	(0.259)
Variable-Deterministic	6.966*	(4.067)	0.807***	(0.265)
<i>Demographics</i>				
Age	0.266	(0.628)	0.036	(0.023)
Male	0.121	(2.492)	0.421***	(0.120)
Italian	6.637	(7.561)	0.086	(0.370)
Education	0.977	(2.411)	0.045	(0.144)
Missing edu.	16.371***	(6.199)		
Non-student	2.285	(8.096)	0.671*	(0.361)
<i>Field of study</i>				
Medicine, Sciences, Engineering	2.537	(4.621)	0.391	(0.445)
Humanities	-1.831	(5.221)	0.080	(0.446)
Other	-4.544	(7.435)	0.065	(0.304)
<i>Task comprehension</i>				
Wrong answers	0.037	(0.419)	-0.006	(0.010)
Response time	-0.002	(0.020)	-0.001	(0.001)
<i>Academic background</i>				
Economic classes	-4.075	(3.693)	-0.206	(0.171)
Statistics classes	3.470	(3.405)	0.284*	(0.151)
Game theory classes	0.037	(3.362)	0.205	(0.331)
<i>Questionnaire measures</i>				
Risk attitudes	-0.715	(0.674)	-0.181***	(0.035)
Generalized trust	-0.599	(3.159)	0.005	(0.156)
Missing trust	-1.432	(3.949)	-0.972***	(0.218)
IQ quiz 1	-0.217	(3.729)	-0.115	(0.351)
IQ quiz 2	4.402*	(2.582)	0.467***	(0.150)
Constant	39.488**	(16.209)	0.355	(0.637)
N	3312		3312	
R-squared overall	0.029			
R-squared between	0.155			
R-squared within	0.000			
Log-likelihood			-1524.5	

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 supergame. 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. Defaults: Short-Stochastic treatment, major in Economics and Business. Standard errors are reported in parentheses.

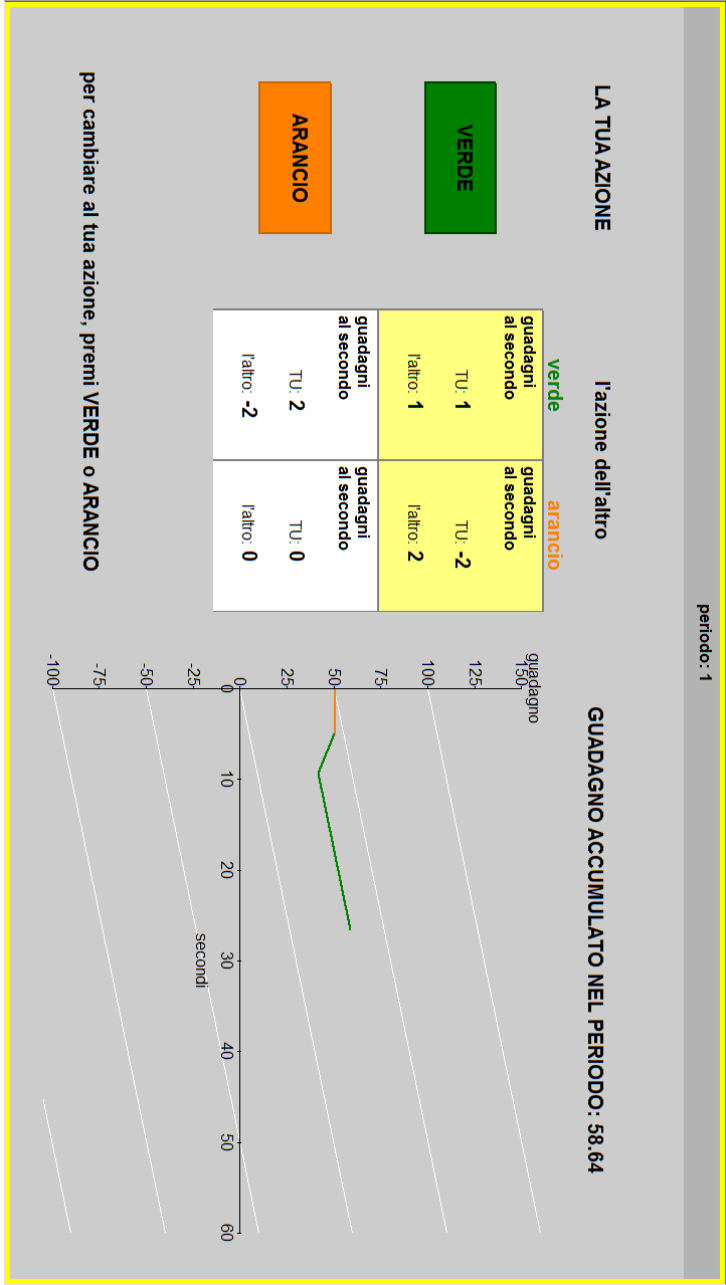
Table B.10: Panel regression on mean and initial cooperation rates

Dependent variable: initial cooperation				
	Short-Stochastic		Variable-Deterministic	
Initial coop. S1	0.498	(0.310)	1.360***	(0.209)
Duration _{t-1}	0.011*	(0.006)	0.007	(0.006)
Duration _{t-1} ²	-0.000	(0.000)	-0.000	(0.000)
Initial coop. _{t-1} ^j	0.452***	(0.100)	0.204*	(0.120)
<i>Demographics</i>				
Age	0.083	(0.068)	-0.038	(0.039)
Male	0.100	(0.314)	0.321	(0.210)
Italian	0.451	(1.103)	-0.780	(0.824)
Education	-0.203	(0.311)	0.252	(0.161)
Missing edu.	5.219	(176.581)		
Non-student	-0.761	(0.761)	0.054	(0.391)
<i>Field of study</i>				
Medicine, Sciences, Engineering	0.123	(0.649)	-0.172	(0.281)
Humanities	-1.064*	(0.640)	0.126	(0.361)
Other	-0.368	(0.919)	0.156	(0.289)
<i>Task comprehension</i>				
Wrong answers	0.026	(0.026)	0.050	(0.041)
Response time	-0.000	(0.001)	-0.000	(0.002)
<i>Academic background</i>				
Economic classes	-0.421	(0.373)	-0.415	(0.312)
Statistics classes	-0.441	(0.301)	0.351	(0.324)
Game theory classes	-0.291	(0.373)	0.160	(0.262)
<i>Questionnaire measures</i>				
Risk attitudes	0.084	(0.093)	0.092*	(0.051)
Generalized trust	0.295	(0.322)	0.382*	(0.207)
Missing trust	0.804	(0.495)	1.274*	(0.671)
IQ quiz 1	-0.696	(0.509)	0.391	(0.386)
IQ quiz 2	0.233	(0.279)	0.740***	(0.204)
Constant	-1.548	(1.572)	-0.854	(1.466)
N	1056		1056	
Log-likelihood	-504.157		-455.317	

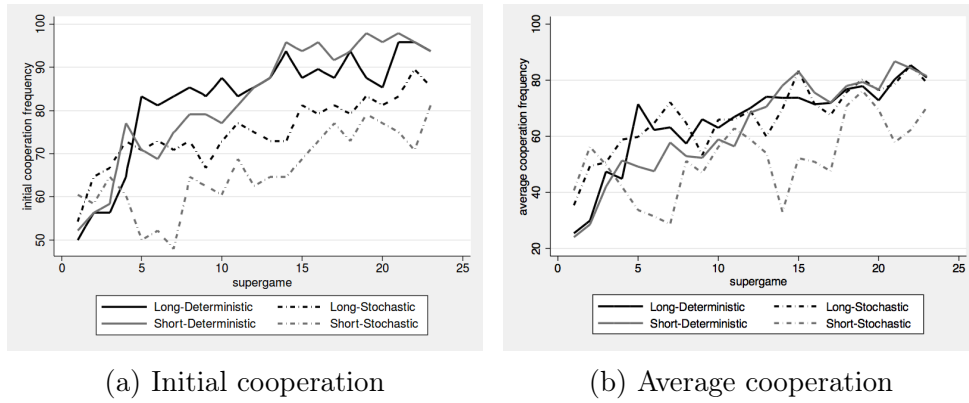
Notes: correlated random effects probit models with the initial decision as the dependent variable. Standard errors are reported in parentheses. The regressors we used are: initial decision of the partner in the previous match (*Initial coop.*_{t-1}^j), the duration of the previous supergame (*Duration*_{t-1}), the duration squared (*Duration*_{t-1}²), and the initial decision in the first supergame (*Initial coop. S1*). Default major: Economics and Business.

Table B.11: Linear regression on mean and initial cooperation rates

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

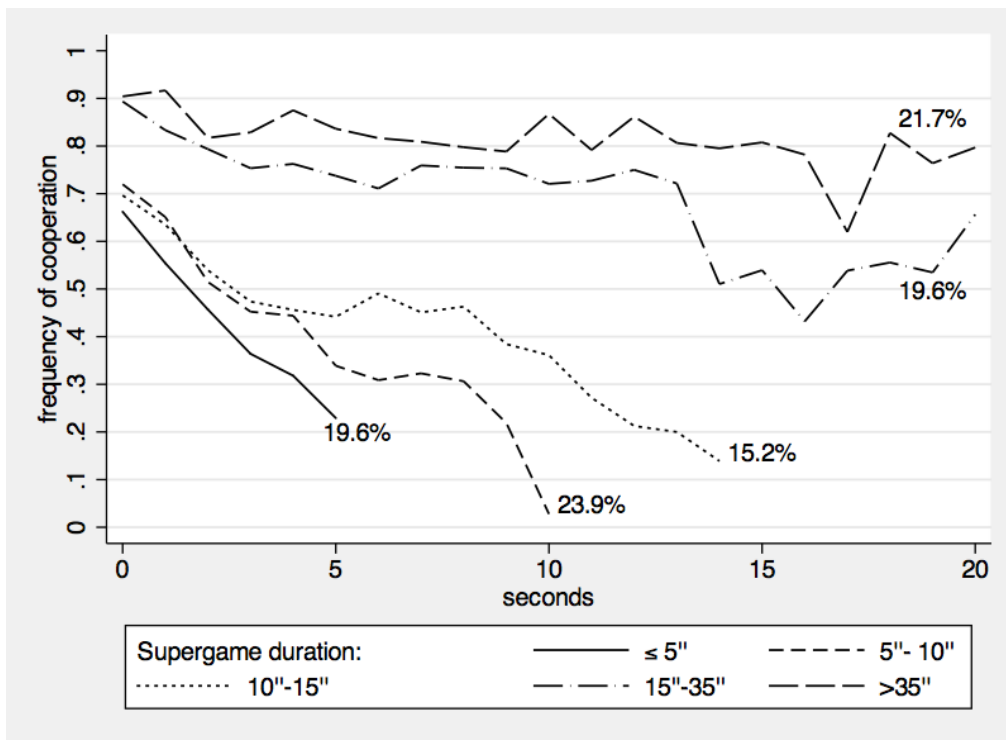


Notes: VERDE = green, ARANCIO = orange, l'azione dell'altro = your opponent's action, guadagno = earnings.



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

Figure B.2: Evolution of cooperation



Notes: The percentage next to each line represents the fraction of supergames having the corresponding duration. The unit of observation is the share of cooperative actions in every second of a supergame.

Figure B.3: Time profile of the share of cooperative actions, in the Variable-Deterministic treatment

C 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 C.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 C.1 above represent earnings per second.

LA TUA AZIONE	l'azione dell'altro	
	verde	arancio
VERDE	guadagni al secondo TU: 1 l'altro: 1	guadagni al secondo TU: -2 l'altro: 2
ARANCIO	guadagni al secondo TU: 2 l'altro: -2	guadagni al secondo TU: 0 l'altro: 0

Figure C.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 C.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.

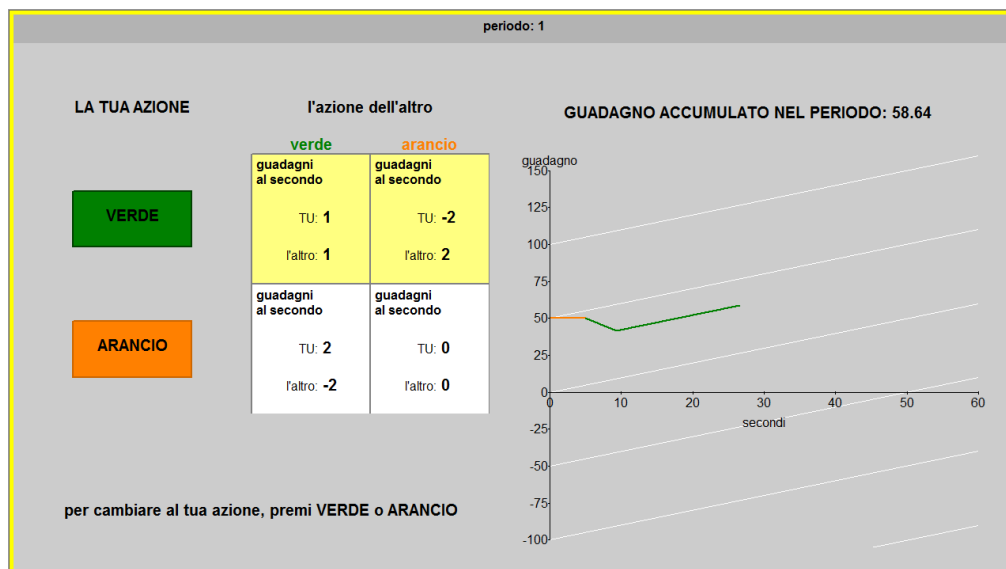


Figure C.2: Earnings table

Earnings

During the period you will receive information in real time on your earnings. In the screen pictured in Figure C.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.